pedocs
Open Access Erziehungswissenschaften

DIPF

Wilke, René [Hrsg.]; Knoblauch, Hubert [Hrsg.]

Videographie und Videoanalyse. Beiträge zur Erhebung, Analyse und Nutzung von Videodaten in der Qualitativen Forschung

Weinheim : Beltz Juventa 2025, 513 S. - (Standards standardisierter und nichtstandardisierter Sozialforschung)

Quellenangabe/ Reference:
Wilke, René [Hrsg.]; Knoblauch, Hubert [Hrsg.]: Videographie und Videoanalyse. Beiträge zur Erhebung, Analyse und Nutzung von Videodaten in der Qualitativen Forschung. Weinheim : Beltz Juventa 2025, 513 S. - (Standards standardisierter und nichtstandardisierter Sozialforschung) - URN: urn:nbn:de:0111-pedocs-337347 - DOI: 10.25656/01:33734

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-337347
https://doi.org/10.25656/01:33734

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

http://www.juventa.de

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: http://creativecommons.org/licenses/by/4.0/deed.de - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: http://creativecommons.org/licenses/by/4.0/deed.en - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der Leibniz-Gemeinschaft

Standards standardisierter und nichtstandardisierter Sozialforschung

René Wilke | Hubert Knoblauch (Hrsg.)

Videographie und Videoanalyse

Beiträge zur Erhebung, Analyse und Nutzung von Videodaten in der Qualitativen Forschung

BELTZ JUVENTA

René Wilke | Hubert Knoblauch (Hrsg.)
Videographie und Videoanalyse

Standards standardisierter und
nichtstandardisierter Sozialforschung

Herausgegeben von Nicole Burzan |
Paul Eisewicht | Ronald Hitzler

René Wilke | Hubert Knoblauch (Hrsg.)

Videographie und Videoanalyse

Beiträge zur Erhebung, Analyse und Nutzung von Videodaten in der Qualitativen Forschung

BELTZ JUVENTA

Dieser Sammelband, herausgegeben von René Wilke und Hubert Knoblauch (Technische Universität Berlin), wurde aus dem Open-Access-Publikationsfonds der Technischen Universität Berlin unterstützt.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Der Text dieser Publikation wird unter der Lizenz **Creative Commons Namensnennung 4.0 International (CC BY 4.0)** veröffentlicht. Den vollständigen Lizenztext finden Sie unter: https://creativecommons.org/licenses/by/4.0/deed.de. Verwertung, die den Rahmen der **CC BY 4.0 Lizenz** überschreitet, ist ohne Zustimmung des Verlags unzulässig. Die in diesem Werk enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Quellenangabe / Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Dieses Buch ist erhältlich als:
ISBN 978-3-7799-8524-2 Print
ISBN 978-3-7799-8525-9 E-Book (PDF)
DOI 10.3262/978-3-7799-8525-9

1. Auflage 2025

© 2025 Beltz Juventa
in der Verlagsgruppe Beltz · Weinheim Basel
Werderstraße 10, 69469 Weinheim
service@beltz.de
Einige Rechte vorbehalten

Satz: Datagrafix, Berlin
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Beltz Grafische Betriebe ist ein Unternehmen mit finanziellem Klimabeitrag
(ID 15985-2104-1001)
Printed in Germany

Weitere Informationen zu unseren Autor:innen und Titeln finden Sie unter: www.beltz.de

Inhalt

I. Theoretische und methodologische Überlegungen zu (audiovisuellen) Forschungsdaten in der empirischen Qualitativen Sozialforschung 9

Forschungsdaten der qualitativen sozialwissenschaftlichen Videoanalyse. Eine einleitende Systematisierung der audiovisuellen Datensorten in diesem Sammelband
Hubert Knoblauch und René Wilke 10

Daten über Daten über Daten. Das wissenschaftssoziologische Projekt »Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten« an der Universität Konstanz (1978–1982)
Christian Meyer und Christian Meier zu Verl 40

Die interpretative Analyse gefilmter Praxis. Verheißungen und Fallstricke
Michael Corsten 62

II. Videographische und medienanalytische Ansätze mit Vernakularen Videos natürlicher sozialer Situationen und inszenierter Inhalte 81

1. Videographische Vernakulare Videos 82

Ausbruch und Verharmlosung von Gewalt. Kontextualisierungen eines YouTube-Videos
Ulrike T. Kissmann 83

Videoanalytische Rassismusforschung
Christian Meyer, Hanna Grauert und Frank Oberzaucher 98

Zwischen Interaktion und Produkt. Polizeiliche Bodycamaufnahmen als hybrider Gegenstand der Videoanalyse
Simon Egbert und Jasper Janssen 120

Polizeiliche Vernehmungen. Zum Wandel einer hochintensiven Kommunikationsform durch ihre audiovisuelle Aufzeichnung
Mina Godarzani-Bakhtiari und René Tuma 144

II. Videographische und medienanalytische Ansätze mit Vernakularen Videos natürlicher sozialer Situationen und inszenierter Inhalte 167

2. Medial-gattungsförmig gestaltete Vernakulare Videos 168

Videoanalyse und Videointeraktionsanalyse als Methoden der empirischen Sexualforschung
Sven Lewandowski 169

Audiovisuelles Datenmaterial in der ethnographischen (Jugend-)Szeneforschung. Rapvideo-Eigenproduktionen als Zugang zu kulturellen Wissens- und Sinnzusammenhängen
Katharina Bock 186

Analyse von Videodaten. Unterschiede zwischen Feldvideos und Plattformvideos
Bernt Schnettler 201

III. Interaktionsanalysen mit Ethnographischen Videos 223

1. Interaktion in Paarbeziehungen, (Klein-)Gruppen, Kollektiven und der Öffentlichkeit 224

Videodaten bei der Analyse der Kommunikation mit Menschen mit der Diagnose Demenz
Jo Reichertz und Anna-Eva Nebowsky 225

Konversationsanalyse transsituativer Kollektionen. Die Erforschung projektiver Gattungen in ihren temporalen Zusammenhängen
Jonas Kramer und Sarah Hitzler 237

Klangorientiertes kommunikatives Handeln. Herausforderungen und Potenziale der Videoanalyse gemeinsamen Musizierens
Theresa Vollmer 254

Erhebung – Analyse – Nachnutzung. Audio-/Videodaten im Kontext sprachlicher Begabungsförderung
Caterina Mempel und Jenny Winterscheid 272

Verdeckte Videographie. Datenerhebung und -analyse im Spannungsverhältnis von Erkenntnisinteresse und Forschungsethik
Ajit Singh 290

III. Interaktionsanalysen mit Ethnographischen Videos 309

2. Interaktion und Interaktivität mit Gegenständen, Technik und digitalen Technologien 310

Analyse atypischer Kommunikation. Transkription von Videosequenzen als analytischer Ansatzpunkt von Mehrpersonengesprächen im Kontext Unterstützter Kommunikation
Imke Niediek 311

Digitale Materialitäten in multizentrischen Interaktionen. Mikroethnographische Verfahren zur Untersuchung gruppenförmigen Arbeitens im Tablet-gestützten Unterricht
Matthias Herrle, Matthias Proske, Aline Puzicha und Anne Zimmer 335

Eye-Viewing als Methode der qualitativen Sozialforschung. Durch Eye-Tracking-Technologie erweiterte Videographie zur Erforschung von (medienbezogenen) Praktiken in Schule und Unterricht
Isabel Neto Carvalho 356

Sehen und soziale Interaktion im Kunstmuseum. MET x EMKA als neuer Ansatz der Videoanalyse
Luise Reitstätter, Seda Pesen und Dirk vom Lehn 372

Autovideographie. Zur Analyse digitaler Spielpraktiken mithilfe von Multi-Kamera-Videos
Marcel Thiel-Woznica 392

Nonverbale Interaktionsmuster in Mixed Reality. Videographische Analysen von Human-Agent-Interactions
Jonathan Harth 408

IV. Nutzungsformen: Forschungsvideos im wissenschaftlichen Film und Forschungsdatennachnutzung 433

Wie Videoanalyse und soziologischer Film voneinander lernen können. Versuch einer theoretischen Grundlegung visualisierter Kommunikation in der Soziologie
Katharina Miko-Schefzig 434

Audiovisuelle Forschungsdaten und ihre Kontexte teilen. Archivierung und Nachnutzung von Daten aus der ethnografischen Filmforschung
Martin Gruber und Michaela Rizzolli 452

Bildungshistorische Perspektiven auf audiovisuelle Dokumente aus pädagogischen Fortbildungskontexten und der Unterrichtsforschung
May Jehle 470

De- und Rekontextualisierung in der Qualitativen Sekundäranalyse von Unterrichtsvideos
Anna Hamer, Jana Helbig und Michael Urban 484

Qualitative Sekundäranalyse. Zu Chancen und Herausforderungen der Nachnutzbarmachung qualitativer Daten für Forschung und Methodenlehre
René Wilke 500

I. Theoretische und methodologische Überlegungen zu (audiovisuellen) Forschungsdaten in der empirischen Qualitativen Sozialforschung

Forschungsdaten der qualitativen sozialwissenschaftlichen Videoanalyse

Eine einleitende Systematisierung der audiovisuellen Datensorten in diesem Sammelband

Hubert Knoblauch und René Wilke

Zusammenfassung: Dieser Text diskutiert einleitend die Begriffe »Daten« und »Forschungsdaten« im Kontext der qualitativen Sozialforschung. Aufbauend auf dem breiten Spektrum der Beiträge zu Videodaten in der deutschsprachigen qualitativen Sozialforschung, vertreten durch insgesamt 26 Einzelbeiträge in diesem Sammelband, entwickelt dieser Beitrag eine Systematik audiovisueller Forschungsdaten, die in diesem Forschungsfeld verwendet werden. Auf Grundlage der unterschiedlichen Erhebungs- und Analysekontexte audiovisueller Daten unterscheiden wir dabei zwischen *Ethnographischen Videos*, die von Forschenden erzeugt werden, und *Vernakularen Videos*, die aus privaten oder institutionellen Alltagskontexten stammen, sowie dem großen Bereich medialer Videodaten, die Gegenstand anderer Formen der (Medien-)Analyse sind. Wir zeigen, dass, zugleich mit dem Erhebungskontext, teilweise auch die Gegenstände der Daten und ihr Analysekontext variieren. Abschließend wird auf Grundlage der Systematisierung von *Erhebungskontexten*, *Analysekontexten* und *Nutzungsformen*, die den wissenschaftlichen Film sowie die Sekundäranalyse umfassen, ein strukturierter Überblick über die Einzelbeiträge des Sammelbands gegeben.

Schlüsselwörter: Forschungsdaten, empirische Wissenschaftstheorie, Videoanalyse, Ethnographische Videos, Vernakulare Videos

1. Einleitung

Der Sammelband entstand im Zusammenhang mit dem Versuch, eine webbasierte Forschungsdateninfrastruktur für audiovisuelle Daten der qualitativen empirischen Sozialforschung (aviDa[1]) aufzubauen. Weil dieses Vorhaben nicht

1 Das Projekt aviDa wurde durch die Deutsche Forschungsgemeinschaft (DFG) gefördert und erstreckte sich zwischen den Jahren 2018 und 2024 über zwei Förderphasen (Projektnummer: 396913378). Aktuell ist die fertig entwickelte Infrastruktur ohne technischen Betreiber. An dieser Stelle möchten wir uns bei der DFG sowie allen Projektmitarbeiter*innen bedanken. Für Korrektur und Kommentare bedanken wir uns bei Paula Kleine-Berkenbusch.

nur die Erhaltung von videographischen Forschungsdaten verfolgte, sondern vor allem die wissenschaftliche Nachnutzung ermöglichen sollte, stellte sich uns die zunächst einfach erscheinende Frage: *Wie können wir Videodaten so archivieren, dass sie von anderen Forscher*innen für Sekundäranalysen nachgenutzt werden können?* Im Laufe des Projekts zeigte sich, dass neben technischen Details auch allgemeine Standards berücksichtigt werden müssen, um die Nachnutzbarkeit von Forschungsdaten sicherzustellen. Die Erfordernisse der Nachnutzbarkeit von Forschungsdaten berühren heterogene und durch sehr unterschiedliche Komplexitätsgrade charakterisierte Bereiche der Forschung (vgl. Gruber/Rizzolli; Jehle; Hamer/Helbig/Urban; Wilke i. d. B.). Sie reichen von rechtlichen und materialen Aspekten (etwa dem Medium der Daten oder den Formaten ihrer Speicherung) über datenschutzrechtliche und forschungsethische bis hin zu organisatorischen, repräsentativen und methodologischen Fragen der Daten, ihrer Erhebung (Erzeugung bzw. Sammlung), Analyse und Nutzung.

All diese Aspekte sind auch relevant, wenn wir uns allgemein die Frage nach Videos als Forschungsdaten stellen. Diese Relevanz zeigte sich in der zweiten Projektphase von aviDa, als wir zur besseren potenziellen Nutzung des Archivs eine Ausweitung dessen, was wir im engeren Sinne »videographisch« (Tuma/Schnettler/Knoblauch 2013) nannten, auf einen weiteren Kreis audiovisueller Daten vornahmen. Dabei stellte sich heraus, dass qualitative Forschung hinsichtlich ihrer audiovisuellen Daten keineswegs leicht standardisierbar ist. Dies gilt bereits für Daten, die im weiteren Sinne ethnographisch erhoben werden, zeigt sich aber umso mehr, je weiter der Fokus auf audiovisuelle Daten eingestellt wird. Daher stellt sich die Frage, die wir in diesem Band verfolgen wollen: *Was sind audiovisuelle Daten der qualitativen Sozialforschung?*

Wer hofft, bei der Klärung dieser Frage auf bestehendes Wissen über den Begriff der »Daten« zurückgreifen zu können, sieht sich allerdings bald enttäuscht. Abgesehen davon, dass die Digitalisierung von Forschungsdaten selten als solche thematisiert wird, werden die Begriffe (digitale) »Daten« und »Forschungsdaten« häufig nahezu synonym verwendet. Zudem werden die allgemeinen Merkmale, die verschiedenen Arten von Forschungsdaten, heute kaum erörtert. So scheint, was Peter Gross und Thomas Luckmann schon in den 1970er Jahren bemerkt hatten, noch immer zu gelten:[2] Es gibt keine »Theorie der Daten«. Deswegen stellt sich auch die Frage: *Was sind eigentlich Forschungsdaten?*

2 Wir nehmen hier auf ein unveröffentlichtes Manuskript von Peter Gross Bezug, das durch das Sozialwissenschaftliche Archiv Konstanz (SAK) der Universität Konstanz archiviert und unter dem Titel »Theorie und sozialwissenschaftliche Daten« (DE-SAK-LUCK-F1.084-01) auf Anfrage zugänglich gemacht wird. Der Text entstand im Kontext des Forschungsprojekts von Luckmann und Gross »Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten« (Meyer/Meier zu Verl i. d. B.).

Wir wollen hier und mit diesem Sammelband zur Klärung der gestellten Fragen beitragen. Dabei bauen wir *erstens* auf Erfahrungen mit audiovisuellen Daten auf, die einer der Herausgeber dieses Sammelbands im Rahmen von empirischen Forschungsprojekten seit etwa 40 Jahren gemacht hat. *Zweitens* reflektieren wir die Erkenntnisse, die wir gemeinsam im Rahmen des Projektes zur Archivierung audiovisueller Daten (aviDa) gewonnen haben. Schließlich und *drittens* basieren die Vorschläge, die wir im Folgenden erläutern, maßgeblich auf den in diesem Band versammelten Texten etablierter Wissenschaftler*innen und Nachwuchsforscher*innen aus dem gesamten Forschungsfeld im deutschsprachigen Raum.

Mit dieser Sammlung unterschiedlicher Beiträge unternehmen wir den Versuch, die Breite der Videoforschung in der qualitativen empirischen Sozialforschung zumindest im deutschsprachigen Raum zu erkunden und abzubilden.[3] Zu diesem Zweck haben wir zunächst eine Tagung mit sozialwissenschaftlich-qualitativ Forschenden aus dem gesamten deutschsprachigen Raum durchgeführt, die mit einer großen Bandbreite audiovisueller Forschungsdaten arbeiten.[4] Daran anschließend haben wir den Kreis erweitert und über eine weit gestreute offene Ausschreibung weitere Forscher*innen eingeladen, ihren Umgang mit audiovisuellen Daten vorzustellen.[5]

Das Ziel, das wir uns mit diesem Sammelband setzen, ist nicht allein, die ausgewählten Beiträge zu dokumentieren. Vielmehr wollen wir – im Sinne einer *empirischen Wissenschaftstheorie*[6] – zur Beantwortung der genannten Fragestellungen beitragen: Wir behandeln deswegen in Abschnitt 2 die Frage, was *Forschungsdaten* im Allgemeinen sind, und, in Abschnitt 3, was *audiovisuelle Daten*

3 Im Vergleich zum angelsächsischen Raum, in dem die ethnomethodologisch orientierte Videographie dominiert (mustergültig dazu: Heath et al. 2010), zeichnet sich der deutschsprachige Raum erstaunlicherweise durch eine größere Vielfalt methodischer, theoretischer und disziplinärer Ansätze aus, die indessen, ebenso verwunderlich, auch deutlich fragmentierter, zuweilen auch hermetischer erscheinen.

4 Die Tagung fand unter dem Titel »Qualitative Videoanalyse – Audio-visuelle Datensorten und ihre Kontexte« am 9. und 10.11.2023 am Institut für Soziologie der TU Berlin statt. Wir bedanke uns an dieser Stelle bei allen Teilnehmer*innen und Helfer*innen, insbesondere bei Paula Kleine-Berkenbusch.

5 Auf unseren Call for Abstracts: Buchprojekt »Audio-visuelle Daten in der empirischen qualitativen Sozialforschung: Erhebung – Analyse – Nutzung – Transkription« erhielten wir eine große Zahl von Beitragsvorschlägen. Dass wir nicht alle vorgeschlagenen Kontributionen in den Band aufnehmen konnten, liegt zum einen an Platzgründen, zum anderen aber maßgeblich an den äußeren Grenzen des gesetzten Themas und dem explizit methodologischen Fokus dieses Sammelbands. Wir bedanken uns an dieser Stelle ausdrücklich bei allen, die einen Beitrag vorgeschlagen hatten.

6 Es geht uns hier also nicht nur um die wissenssoziologische Beobachtung der forschenden Praxis, der Vorstellungen über Videodaten und deren technisch erfassbare Eigenheiten, sondern auch um eine Klärung dessen, was Videodaten sind, die als Orientierung für die Arbeit in der Archivierung und in der Forschung selbst dienen soll. Zur empirischen Wissenschaftstheorie vgl. Knoblauch (2021).

der qualitativen Sozialforschung sind. Dabei grenzen wir ein, was wir als *Ethnographische Videos* bezeichnen, indem wir diese von *Vernakularen Videos* unterscheiden. Letztere differenzieren wir weiter, indem wir *ethnographische* von *medial-gattungsförmigen Vernakularen Videos* unterscheiden. Schließlich geben wir in Abschnitt 4 einen nach Erhebungs- und Analysekontexten sowie Nutzungsformen systematisierten Überblick über die in diesem Band versammelten Einzelbeiträge.

Der Band gliedert sich in die folgenden Kapitel, denen wir die Einzelbeiträge gemäß unserer Datentypologie systematisch zuordnen. Verweise auf diese Kapitel finden sich auch in dem systematischen Überblick am Ende dieses Beitrags (Abschnitt 4):

I. Theoretische und methodologische Überlegungen zu (audiovisuellen) Forschungsdaten
II. Videographische und medienanalytische Ansätze mit *Vernakularen Videos* natürlicher sozialer Situationen und inszenierter Inhalte
III. Interaktionsanalysen mit Ethnographischen Videos
IV. Nutzungsformen: Forschungsvideos im wissenschaftlichen Film und Forschungsdatennachnutzung.

2. Was sind Forschungsdaten?

Die Frage, was Forschungsdaten sind, erscheint zunächst leicht zu beantworten, wird der Begriff doch in jedem Methodenbuch verwendet. Dort werden seine Ausprägungen zumeist nur beispielhaft aufgelistet, wie auch in den Leitlinien der Deutschen Forschungsgemeinschaft (DFG): »Zu Forschungsdaten zählen u. a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten, Objekte aus Sammlungen oder Proben, die in der wissenschaftlichen Arbeit entstehen, entwickelt oder ausgewertet werden« (DFG 2015, S. 1). Während die DFG daneben auch »methodische Testverfahren, wie Fragebögen, Software und Simulationen« (ebd.) unter Forschungsdaten fasst,[7] werden Daten in anderen Definitionen als Ergebnisse von Verfahren dargestellt, da sie im Zuge wissenschaftlicher Vorhaben durch »Digitalisierung, Quellforschung, Experimente, Messungen, Erhebungen oder Befragungen entstehen«.[8] Diese Unklarheit hinsichtlich des »Datenbegriffs« wird zusätzlich verstärkt, indem viele Definitionen tautologisch

7 Ähnlich definiert und systematisiert auch Peter Andorfer (2015), der Forschungsdaten als Quellen (Archivdokumente) und Arbeitsdaten (Bibliographie, Exzerpte, Textentwürfe, Transkriptionen) fasst.

8 Schwerpunktinitiative »Digitale Information«: Forschungsdaten, Beschreibung des Handlungsfeldes der Allianzinitiative Digitale Information, zit. n. Wikipedia, »Forschungsdaten« (16.08.2024).

sind, d. h. dass sie »Daten« sowohl als *definiens* wie auch als *definiendum* nutzen.[9] Die Situation wird noch undurchsichtiger durch den ›doppelten Datenbegriff‹: Oft bleibt unklar, ob »Daten« etwas spezifisch Forschungsbezogenes, also »Forschungsdaten«, meint oder lediglich auf den Umstand hinweist, dass diese Daten technisch als digitale Daten vorliegen. Auch wenn heutzutage die meisten Forschungsdaten digital sind, sollte der Kurzschluss vermieden werden, dass digitale Daten mit Forschungsdaten gleichzusetzen seien.

Wissenschaftshistorisch erscheint der Begriff »Datum« keineswegs geklärt.[10] Seine Verwendung hängt offenbar eng mit der Ausbreitung des Positivismus im 19. Jahrhundert zusammen: So suggeriert das Wort »Datum« epistemologisch den Bezug zur »Wirklichkeit«, als etwas, das unmittelbar beobachtet werden kann, sozusagen von Natur aus gegeben ist. Dagegen weist die Herkunft des Begriffs »Datum« aus der Mathematik auf einen zusätzlichen Aspekt hin, der durch das Wort »Maß« hervorgehoben wird. Erkenntnistheoretisch folgt diese Verbindung noch der (seit Galileo eingeführten) Unterscheidung zwischen den (messbaren) »primären Qualitäten«, die dem Dinge anhaften, und den sekundären Qualitäten, in denen sich Objekte den menschlichen Sinnen zeigen (Luckmann 1980, S. 17 ff.). Diese Vorstellung der sekundären Qualitäten, die Gegenstand der »Reflexion« werden, wird im Begriff der »Sinnesdaten« (»sense data«) von John Locke formuliert.

Auch wenn Forschungsdaten gerade in der jüngeren Zeit zunehmend thematisiert wurden, hat sich die Wissenschaftstheorie damit bisher wenig beschäftigt. So findet sich etwa in der von Jürgen Mittelstrass (2024) herausgegebenen achtbändigen Enzyklopädie der Wissenschaftstheorie kein Eintrag zu »Daten«. Eine der wenigen namhaften wissenschaftstheoretischen Ausarbeitungen schlagen James Bogen und James Woodward (1992) vor. Um den Datenbegriff zu begründen, gehen sie von einem erkenntnistheoretischen Begriff des *Phänomens* aus, als dem, was sich dem wahrnehmenden Subjekt als Gegenstand der Beobachtung zeigt, über das eine Erkenntnis gemacht und damit Wissen erworben wird. Sie beschränken den Begriff aber nicht auf die Erkenntnis im Allgemeinen; vielmehr gehen sie davon aus, dass Phänomene insbesondere für die moderne Wissenschaftstheorie relevant sind: Dem »zweigliedrigen Wissenschaftsmodell« zufolge (siehe Abb. 1) beziehen sich *wissenschaftliche Theorien* demnach auf *Phänomene*, die sie *belegen* und die von ihnen *erklärt* werden. Allerdings erweitern sie dieses Modell: Im Unterschied zur nicht-wissenschaftlichen Erkenntnis spielen *Daten* in der Wissenschaft eine Sonderrolle. Bogen und Woodward bestimmen sie als

9 Diese Tautologie findet sich auch im Wikipedia-Artikel zu (hier auch digitalen) »Forschungsdaten«: »Unter digitalen Forschungs*daten* verstehen wir alle digital vorliegenden *Daten*, die während des Forschungsprozesses entstehen oder ihr Ergebnis sind« (16.08.2024; Herv. H. K./R. W.).

10 So enthält etwa die wissenschaftshistorische Arbeit von Lorraine Daston und Peter Galison (2007) zur »Objectivity« der Naturwissenschaft keinen Registereintrag zu »Daten«.

Messergebnisse, die dem menschlichen Wahrnehmungssystem zugänglich sind und die man daher öffentlich überprüfen kann. Daten ergänzen das zweigliedrige Wissenschaftsmodell daher auf eine Weise, die sich mit Bogen und Woodward (und in Anlehnung an Apel 2011) als das »dreigliedrige Wissenschaftsmodell« darstellen lässt (siehe Abb. 1).

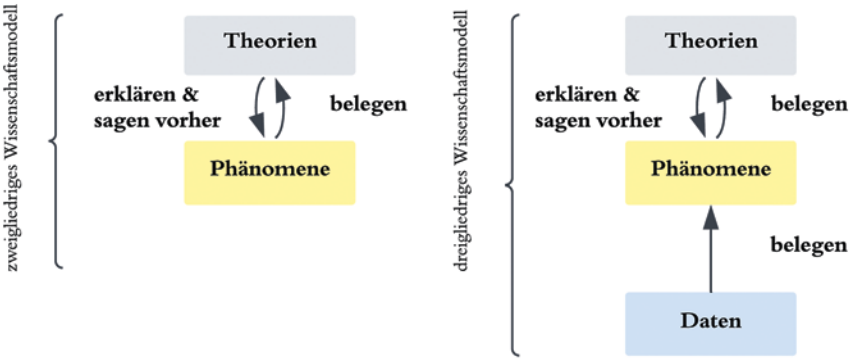


Abbildung 1: Links das »zweigliedrige Wissenschaftsmodell«, rechts das »dreigliedrige Wissenschaftsmodell« nach Bogen und Woodward (eigene Darstellung nach Apel 2011, S. 17–22)

Diese Definition des Datenbegriffs ist zwar sehr grundlegend und damit sehr abstrakt, erlaubt aber immerhin einen Vorbegriff der »Daten«, indem sie von *Phänomenen* und *Theorien* unterschieden werden. Diese Unterscheidung erfolgt allerdings gleichsam von außen (bzw. *semantisch extensional)*, während die innere (*semantisch intensionale*) Bestimmung der sie auszeichnenden Merkmale dabei unzulänglich bleibt. Wenn sie Daten als »Messergebnisse« bezeichnen, haben wir zwar einerseits den Hinweis auf eine (implizit technisch vermittelte) Handlung (»Messen«) und deren Objektivierung (als »Ergebnis«). Allerdings fragt sich gerade bei Daten der qualitativen Sozialforschung, ob wir sie alle unter dem Begriff des Messens fassen können. Selbst wenn wir eine breite Theorie des Messens unterstellen, wie sie in den Sozialwissenschaften möglich ist, so können doch gerade qualitative audiovisuelle Forschungsdaten selten darunter gefasst werden. Zwar gibt es auch im qualitativen Bereich einzelne Ansätze, die etwa Bilder in einem gewissen Sinne vermessen.[11] Dennoch kann Messen als lediglich *eine* für die Herstellung von Daten relevante Handlungsform angesehen werden, neben der es viele andere gibt (und zwar auch in den Naturwissenschaften, wie etwa bei der biologischen Klassifikation von Pflanzen oder Tieren).

11 So z. B. Techniken aus der Kunstwissenschaft wie die Planimetrie oder der Goldene Schnitt, die in der wissenssoziologischen Videohermeneutik (Jürgen Raab), der Dokumentarischen Bild- und Videoanalyse (Ralf Bohnsack) oder der visuellen Diskursanalyse (Boris Traue) zum Einsatz kommen.

Für eine erste Klärung der Frage, was Daten (hier: sozial-)wissenschaftlicher Forschung sind, hat ein Forschungsprojekt von Luckmann und Gross mit der Kurzbezeichnung »Daten über Daten« sehr wesentlich beigetragen, das schon in den späten 1970er Jahren eine Art empirische Erforschung wissenschaftlicher Methoden vorgenommen und sich insbesondere mit der Frage der Konstitution sozialwissenschaftlicher Daten beschäftigt hat. Das wissenschaftsgeschichtlich gesehen sicherlich sehr innovative Projekt wurde an der Universität Konstanz durchgeführt, ist jedoch wenig bekannt und wird deswegen hier in diesem Band etwas detaillierter vorgestellt (Meyer/Meier zu Verl i. d. B.). Ähnlich wie das berühmtere Vorgängerprojekt der »First Five Minutes« (Pittenger/Hockett/Dassehy 1960) bezog sich das »Daten über Daten«-Projekt u. a. auf die Frage nach Erzeugung und Verarbeitung von Interviewdaten. Im Unterschied zu seinem amerikanischen Vorbild ging es hier jedoch nicht darum, die Möglichkeiten der Analyse und Interpretation des Datums zu eruieren; vielmehr lag das Augenmerk darauf, zu untersuchen, wie sozialwissenschaftliche Daten erzeugt werden. Dazu waren Interviews, die von den Forschenden selbst durchgeführt wurden, audiovisuell aufgezeichnet worden. Diese Aufzeichnungen wurden dann auf sehr unterschiedliche Weisen transkribiert, die den Fachanforderungen der beteiligten Disziplinen entsprachen. Auf diese Weise wurden die Interviews *in Daten transformiert*, die es erlaubten, sie aus verschiedenen Perspektiven zu lesen und zu interpretieren. Die Verbindung der verschiedenen Transkriptionsweisen (vgl. Niediek; Vollmer i. d. B.) wurde »Partitur« genannt und entsprach deutlich dem, was in der linguistischen Forschung Jahrzehnte später als Multimodalität bezeichnet wurde.

Im Zusammenhang mit diesem Projekt wurden von Luckmann und Gross auch verschiedene Typologien von Daten vorgeschlagen, die damals noch auf der Grundlage zeitgenössischer analoger Datensorten beruhten. Für unsere Frage nach dem, was Forschungsdaten sind, ist daher die Unterscheidung verschiedener Objektivierungsstufen sehr viel einschlägiger, die sich aus der Transformation der Daten ergeben. Auch wenn das Projekt einen sehr großen Wert auf die Wahrnehmung und eben Konstitution der Interview-Daten als Erfahrungsgegenstände legte, so verweist der Begriff der Objektivierung (auch hier) auf deren handelnde Erzeugung und ihre Materialität als Erzeugnis. Erst die Be-Handlung der »Rohdaten« – in diesem Fall durch die Forschenden im Labor zunächst auf (U-matic-)Videokassetten festgehaltene Aufzeichnungen – durch Anhören und Transkribieren, macht sie zu (Daten-)Transkripten, die schließlich in Datensitzungen analysiert werden (vgl. dazu Reichertz 2013). Entsprechend lassen sich Forschungsdaten nicht nur nach ihrer Modalität und Medialität unterscheiden (z. B. akustische oder visuelle Aufzeichnung, geschriebener Text auf Papier, digitaler Text), sondern auch nach den Forschungstechniken, mit denen sie erzeugt (z. B. Beobachtung, Interview) und dann weiterverarbeitet bzw. transformiert

werden. Daten sind also handelnd konstruiert. Für die Art und Weisen, wie sie konstruiert werden, haben wir den Begriff der *Datensorten* vorgeschlagen. Datensorten werden kommunikativ konstruiert und sind deswegen geprägt von den »Verfahren, mit denen sie erzeugt werden« (Knoblauch 2003, S. 59). So sind Gespräche, die etwa als Leitfaden-Interviews von den Forschenden anberaumt, geleitet, strukturiert und aufgezeichnet werden, eine andere Datensorte als ›natürliche‹ Gespräche, die sich zwischen Handelnden abspielen und dabei von Dritten aufgezeichnet werden; entsprechend stellen große Bild-Datenkorpora aus Instagram oder Textkorpora von X (vormals Twitter) andere Datensorten dar als z. B. Hochzeitsvideos, die von Laien oder Professionellen erzeugt werden (Raab 2002). In allen drei Fällen sind die Datensorten jedoch mit einem technologischen Medium verbunden, das sie aufzeichnet bzw. objektiviert.

Wie Gross (1985) betont, zeichnen sich Daten dadurch aus, dass sie »erfahren« werden können. Eine Voraussetzung dafür ist, dass die Daten (im möglichen Unterschied zu dem, worauf sie sich beziehen) sinnlich wahrnehmbar sind – und das gilt selbst für Atome oder Ultraschallklänge. Auch wenn es sich bei den Gegenständen um abstrakte Phänomene handelt, die der Wahrnehmbarkeit entrückt sind (wie etwa subjektive Empfindungen, Strahlungen aus dem Weltraum oder Deutungen von Träumen), so zeichnen sich Forschungsdaten dadurch aus, dass sie objektiviert sind, um von den Forschenden sinnlich wahrgenommen und gedeutet werden zu können. Daten sind immer Objektivationen (Knoblauch 2017, S. 161 ff.). Häufig werden Daten z. B. in zeichenhafte Objektivationen »transformiert«, wenn etwa neuronale Prozesse in konventionalisierten Farben repräsentiert (Wilke 2022, S. 215 ff.) oder musikalische Klänge in Transkriptionsmodelle integriert werden (Vollmer i. d. B.). Die oft zeichenhafte Objektivität macht sehr deutlich, dass nicht nur die Erzeugung der Daten, sondern auch ihre Transformation selbst ein Handlungsprozess ist, der zuweilen von (Mess-, Transkriptions- oder Aufnahme-)Geräten unterstützt oder heute immer häufiger auch durch sie ersetzt wird. In jedem Fall haben wir es mit Objektivationen zu tun, die in technisch vermitteltem kommunikativem Handeln erzeugt und dadurch auch für andere (intersubjektiv) beobachtbar gemacht werden.

Denken wir diese handelnde Erzeugung von Daten als Objektivationen hinzu, erscheint die vorgeschlagene wissenschaftstheoretische Unterscheidung der Daten von den Phänomenen sehr hilfreich: Daten *›repräsentieren‹* (auf Weisen, die wir unten erläutern werden) Phänomene als Objektivationen im wissenschaftlichen Handeln, sodass sie auch von anderen wahrnehmbar werden. Die Form der Daten selbst ist sicherlich abhängig von Methodologien, Forschungsmethoden und den Objekten. In allen Fällen allerdings sind Daten das Gegenteil von dem, was das Wort bedeutet: Sie sind nicht ›gegeben‹, sondern das Resultat von Handlungen; sie werden von den Forschenden erzeugt, mithilfe von Geräten

produziert oder allgemeiner: kommunikativ konstruiert.[12] Das Datum ist Ergebnis eines Handlungszusammenhangs, der Objekte (etwa in Laboratorien) isoliert, Zahlen oder Namen identifiziert oder sie mit Gerätschaften und soziotechnischen Systemen selbst generiert.

Bei der Frage, was als Datum betrachtet wird, spielt das, was Theorie heißt, eine zentrale Rolle. Denn die Frage, was objektiviert wird, ist unmittelbar mit dem i. d. R. diskursiven Wissen verbunden, das die Phänomene und Phänomenbereiche beschreibt und bestimmt. Der Zusammenhang zwischen Theorie, Daten und Phänomenen[13] besteht jedoch nicht nur im deduktiven Erklären und Vorhersagen, wie Bogen und Woodward nahelegen. Er spielt auch in der induktiv und abduktiv verfahrenden qualitativ orientierten empirischen Sozialforschung eine mindestens ebenso große Rolle. Angesichts der Diversität der dabei verwendeten Datensorten können wir sie hier nicht behandeln, sondern wollen uns auf die audiovisuellen Daten beschränken, die in unserem Forschungszusammenhang im Zentrum stehen.

3. Was sind audiovisuelle Daten der qualitativen Sozialforschung?

Wenn wir uns dieser Frage zuwenden, geht es uns keineswegs darum, eine scholastische Bestimmung aus einer besonderen Position im Feld heraus willkürlich vorzunehmen, was diese Daten sind. Vielmehr nutzen wir die Unterschiedlichkeit und Vielzahl der Beiträge in diesem Sammelband selbst, um näher zu bestimmen, was audiovisuelle Daten sind, die in der videoanalytischen qualitativen Sozialforschung behandelt werden. Der Begriff und die Systematik der audiovisuellen Daten werden also im Sinne der empirischen Wissenschaftstheorie (Knoblauch 2021) gewonnen, indem wir die Ähnlichkeiten und Unterschiede der untersuchten Gegenstände bzw., im wissenschaftstheoretischen Sinne, Phänomene betrachten. Daneben spielen aber auch verschiedene Kontexte eine zentrale Rolle. Im Unterschied zur klassischen wissenschaftstheoretischen Unterscheidung zwischen den »Contexts of Discovery« und »Contexts of Justification« (Hoyningen-Hüne 1987) spielen hier die (Handlungs-)Kontexte der (Daten-)Erhebung, der (Daten-)Analyse und der Nutzung von Daten eine tragende Rolle. Dies erlaubt uns nicht nur (a) die epistemische Unterscheidung verschiedener Datensorten und b) die Abgrenzung zu anderen audiovisuellen Daten (z. B. der

12 Vgl. Knorr-Cetina (1984). Dabei sollten wir betonen, dass sich diese Aussagen keineswegs nur auf die qualitative oder sozialwissenschaftliche Forschung beziehen, sondern möglicherweise auch auf andere Felder der Wissenschaft übertragen werden können.

13 Wir werden im Weiteren den wissenschaftstheoretischen Begriff des »Phänomens« durch den des (Forschungs-)Gegenstands ersetzen, der in der empirischen Forschung gebräuchlicher ist.

Film-, TV- oder Laborforschung), sondern ist auch (c) Grundlage für den praktischen Umgang mit ihnen in der Forschung, der Archivierung und der Nachnutzung von Forschungsdaten.

Da ›Videodaten‹ durch die Digitalisierung mittlerweile von einer (aus wissenschaftlicher Sicht) »Veralltäglichung« geprägt sind, werden sie nicht nur in hochgradig fragmentierten wissenschaftlichen, sondern auch in vielfältigen sozialmedialen und nicht-wissenschaftlichen Kontexten genutzt, die (im systemtheoretischen Sinne) sozusagen die ›Umwelt‹ der Daten darstellen, mit denen wir es hier im Kern zu tun haben. Angesichts der enormen Vielzahl *Vernakularer Videos*, ihrer Genres und Formen sowie ihrer Behandlung in der Forschung, müssen wir das, was als Datum der videoanalytischen qualitativen Sozialforschung genutzt wird, sehr viel genauer fassen. Dazu verwenden wir den Begriff »videographisch«: Wie immer auch qualitative Videodaten erhoben werden (gesammelt oder »im Feld« (von Forscher*innen) erzeugt), geht es in unserem Forschungszusammenhang im Kern nicht um die Analyse von Videos an sich und auch nicht um fiktive Inhalte oder andere Genres, die Gegenstand medienwissenschaftlicher Film- und Fernsehanalysen sind (Peltzer/Keppler 2015). Vielmehr geht es i. d. R. um Videos, die sich durch *bestimmte* Gegenstände auszeichnen, die auf ihnen *abgebildet* sind. Bei diesen Gegenständen handelt es sich um soziale Prozesse, die mit allgemeinen theoretischen Begriffen als »Interaktionen«, »Praktiken«, (kommunikative) »Handlungen« oder in einem sehr engen (Goffman'schen) Sinn als »soziale Situationen«, »Ereignisse« oder »events« bezeichnet werden. Im Unterschied zu z. B. Interviews, Surveys oder Laborforschungen, zeichnen sich diese Situationen dadurch aus, dass sie nicht von den Forschenden initiiert, organisiert oder manipuliert wurden und daher »(quasi-)natürliche« soziale Situationen darstellen.[14] In diesem Sinne schlagen wir die Bezeichnung *videographisch* (im weiteren Sinne) für die Vielzahl (erzeugter oder gesammelter) Videodaten vor, deren Gegenstand natürliche soziale Prozesse darstellen. Es geht uns also hier nicht nur um ethnographisch (von Forschenden im sozialen Feld) erhobene Daten (Tuma/Schnettler/Knoblauch 2013). Vielmehr geht es uns um Daten, deren Gegenstand (bzw. »Phänomen«) soziale Prozesse sind.[15]

Videodaten, die von Forschenden in ethnographischen Forschungsfeldern erzeugt wurden, bezeichnen wir als *Ethnographische Videos*. Damit grenzen wir sie hinsichtlich ihrer Erzeugungsweise von solchen audiovisuellen Daten ab, die in alltäglichen, privaten oder Arbeitskontexten produziert und von Forschenden zwar nicht erzeugt, aber zur Analyse gesammelt wurden. Diese bezeichnen wir

14 Bei den sozialen Situationen kann es sich auch um besondere »gerahmte« Situationen handeln, wie etwa Videoaufzeichnungen von Menschen bei der Interaktion mit KI (vgl. Harth i. d. B.) oder beim Computerspielen (vgl. Thiel-Woznica i. d. B.).

15 Vgl. dazu auch Meier zu Verl (2018, S. 76).

als *Vernakulare Videos*.[16] Innerhalb der Gruppe der *Vernakularen Videos* muss allerdings ein weiteres Mal differenziert werden: Wenn *Vernakulare Videos* soziale Prozesse in natürlichen Situationen abbilden und diese Gegenstände als solche von Forschenden für Analysen genutzt werden, bezeichnen wir diese Videos als *videographische Vernakulare Videos*. Wie allerdings schon Raab (2002) am Beispiel von (analogen) Hochzeitsvideos zeigte, handelt es sich bei nicht zu wissenschaftlichen Auswertungszwecken erzeugten und von Forschenden gesammelten Videodaten nicht selten auch um besonders ›gestaltete‹ Gattungen. Diese *Vernakularen Videos* grenzen wir daher nochmals ab und bezeichnen sie als *medialgattungsförmige* (oder gestaltete) *Vernakulare Videos* (siehe Abb. 2).

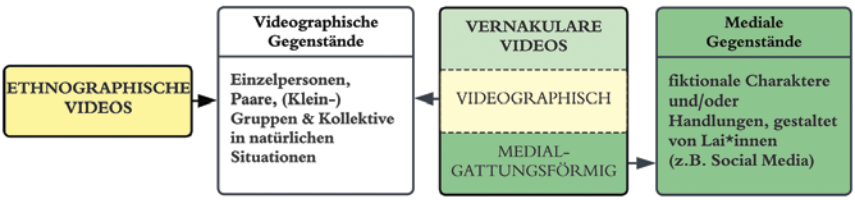


Abbildung 2: Typologie der audiovisuellen Forschungsdaten dieses Sammelbands

3.1 Zur Entwicklung der sozialwissenschaftlichen Videoanalyse

Die spezifische Ausrichtung der Videoanalyse als sozialwissenschaftliche Erhebung von Videodaten sozialer Prozesse folgt verschiedenen Pfaden der Forschung mit audiovisuellen Medien und Videos, die sich in unterschiedlichen Disziplinen, Richtungen und Ansätzen der qualitativen sozialwissenschaftlichen Forschung historisch ausgebildet haben und, wie Ramón Reichert (2007) gezeigt hat, auch durch die Entwicklung der Medientechnologie geprägt wurden. Eine wichtige Rolle spielt sicherlich der ethnologische Film, der sich schon sehr früh der Analyse des ›Verhaltens‹, ›Handelns‹, oder ›Interagierens‹ von Menschen zugewandt hat. Berühmte Autor*innen sind Alfred C. Haddon, Robert Flaherty oder Margret Mead, die den Film zur Analyse menschlichen Verhaltens einsetzten. Damit produzierte die Anthropologie eine große Sammlung von Filmdaten. Hauptsächlich setzte sie diese zur Dokumentation ein, etwa von Arbeitsweisen, -instrumenten und -produkten in Gesellschaften oder Kulturen. Dabei brachte sie diese Daten auch gestalterisch in die Form des dokumentarischen, ethnologischen oder allgemeiner: des wissenschaftlichen Films (vgl. Gruber/Rizzolli; Miko-Schefzig i. d. B.).

16 Vgl. dazu Tuma (2016). Tuma zeigt zudem die Ausbildung »vernakularer« Formen der Videoanalyse, die sich außerhalb der akademischen Forschung ausgebildet haben.

Seit den 1920er Jahren führten Kurt Lewin und Arnold Gesell filmische Analysen der Handlungen von Kindern durch (Thiel 2003). Berühmt wurden auch die Analysen des Tanzens in Bali von Gregory Bateson und Margaret Mead (1942). Darauf aufbauend wurden in der sogenannten Palo-Alto-Gruppe Filme eingesetzt, um Interaktionen zwischen Familienmitgliedern in westlichen Gesellschaften zu untersuchen. Eine ebenso »mikrosoziologische« Perspektive nahm Ray Birdwhistell (1952) ein, der das Zusammenspiel von nonverbalem und verbalem Verhalten bis ins kleinste Detail analysierte, oder Albert E. Scheflen (1965), der die Rolle der Körperhaltung in psychotherapeutischen Gesprächen auf Grundlage von Filmaufzeichnungen untersuchte. Weit über die Wissenschaft hinaus berühmt wurden die experimentellen Untersuchungen von Paul Ekman und Wallace V. Friesen (1969), die mithilfe von Filmmaterial »nonverbales Verhalten« analysierten, um menschliche »Universalien« zu erkennen (z. B. Lachen).

Während diese Analysen auf der Grundlage von Filmen durchgeführt wurden, änderte sich Forschung mit der Einführung der Videoaufnahmetechnik in den 1970er Jahren erheblich. Mit der Miniaturisierung der Technologie wurde der Einsatz des Videos zur Erhebung von Daten in ›natürlichen‹ Situationen erleichtert. Wichtige Anstöße erhielt die Videographie dabei durch die Ethnomethodologie. So analysierte etwa Charles Goodwin (1981) Alltags-Gespräche, um zu zeigen, wie z. B. Blicke dazu beitragen, Interaktionen zu koordinieren. Frederick Erickson und Jeffrey Shultz (1982) setzten Video zur Untersuchung von Gesprächen zwischen Schulberater*innen und Schüler*innen ein. Christian Heath führte Videostudien durch, die komplexe soziale Situationen fokussierten, z. B. die Rolle der neu eingeführten Computer bei der Interaktion zwischen Ärzt*innen und Patient*innen in medizinischen Sprechstunden (Heath 1986).

Angesichts der großen Rolle der Technologie für die Entwicklung der Videographie hat der Übergang zur Digitalisierung weitreichende Folgen für den gesamten Forschungsbereich. Er zieht sich über die verschiedenen technischen Entwicklungen, wie etwa der Schritt vom Video-Magnetbandrekorder (VTR) zum Videokassettensystem (VCR) (ab 1969), bis schließlich zu der Integration von Videofunktionen in mobilen Telefongeräten (ab 2003) sowie der translokalen Zugänglichkeit von mit mobilen Endgeräten erstellten Videoinhalten in Clouds und auf einer großen Zahl von Social-Media-Plattformen. Diese technische Entwicklung hat nicht nur die wissenschaftliche Forschung mit audiovisuellen Daten massiv erleichtert, sondern auch zu einer Art Explosion der vernakularen Videoproduktion geführt, die von einer immer umfänglicher werdenden Zahl an individuellen, aber auch institutionellen Akteuren erzeugt und öffentlich zugänglich gemacht wird. Mit *Vernakularen Videos* ist dabei eine neue audiovisuelle Datensorte entstanden, die sich als von wachsender Bedeutung für die qualitative sozialwissenschaftliche Videoanalyse erweist. Sie überschreitet mit dieser neuen Datensorte, wie zuvor erläutert, teilweise den Bereich des *Videographischen* und

greift in den Bereich der eher kunst- und/oder medienwissenschaftlichen *Medienproduktanalyse* über (vgl. Bock; Schnettler i. d. B.).

3.2 Mediale Grundzüge von audiovisuellen Daten

Die spezifisch auf die Verwendung von Videotechnologie zielende methodische Reflexion setzte in den 1980er Jahren (z. B. Erickson 1982; Heath 1997) ein und ist inzwischen in Sammelbänden (z. B. Knoblauch et al. 2006; Kissmann 2009; Bohnsack/Fritzsche/Wagner-Willi 2015), englischsprachigen Lehrbüchern (z. B. Dinkelacker/Herrle 2009; Heath et al. 2010; Knoblauch et al. 2014), in einem Handbuch der Videoanalyse (Moritz/Corsten 2018) sowie in Sammlungen für bestimmte Anwendungsbereiche wie Ethnologie, Erziehungswissenschaften, Religionswissenschaft, Linguistik, Soziologie und andere Sozialwissenschaften dokumentiert. Wir wollen hier einige Grundzüge von Videodaten zusammenfassen.

Im Unterschied zum Film besteht ein Prinzip von Video darin, die audiovisuellen Aufzeichnungen im selben technischen System betrachten und behandeln zu können: wiesen analoge Videogeräte dies durch die Zweiteilung von Aufzeichnungs- und Wiedergabefunktionen aus, so finden sich beide Funktionsbereiche auch in heutigen Smartphones zwar integriert, aber immer noch unterteilt. Wie beim Film steht auch bei Video die Möglichkeit der Aufzeichnung auf externen Datenträgern zur Verfügung, die unabhängig vom Aufzeichnungsgerät abgespielt, gesichtet und bearbeitet werden können. Diese bereits im historischen Entwicklungspfad der Technologie angelegte Zweiteilung in Aufzeichnung und Wiedergabe ist für die sozialwissenschaftliche Videographie folgenreich, weil sie technisch, zeitlich und dann auch institutionell (etwa in Gestalt von »Datensitzungen«) die Differenz zwischen Erhebung und Analyse (Aufbereitung, Codierung, Transkription etc.) zur Folge hatte.

Videodaten sind zudem *audiovisuelle* Daten. Das bedeutet nicht nur, dass Ton und Bild gleichzeitig aufgezeichnet, sondern auch, dass sie getrennt werden können. Zwar erlaubten analoge Systeme noch keine Trennung von Ton- und Bildspur, doch wurden beide Modalitäten schon früh zu eigenen Datensorten weiterverarbeitet: der Ton konnte paraphrasiert, zusammengefasst oder transkribiert werden, wobei die Art der Transkription häufig schon mit besonderen Forschungsinteressen, -disziplinen oder -ansätzen verbunden war (z. B. linguistische, prosodische oder konversationsanalytische Transkription); auch die Bilder konnten als eigene »Spur« in zum Teil sehr aufwändigen Verfahren verschriftlicht und damit zu einer eigenen Datensorte werden (wobei sich auch hier methodologische und disziplinäre Differenzen in der Art der Verschriftlichung niederschlugen).[17]

17 Zur Transkription von audiovisuellen Forschungsdaten siehe auch Moritz (2018).

Die Videoanalyse nutzt die allgemeinen Merkmale des Videos als audiovisuelles Medium, dessen Besonderheiten in den 1980er Jahren von Grimshaw (1982, S. 122, dt. H. K.) so benannt wurden: »Die beiden Hauptvorteile von SIR [Sound-Image Data Records] sind Dichte und Dauer. Andere Aufzeichnungen können das eine oder das andere dieser Attribute aufweisen; keine andere hat beides«. Was die *Dauer* anbelangt, so zeichnet sich das Video, wie der Film, als ein diachrones *zeitliches* Medium aus. Diachron werden Bildsequenzen so wiedergegeben, dass die lineare Zeitlichkeit der von ihnen abgebildeten Abläufe wiedererkannt werden kann (das gilt auch für die akustischen Phänomene, die von den technischen Systemen im Regelfall nur als zeitliche Abläufe wahrnehmbar gemacht werden können). Im Unterschied zum Foto registriert es Abläufe in der Zeit und gibt sie in zeitlicher Abfolge wieder. Weil die Aufzeichnungen gespeichert werden, stehen sie daneben auch dauerhaft zur Verfügung. Dies galt für Videobänder oder CDs, es gilt noch mehr für digitale Speichermedien. Damit sind die aufgezeichneten Abläufe im Wesentlichen unverändert reproduzierbar.

Im Unterschied zu Feldprotokollen oder auch Audioaufzeichnungen weisen audiovisuelle Daten zum zweiten eine besondere *Dichte* auf. Dichte bedeutet, dass sie – neben dem zeitlichen Ablauf – in statischen Einzelbildern reichhaltige synchrone Informationen enthalten, die simultan betrachtet werden können. Sie können so zum Gegenstand semiotischer, bildhermeneutischer oder ethnographischer Elizitierungsverfahren gemacht werden (Knoblauch 2004, S. 134 f.).

Wie der Film oder die Fotografie tritt das Video mit dem Anspruch einer *Abbildung* auf. Es ist auch *mimetisches Medium* genannt worden, weil es audiovisuell registriert, worauf die Videokamera gerichtet ist.[18] Mimetik sollte aber nicht als bloße Abbildung verstanden werden. Denn wie alle Forschungsdaten sind auch Videodaten, wie schon betont, Ergebnisse kommunikativer Konstruktionen, die menschliche Handlungen und Wahrnehmungen mit den Leistungen materialer-technologischer Systeme verbinden. Auch wenn der Abbildungscharakter selbst eine kulturelle Leistung ist, zeichnen sich Videos dadurch aus, dass ihr mimetischer Charakter über die Kulturen hinweg sehr leicht erlernt und von den aufgezeichneten Akteuren zumeist (selbst von kleinen Kindern) leichthin erkannt werden kann.

Weil die aufgezeichneten sozialen Prozesse auf diese (eingeschränkte) Weise die Phänomene repräsentieren, auf die sich auch grundlegende sozialtheoretische Analyse-Begriffe (wie soziales, kommunikatives Handeln oder Interaktion)

18 Auch wenn dieses Prinzip schon für den Fotoapparat und die Filmkamera galt, hebt schon der Name »Video« das »Ich« hervor, das (mit dem Gerät) sieht. Es ist dabei sicherlich bezeichnend, dass sich dieses »Ich« nicht notwendig auf den aufnehmenden Menschen beziehen muss, unbedingt aber auf das Gerät. Auf den Zusammenhang zwischen der technischen Fokussierung und der gegenständlich-phänomenalen Fokussierung können wir hier nicht eingehen (vgl. dazu Knoblauch/Vollmer 2019).

beziehen, kann diesen Daten eine außergewöhnliche *Validität* zugeschrieben werden. Wissenschaftstheoretisch wird mit Validität bekanntlich die inhaltliche Übereinstimmung einer empirischen Messung oder Beobachtung mit einem Konzept bezeichnet: Videos können das, worauf sie gerichtet sind, so gut wiedergeben, dass es auch von Menschen wieder-erkannt wird, die es nur über die Videos wahrnehmen. Wir haben es daher genauer mit dem zu tun, was Roger E. Kirk und Samuel A. Miller (1986) als *apparant validity*, als offenkundige Validität, bezeichnen. Das, was analytisch bezeichnet wird, kann auch empirisch – etwa anhand von Transkripten – gezeigt werden.

Ein zentraler Grundzug der Videoanalyse besteht schließlich darin, dass sie die sozialen Abläufe, die sie aufzeichnet oder deren Aufzeichnungen sie sammelt, immer auch anhand dieser Aufzeichnungen untersucht. Sie ist also nicht eine »rekonstruktive Methode«, in der das, worum es geht, nur durch die sprachliche Referenz von z. B. Interviews oder Fragebögen präsent ist; sie zielt vielmehr darauf, dass das, was Gegenstand der Untersuchung ist, selbst auch audiovisuell abgebildet ist – und zwar im prozessualen Zeitverlauf. Daher geht es in der Analyse nicht etwa darum, wie Menschen z. B. ihre Gewalt oder ihren Rassismus (vgl. Kissmann; Egbert/Janssen; Meyer/Grauert/Oberzaucher i. d. B.) (später) deuten oder legitimieren, sondern sie fokussiert auf die gewalttätigen bzw. rassifizierenden Handlungen selbst; es geht auch nicht um die Frage, wie z. B. schulische Interaktionen (vgl. Mempel/Winterscheidt; Herrle/Proske/Puzicha/Zimmer; Jehle; Hamer/Helbig/Urban i. d. B.) später gedeutet werden, sondern um deren Ablauf.

4. Die Struktur dieses Sammelbands und die Systematik seiner Beiträge

Der Sammelband versammelt insgesamt 26 Beiträge und ist in vier Kapitel unterteilt: Kapitel I. *Theoretische und methodologische Überlegungen zu (audiovisuellen) Forschungsdaten in der empirischen Qualitativen Sozialforschung*, Kapitel II. *Videographische und medienanalytische Ansätze mit Vernakularen Videos natürlicher sozialer Situationen und inszenierter Inhalte* (1. Videographische Vernakulare Videos, 2. Medial-gattungsförmig gestaltete Vernakulare Videos), Kapitel III. *Interaktionsanalysen mit Ethnographischen Videos* (1. Interaktion in Paarbeziehungen, (Klein-)Gruppen, Kollektiven und der Öffentlichkeit, 2. Interaktion und Interaktivität mit Gegenständen, Technik und digitalen Technologien) und Kapitel IV. *Nutzungsformen: Forschungsvideos im wissenschaftlichen Film und Forschungsdatennachnutzung.* Neben diesem einleitenden Beitrag finden sich noch zwei weitere Beiträge in Kapitel I., die sich aus theoretischer und/oder methodologischer Perspektive mit audiovisuellen Forschungsdaten der qualitativen empirischen Sozialforschung beschäftigen:

- *Christian Meyer und Christian Meier zu Verl* beschäftigen sich mit einem Projekt von Thomas Luckmann und Peter Gross (1978–1982), das bereits Ende der 1970er Jahre im Rahmen eines Laborsettings die Erzeugung sozialwissenschaftlicher Daten durch die audiovisuelle Analyse von *face-to-face*-Interaktionen in Interviews untersuchte.
- *Michael Corsten* beleuchtet die Möglichkeiten, soziale Praxis durch Video zu dokumentieren, und diskutiert methodologische Fragestellungen zur Analyse solcher Videos. Ein Fokus seines Textes liegt darauf, Praxis als symbolisch und materiell strukturierte Realität zu verstehen.

In den Kapiteln II bis IV folgen dann 22 Beiträge, die die Grundlage für unsere folgende Systematisierung bilden, anhand derer wir sie den Kapiteln des Sammelbands zugeordnet haben und unten im Einzelnen vorstellen werden. Zum Abschluss des Bands findet sich am Ende von Kapitel IV außerdem ein Beitrag, der sich ohne eigene Empirie dem Gegenstand der Sekundäranalyse in der qualitativen Forschung zuwendet.

Mit der folgenden Systematik, die wir aus den empirischen Beiträgen zu diesem Sammelband gewonnen haben, wollen wir dazu beitragen, sehr genau zu bestimmen, was audiovisuelle Forschungsdaten der qualitativen Sozialforschung sind und wie sie in der zeitgenössischen Videographie und Videoanalyse behandelt, d. h. *erhoben*, *analysiert* und *genutzt* werden. Wir unterscheiden dabei drei Kategorien, nach denen sich die audiovisuellen Daten bzw. Forschungsansätze in diesem Band (nicht immer völlig trennscharf) zusammenfassen und differenzieren lassen:

- *Erhebungskontexte* (einschließlich der fokussierten Gegenstände),
- *Analysekontexte* (einschließlich der Triangulation und Weiterverarbeitung der Daten),
- *Nutzungsformen*.

Quer zu dieser Systematik verläuft die bereits erläuterte Typologie audiovisueller Forschungsdaten (vgl. Abb. 2): Wir unterscheiden zwischen *Ethnographischen* und *Vernakularen Videos* sowie zwischen *videographischen* und *medialen Gegenständen* audiovisueller Daten. Auf diese Weise lassen sich auch eine Reihe von *Vernakularen Videos* aus verschiedenen Quellen (Social Media und Offline-Kontexte) der Videographie im zuvor erläuterten Sinne zuordnen, während wir eine andere Sorte als Mediendaten betrachten, die, wie wir sehen werden, in der Forschung auch anders behandelt, d. h. anders erhoben und analysiert wird.

Zugleich mit der Darstellung unserer Systematisierung werden wir die in Kapitel II bis IV dieses Sammelbands zusammengestellten Beiträge in der Reihenfolge ihrer Wiedergabe einzeln kurz zusammenfassen. Dabei werden wir diese Beiträge entsprechend ihren Forschungsansätzen und Datensorten den

drei Kategorien unserer Systematik sowie den querverlaufenden Datentypen zuordnen. Diese Zuordnung erfolgt entsprechend charakteristischer Merkmale der Beiträge. Die Systematik selbst, die auf den genannten Kategorien (*Erhebungs*kontexte, *Analysekontexte*, *Nutzungsformen*) und Datentypen (*Ethnographische Videos* (gelb), *Vernakulare Videos* (grün)) basiert, haben wir in Abbildung 3 visualisiert.

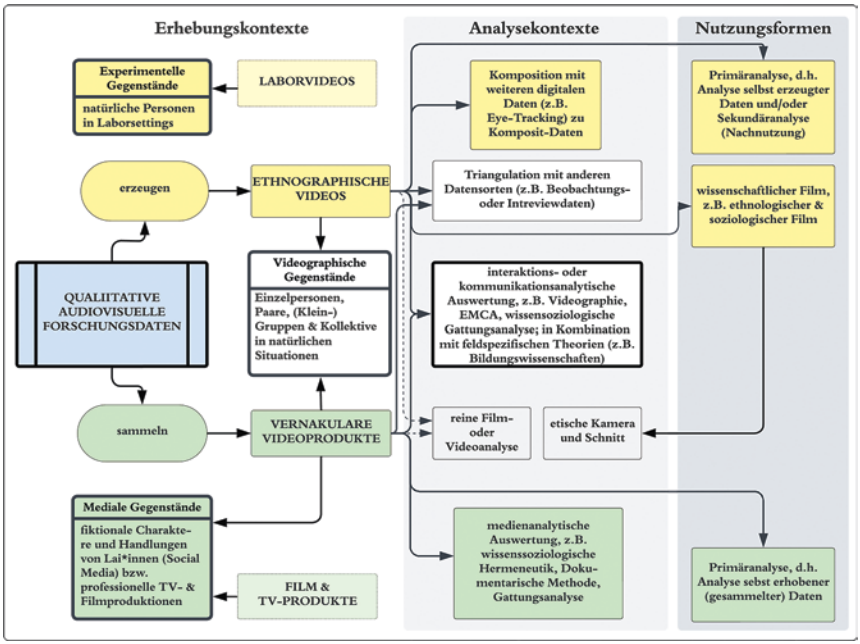


Abbildung 3: Systematisierung der qualitativen audiovisuellen Forschungsdaten der Beiträge in diesem Band

4.1 Erhebungskontexte

Als Erhebungsmethode findet sich in der qualitativen Videoanalyse, neben der *Erzeugung von Ethnographischen Videos*, auch die *Sammlung* von Daten. Diesen Datentyp wollen wir, wie bereits kurz erläutert, als *Vernakulare Videos* bezeichnen. Der Begriff »vernakular« soll dabei signalisieren, dass diese Videos außerhalb der Forschungswelt und ohne wissenschaftliche Intention erstellt werden. Damit eignet sich der Begriff, um die Differenz zu *Ethnographischen Videos*, die von Forscher*innen erzeugt werden, zu markieren. *Vernakulare Videos* können verschiedene Quellen umfassen – sie werden von privaten oder institutionellen Akteuren (z. B. in Arbeitskontexten) produziert und können *online* (z. B. auf Social Media) vorgefunden oder »im Feld« *offline* gesammelt werden.

Ein folgenreicherer Unterschied macht eine Differenzierung der *Vernakularen Videos* nötig (vgl. Abb. 2 und 3) und erlaubt dabei innerhalb des Typs unterschiedliche Datensorten zu differenzieren: Deutlich lassen sich *medialgattungsförmige* (bzw. gestaltete) von *videographischen Vernakularen Videos* unterscheiden. Medial-gattungsförmig gestaltete *Vernakulare Videos* bilden Gegenstände ab, die seltener dafür genutzt werden, ausschließlich interaktionsanalytische Fragestellungen zu fokussieren. Häufiger werden an diese Datensorte hermeneutische kunst- oder medienanalytische Analyseverfahren angelegt und auch das Handeln hinter der Kamera (Kamera-Handeln) und/oder Kamera- und Schnitttechniken analysiert.

Das Vorgehen erinnert dabei dann durchaus an die älteren Film- und Fernsehanalysen: gemeinsam sind beiden (zum Teil fiktive) *mediale Gegenstände*. *Vernakulare Videos* hingegen, die wir als *videographisch* bezeichnen, zeichnen sich, vergleichbar mit *Ethnographischen Videos*, durch *videographische Gegenstände* aus. Die Forschung mit diesem Material fokussiert das, was »soziale Situationen«, »soziale Interaktionen«, »Praxis« oder kommunikative »Handlungen« genannt wird, bei denen einzelne Personen oder Gruppen vor der Kamera so handeln, dass es von den Forschenden (audiovisuell) wahrnehmbar ist. Während Forscher*innen bei der Erzeugung von *Ethnographischen Videos* allerdings selbst im Feld zugegen sind und dabei auch das Problem der Reaktivität ihrer Aufzeichnung reflektieren können, gibt es im Fall von *videographischen Vernakularen Videos* nicht immer die Möglichkeit, genau zu bestimmen, ob eine Situation sich tatsächlich unter »natürlichen Gegebenheiten«, also z. B. ohne Planung oder externe Anweisungen zugetragen hat. Zur Kontrolle der Validität von videographisch analysierten *Vernakularen Videos* kann aber z. B. die Dokumentenanalyse dienen (vgl. Egbert/Janssen; Meyer/Grauert/Oberzaucher i. d. B.).

Videographische Vernakulare Videos finden sich in Kapitel II.1 dieses Sammelbands:

- *Ulrike T. Kissmann* beschäftigt sich in ihrem Beitrag mit der videographischen Analyse eines spontan-authentischen YouTube-Videos, das einen eskalierenden Konflikt zwischen einer weißen Ladendiebin und einem schwarzen Drogerieangestellten zeigt.
- *Christian Meyer, Hanna Grauert und Frank Oberzaucher* untersuchen ebenfalls den Einsatz von Social-Media-Videos in der Erforschung rassifizierender Interaktionen und der Eskalation von Konflikten. Als Daten dienen öffentlich zugängliche Videoaufzeichnungen, die zum Teil von Betroffenen selbst publiziert wurden.
- *Simon Egbert und Jasper Janssen* untersuchen die zunehmende Verwendung von Bodycams im Kontext der deutschen Polizeiarbeit und wie diese Videos zur Legitimierung polizeilicher Praktiken verwendet werden. Zu diesem

Zweck wenden sie eine Kombination von Interaktions- und Medienanalyse auf das in Folge eines »Leaks« einzige öffentlich zugängliche Bodycam-Video eines Polizeieinsatzes in Deutschland an.
- *Mina Godarzani-Bakhtiari und René Tuma* widmen sich ebenfalls der Polizeiarbeit und analysieren die Auswirkung der Videoaufzeichnungen von polizeilichen Vernehmungen auf die Kommunikation in der Vernehmungssituation.

Die *medial-gattungsförmigen Vernakularen Videos*, die im Folgenden versammelt sind, zeichnen sich durch eine Orientierung an Mediengattungen aus, die entweder im Kontext der Sozialen Medien neu entstanden sind (wie das ›Reel‹) oder aber sich bereits viel früher im TV (wie das Rap-Video) bzw. im Film entwickelt haben. In diesem Sinne zählen wir auch amateur-pornographische Videos (Lewandowski i. d. B.) zu den *medial-gattungsförmig* gestalteten *Vernakularen Videos*. Schon ihre Bezeichnung als »pornographisch« legt nahe, dass es sich hierbei nicht einfach um die Abbildung von Menschen in natürlichen sozialen Situationen handelt, sondern dass ihre spezifische Form sich an die sehr alte (mediale) Gattung der Pornographie anschließen lässt. Das Video ist in diesen Fällen selbst Repräsentant eines Phänomens und wird selbst zum Gegenstand, indem es die in ihm eingeschriebenen Techniken des Filmens, Beleuchtens, Darstellens, Schneidens etc. – in einem Wort: Film- bzw. Videotechniken – seiner Ersteller*innen vor und hinter der Kamera repräsentiert. Demgegenüber umfasst die *Sammlung* von *videographischen Vernakularen Videos*, wie auch die *Erzeugung Ethnographischer Videos*, zwar die Reflexion des Erzeugungsprozesses, einschließlich der Standortgebundenheit der (wissenschaftlichen) Erzeuger*innen, das Video selbst, seine (ästhetischen) Qualitäten sowie die Differenz zwischen abbildenden und abgebildeten Bildproduzent*innen (Bohnsack/Fritzsche/Wagner-Willi 2015, S. 14), wird dabei i. d. R. aber nicht zum Gegenstand der Analyse gemacht.

Medial-gattungsförmige Vernakulare Videos finden sich in Kapitel II.2 dieses Sammelbands:

- *Sven Lewandowski* erläutert in seinem Beitrag das Vorgehen bei der Kombination aus Video- und Videointeraktionsanalyse von amateurpornographischen Videos im Kontext der Sexualforschung. Die untersuchten *Vernakularen Videos* stellen (quasi-)autoethnographische Dokumente dar, die Lewandowski von den Ersteller*innen persönlich zum Zweck der wissenschaftlichen Forschung zur Verfügung gestellt wurden.
- *Katharina Bock* beschäftigt sich mit der Analyse von gattungsförmig gestalteten sozialmedial geteilten Musikvideoproduktionen Jugendlicher im Gangsta-Rap-Milieu. Sie trianguliert dabei verschiedene qualitative Methoden, darunter eine ethnosemantische Analyse von Songtexten, die Transkription und

Analyse von Videoaufnahmen sowie die Analyse von Symbolen und spezifischen Darstellungsformen.
- *Bernt Schnettler* widmet sich aus methodologischer Perspektive der Analyse von Videodaten, mit besonderem Fokus auf den Unterschied zwischen Feldvideos (von Forschenden im Rahmen eigener Feldforschung erstellte Videos) und Plattformvideos (Videos, die auf Plattformen wie Instagram, TikTok oder YouTube geteilt werden). Durch diesen Schwerpunkt schlägt der Text eine Brücke von Abschnitt II zu Abschnitt III dieses Sammelbands.

Mit den gattungsförmig gestalteten medialen Gegenständen der *Vernakularen Videos* begeben wir uns im zweiten Teil von Abschnitt II an eine der äußeren Grenzen der qualitativen Videoanalyse, die an dieser Stelle die medienanalytische Film- und Fernsehanalyse (Peltzer/Keppler 2015), die sich der Analyse medialer Gattungen widmet, berührt oder in sie übergeht.

Seitens der *Ethnographischen Videos*, als von Forscher*innen selbst erstellte Videos, die wir in den folgenden Kapiteln III bis IV dieses Bands zusammengestellt haben, wird die Grenze der qualitative Videoanalyse durch den Übergang zu Laborvideos (vgl. Meyer/Meier zu Verl; Harth i. d. B.) markiert (siehe Abb. 3), die, insofern die Laborumgebung nicht selbst als natürliche Situation in den Analysefokus gerückt wird, ihrerseits die Grenzen der qualitativen Videoanalyse überschreiten. Zwar sind auch Laborvideos von Forschenden selbst erzeugt, die »sozialen Situationen«, »sozialen Interaktionen«, die »Praxis« oder kommunikativen »Handlungen« die sie als Gegenstände abbilden, sind allerdings stark von den Forscher*innen beeinflusst, die sie initiieren, organisieren oder manipulieren. Die Arbeit mit Laborvideos berührt oder überschreitet dabei oft die Grenze zum Bereich der quantitativ-standardisierten Verfahren der Videoanalyse. Häufig werden solche Videos quantitativ ausgewertet, d. h. die in experimentellen Settings erzeugten audiovisuellen Aufzeichnungen werden anschließend systematisch kodiert und die Kategorien statistisch ausgezählt, um Hypothesen zu testen oder bestimmte Muster zu identifizieren (z. B. Bull 1985; Liszkowski et al. 2004; Tomasello et al. 2005).

4.2 Analysekontexte

Wie die Betrachtung der Einzelbeiträge dieses Bands bereits verdeutlichte, sind die methodologischen und theoretischen *Analysekontexte* audiovisueller Forschungsdaten ebenso vielschichtig wie ihre *Erhebungskontexte* und *Gegenstandsbereiche*. Auch hier lässt sich allerdings eine Systematisierung entlang der querverlaufenden Typologie audiovisueller Daten, zwischen *Ethnographischen Videos* und *videographischen* bzw. *medial-gattungsförmig* gestalteten *Vernakularen Videos* treffen: Insbesondere für die *medial-gattungsförmig* gestalteten

Vernakularen Videos erweisen sich, wie erläutert, auch stärker hermeneutisch orientierte Auswertungsverfahren geläufig, z. B. die *wissenssoziologische Hermeneutik* (Raab/Stanisavljevic 2018), die *Videohermeneutik* (Kissmann 2014) oder die *Dokumentarische Video- und Filmanalyse* (Bohnsack/Fritzsche/Wagner-Willi 2015). *Ethnographische Videos* (und s. o.: *videographische Vernakulare Videos*) hingegen, Videos also, die *videographische Gegenstände* fokussieren, wie Einzelpersonen, Paare, (Klein-)Gruppen oder Kollektive in natürlichen sozialen Situationen, in Interaktion und/oder interaktiv bei der Nutzung von Gegenständen, Technik oder komplexen Technologien, werden zumeist (obgleich diese Differenzierung nicht völlig trennscharf ist) mittels einer Kombination aus forschungsfeldspezifischen Theorien (z. B. Bildungswissenschaften) und interaktions- bzw. kommunikationsanalytischen Auswertungsverfahren analysiert, insbesondere *Videographie* (Tuma/Schnettler/Knoblauch 2013) oder *ethnomethodologische Konversationsanalyse* (Sacks/Schegloff/Jefferson 1974).

Ethnographische Videos finden sich in den Kapiteln III und IV dieses Sammelbands. Wir konnten sie aufgrund ihrer großen Zahl weiter unterteilen. In Kapitel III.1 finden sich Beiträge zu Daten, die (in dieser Reihenfolge) Interaktion in Paarbeziehungen, (Klein-)Gruppen und Kollektiven fokussieren:

- *Jo Reichertz und Anna Eva Nebowsky* analysieren in ihrem Beitrag die Kommunikation in Paarbeziehungen bei Demenzerkrankungen. Ziel ist es, die Veränderungen in der Kommunikation und die Auswirkungen dieser Diagnose auf interpersonelle Beziehungen zu untersuchen. Der Fokus liegt auf Interaktionen innerhalb von Paarbeziehungen und mit den anwesenden Forscher*innen.
- *Jonas Kramer und Sarah Hitzler* widmen sich in Familien der Entwicklung sog. projektiver Gattungen, die sich transsituativ erstrecken. Die Autor*innen untersuchen dabei, wie Wissensbestände über Zeit und verschiedene Situationen hinweg rekursiv hergestellt und reproduziert werden. Es geht insbesondere um die kommunikative Verfertigung von Zukunft in alltäglichen Interaktionen und die longitudinalen Prozesse, die hierbei eine Rolle spielen.
- *Theresa Vollmer* beschäftigt sich mit dem Forschungsfeld des gemeinsamen Musizierens eines Streich-Ensembles und eines Chores. Der Text behandelt die Herausforderungen und Potenziale der Videoanalyse des Musizierens. Ein methodologischer Schwerpunkt liegt auf der Transkription entsprechender audiovisueller Daten.
- *Caterina Mempel und Jenny Winterscheid* erläutern die Erhebung, Transkription und Analyse audiovisueller Daten im Kontext der sprachlichen Begabungsförderung im Klassenverband. Im Zentrum ihrer Forschung steht die Förderung von sprachlich-literarischen Kompetenzen bei leistungsstarken und potenziell besonders leistungsfähigen Schüler*innen im Deutschunterricht.

- *Ajit Singh* widmet sich der Methode der verdeckten Videographie im öffentlichen Raum und ihrer ethischen Implikationen. Der Text untersucht die Rolle verdeckter Videoaufnahmen im Rahmen der Pandemieforschung. Er diskutiert, welche besonderen Herausforderungen diese Form der Datenerhebung für die Forschungsethik und die Analyse darstellt.

Bei den Analysekontexten zeigt sich, dass nur wenige Daten der hier versammelten Beiträge aus reinen Datenerhebungsprojekten (explizit: Singh i. d. B.) stammen. Die reine Videoanalyse, die ohne Hinzuziehung von anderen Datensorten arbeitet, stellt noch eine seltene Ausnahme in der Videoanalyse dar. Der Großteil der hier vertretenen Ansätze basiert hingegen auf einer für die qualitative empirische Sozialforschung typischeren Kombination verschiedener Datensorten, der Datentriangulation (Flick 2011): Zur Analyse werden i. d. R. weitere Datensorten hinzugezogen, um die Sicht auf den Gegenstand perspektivisch und im Kontext der jeweiligen Forschungsfrage zu vervollständigen.

Hierbei kann, wie sich auch an den versammelten Beiträgen anschaulich verdeutlichen lässt, das gesamte Spektrum von (auto-) (Thiel-Woznica i. d. B.) ethnographischen Methoden zum Einsatz kommen, insbesondere die Teilnehmende Beobachtung bzw. Beobachtende Teilnahme (Reichertz/Nebowsky i. d. B.), Ethnographische Gespräche (ebd.) bzw. verschiedene Interviewformate (Lewandowski i. d. B.), die Dokumentenanalyse (vgl. Meyer/Grauert/Oberzaucher i. d. B.) etc. Gleichzeitig sind aufwändige Transkriptions- (Niediek; Vollmer i. d. B.) und Annotationsverfahren (Gruber/Rizzolli; Kramer/Hitzler i. d. B.), gerade im Kontext der Analyse von Interaktionen, in die zusätzlich technische Geräte und technologische Systeme involviert sind, ein häufig zu beobachtendes Merkmal der Analyse. Transkripte und Annotationen dienen dazu, die Daten aufzubrechen, gemeinsame Datensitzungen durchzuführen sowie der Repräsentation erhobener Daten und von Forschungsergebnissen in wissenschaftlichen Publikationen.

Ethnographische Videos, die Interaktion und Interaktivität mit Gegenständen, Technik und digitalen Technologien fokussieren, finden sich in Kapitel III.2:

- *Imke Niediek* erläutert in ihrem Beitrag ihre Forschung zur atypischen Kommunikation und die Analyse von Mehrpersonengesprächen, die durch die Verwendung von elektronischen Kommunikationshilfsmitteln (sogenannten »Talkern«) unterstützt werden. Der Text präsentiert die mehrstufige Erzeugung von Videotranskriptionen und diskutiert, wie diese genutzt werden können, um interaktionsanalytische Studien im Bereich der Unterstützten Kommunikation durchzuführen.

Neben dem Einsatz komplexer Transkriptionsverfahren zeigt sich, dass bei der Forschung mit *Ethnographischen Videos*, insbesondere dort, wo es um die

Analyse der Interaktivität zwischen Menschen und technischen Systemen geht, das *Ethnographische Video* häufig mit weiteren digitalen Datensorten zu einem *Komposit-Datum* verbunden wird. Unter *Komposit-Daten* verstehen wir solche Daten, die zum Zweck ihrer Analyse mit weiteren (digitalen) Daten verbunden werden. Hierzu zählen die Integration von Bildschirmaufzeichnungen (»Screensavings«) sowie von Eye Tracking- oder Full Body Tracking-Daten in per Videokamera aufgezeichnete audiovisuelle Daten. Die entstehenden Daten sind überaus vielschichtig und komplex:

- *Matthias Herrle, Matthias Proske, Aline Puzicha und Anne Zimmer* thematisieren in ihrem Beitrag die mikroethnographische Analyse von Interaktions- und Wissensbildungsprozessen im Kontext von Tablet-gestütztem Unterricht. Die erzeugten *Ethnographischen Videos* werden mit *Screenrecordings* der Tablets sowie mit Verbaltranskripten von *Audioaufzeichnungen* an den Arbeitsplätzen der Schüler*innen zusammengefügt. Zur Ergebnisdarstellung werden außerdem Skizzen hergestellt, die die prozessrelevanten Phänomene abbilden, und gegebenenfalls mit Verbaltranskripten kombinieren.
- *Isabel Neto Carvalho* erläutert ebenfalls einen medienpädagogischen Ansatz. Die Daten umfassen *Ethnographische Videos*, die mit *Eye-Tracking* der Akteure kombiniert werden, um deren visuellen Fokuspunkte während der Interaktivität mit digitalen und physischen Artefakten zu analysieren. Zusätzlich werden Beobachtungsdaten genutzt, um die Dynamik der Mensch-Maschine-Interaktionen im Klassenzimmer zu analysieren.
- *Luise Reitstätter, Seda Pesen und Dirk vom Lehn* thematisieren ihre Analyse des Sehens und sozialer Interaktion im Kunstmuseum auf Grundlage einer Kombination von *mobilem Eye-Tracking* (MET) und *Ethnographischen Videos*. Ziel ist es, das Zusammenspiel von visueller Aufmerksamkeit und sozialer Interaktion während des Besuchs eines Kunstmuseums zu analysieren. Die Datensätze bestehen aus MET-Aufnahmen, Videoaufnahmen, Fragebögen und Stimulus-gestützten Interviews mit Museumsbesucher*innen.

Durch ihren spezifischen Status als genuin digitale Daten werfen *Komposit-Daten* die Frage nach ihrer Ontologie auf. Insbesondere sog. »Mixed-Reality-Videodaten« (Harth i. d. B.) haben, was nicht übersehen werden sollte, einen anderen ontologischen Status als klassische Videodaten. Während Letztere i. d. R. reale Ereignisse im physischen Raum und in linearer Zeit erfassen, sind genuin digitale Daten, wie sie bei Full Body-Scans und/oder der Aufzeichnung von Interaktivität im virtuellen Raum entstehen, das Ergebnis einer Verschränkung von physischer Realität und digitaler, virtueller Elemente. Das bedeutet, dass sie nicht nur physische Handlungen und Interaktionen im natürlichen Raum dokumentieren, sondern auch digitale Inhalte und Inter- und *Intraaktivität* (Knoblauch 2017, S. 186 f.) in der virtuellen Realität enthalten. Der Gegenstand muss also in einer

anderen Umgebung, dem hybriden Raum, verortet werden. Daher funktionieren auch die Interaktionen dort anders als zwischen menschlichen Akteuren in der rein physischen Welt, wie insbesondere zwei Beiträge dieses Sammelbands verdeutlichen:

- *Marcel Thiel-Woznica* erläutert in seinem Beitrag ein autovideographisches Vorgehen zur Analyse von digitalen Spielpraktiken. Die audiovisuellen Daten bestehen aus *Ethnographischen Videos*, die während des Spielens erzeugt wurden. Die *Komposit-Daten* zeigen schließlich sowohl das Bildschirmgeschehen als auch die physischen Handlungen des Spielers. Der Text untersucht, wie digitale Spielpraktiken im Kontext von Inter- und Intraaktivität aufgezeichnet und analysiert werden können.
- *Jonathan Harth* widmet sich der Analyse von Mensch-Agent-Interaktionen in Mixed-Reality-Situationen. Er untersucht Interaktionsmuster zwischen menschlichen Nutzer*innen und verkörperten Konversationsagenten. Der Text behandelt *Komposit-Daten*, die durch »Mixed-Reality-Videographie« erhoben wurden. Die Aufnahmen integrieren sowohl physische Handlungen der Nutzer*in als auch die virtuellen Reaktionen des Agenten (Intraaktion).

Insbesondere anhand dieser letzten beiden Beiträge wird die ontologische Differenz von *Komposit-Daten* zu anderen audiovisuellen Forschungsdaten deutlich, eben weil die aufgezeichneten Gegenstände in zwei unterschiedlichen »Realitäten« stattfinden: *physisch und virtuell*. In beiden Ansätzen sind die Interaktionen von Menschen und virtuellen Agenten bzw. Avataren zentral. Letztere haben allerdings keinen festen ontologischen Status als »real«, sondern existieren nur im digitalen Speichermedium, wodurch die menschlichen Aktionen und Reaktionen sowohl auf physische als auch auf digitale Entitäten gerichtet sind. »Mixed-Reality-Daten« dokumentieren daher nicht nur physische Vorgänge, sondern auch virtuelle Handlungen und Interaktivität. Ihre Ontologie ist dadurch komplexer, nicht zuletzt, weil auch die Intraaktivität von technologischen Systemen selbst in Betracht gezogen werden muss.

Im Zusammenhang mit den Analysekontexten möchten wir noch eine weitere Be-Handlung von *Ethnographischen Videos* ansprechen, die zugleich (siehe Abschnitt 4.3) eine besondere Nutzungsform darstellt. Sie findet sich im wissenschaftlichen Film, wo die Vertextlichung der Analyseergebnisse entfällt, insofern stattdessen die Analyse durch Selektions- und Schnittentscheidungen am audiovisuellen Material stattfindet und die Analyseergebnisse daher bereits im fertigen Medienprodukt enthalten sind, ohne zwingend in Textform transformiert zu werden.

Während alle hier vertretenen Beiträge zur qualitativen audiovisuellen Forschung sich dadurch auszeichnen, dass die im Video fixierten Prozesse von den

Forschenden verstanden werden müssen,[19] lassen sich dabei grob zwei unterschiedliche Ansätze unterscheiden: a) solche, die den Großteil ausmachen, bei denen die Erzeugung bzw. Sammlung der Daten räumlich und zeitlich von der Analyse bzw. Interpretation abgetrennt werden und b) die, wie der ethnographische und der soziologische Film, die sich dadurch auszeichnen, dass Erhebung und Interpretation bzw. Analyse der Daten miteinander verknüpft sind.[20] In diesen Fällen geht die Deutung der Gegenstände schon in die hochgradige Selektion dessen ein, was aufgenommen bzw. herausgeschnitten wird (vgl. Miko-Schefzig; Gruber/Rizzolli i. d. B.). Wie beim dokumentarischen Film zielt die Erzeugung des Videos dabei auf die Gestaltung eines Medienproduktes, das über die Forschenden hinaus an eine transdisziplinäre Öffentlichkeit gerichtet ist.

Auch wenn damit die vom ursprünglichen Videogerät geprägte Trennung von Erhebung und Analyse bzw. Interpretation aufgehoben ist, weisen auch diese auf *Ethnographischen Videos* basierenden wissenschaftlichen Filme, neben einer ikonographischen, d. h. einer bildlich-gegenständlichen Gemeinsamkeit mit anderen videographischen Daten, auch die weiter oben bereits erläuterte Gemeinsamkeit auf, die sie allgemein als »videographisch« kennzeichnet: Es geht auch hier um Aufzeichnungen dessen, was wir als »natürliche soziale Situationen« bezeichnen, wobei das, was »sozial« ist bzw. als sozial betrachtet wird, natürlich stets und nicht allein hier von theoretischen und analytischen Perspektiven abhängt.

In den Bereich wissenschaftlicher Film fallen in diesem Sinne videographische Daten, bei denen die audiovisuelle Erhebungs- und Repräsentationsweise als »Film« bereits Konstrukte zweiter Ordnung enthalten, um so unmittelbar wissenschaftliches Verstehen und Erklären zu erzielen. Diese Form der Nutzung von audiovisuellen Daten zeigt sich z. B. auch in der am ethnographischen Film orientierten Methode der *Kamera-Ethnographie* von Bina Mohn (2022) in der (in Anlehnung an Geertz) das Konzept des »dichten Zeigens« eine zentrale Rolle spielt.

4.3 Nutzungsformen

Somit stellt der wissenschaftliche Film, zugleich mit seiner Besonderheit im Rahmen des Analysekontexts, auch eine besondere Nutzungsform von *Ethnographischen Videos* innerhalb der qualitativen Videoanalyse dar, insofern das Videomaterial nicht Grundlage für eine Analyse ist, die sich in bzw. anhand von handelnd

19 Im Gegensatz zu standardisierten Verfahren, die sich mit der theoretisch abgeleiteten Codierung audiovisueller Daten beschäftigen (Nassauer/Legewie 2022).

20 Während die Interpretation ein alltägliches Verständnis der aufgezeichneten Prozesse (*was sagen sie, was tun sie, was geschieht hier?*) bezeichnet, reden wir von Analysen, wenn die Daten mit theoretisch begründeten analytischen Konzepten verbunden werden (Interaktion, Sequenz, Turn etc.).

erzeugten Datentransformation z. B. Transkripten entfaltet, mittels derer die Forschungsergebnisse für andere wahrnehmbar gemacht werden. Vielmehr stellt das Material selbst, wie bereits erläutert, als fertiges Medienprodukt, das analytische Forschungsergebnis dar. Es ersetzt damit, neben der nachgelagerten Analyse, insbesondere die textuelle oder anhand lediglich von Momentaufnahmen illustrierte Repräsentation der Forschungsergebnisse. Kameraführung und Bildschnitt, nicht unähnlich im Dokumentarfilm, verfolgen dabei die Absicht, nicht zuletzt unter Einsatz von ästhetischen Prinzipien, audiovisuelle Beiträge zur (transdisziplinären) Wissensvermittlung zu erzeugen.

Neben der besonderen Nutzungsform des wissenschaftlichen Films, lässt sich noch eine weitere Form unter dieser Kategorie unterscheiden: die Nachnutzung. Zunehmend relevant ist die Sekundäranalyse von qualitativen Daten auch im Bereich der audiovisuellen Sozialforschung mit *Ethnographischen Videos*.[21] Das zeigt eine Reihe von Beiträgen in diesem Band. Dabei erweist sich auch die wachsende Bedeutung von Forschungsdatenzentren (FDZ), da, während früher Daten häufig von Forscher*innen untereinander geteilt wurden, die hier versammelten Beiträge die Nutzung von FDZ als institutionalisierte Intermediäre belegen, die nicht zuletzt die Berücksichtigung rechtlicher und forschungsethischer Aspekte des Datenteilens garantieren. Besonders in der Bildungsforschung (Jehle; Hamer/Helbig/Urban i. d. B.), aufgrund des spezialisierten FDZ-Bildung am DIPF (Frankfurt), spielt die Nachnutzung *Ethnographischer Videos* inzwischen eine große Rolle. Mit der Zusammenarbeit von Qualiservice (Universität Bremen) und dem Fachinformationsdienst der Sozial- und Kulturanthropologie (FID-SKA) entwickelt aktuell ein FDZ auch eine Lösung für die Nachnutzung audiovisueller Forschungsdaten aus der ethnologischen Forschung (Gruber/Rizzolli i. d. B.). Trotz dieser Entwicklungen erweist sich allerdings, wie zum Schluss und ohne den Bezug auf konkrete Daten gezeigt wird, dass die Kultur des Datenteilens in der qualitativen Forschung noch am Anfang steht (Wilke i. d. B.).

21 Die Nachnutzung von Korpora *Vernakularer Videos* scheint hingegen nicht etabliert. Dies mag zum einen an rechtlichen, vor allem urheberrechtlichen Aspekten liegen, die eine Archivierung und damit eine systematische Nachnutzung erschweren (dies geht mit der Problematik einher, dass diese Videos i. d. R. nicht aus dem Netz heruntergeladen und archiviert werden dürfen, womit die Gefahr eines Datenverlusts mir schwerwiegenden Herausforderungen für die Belegpflicht im Sinne guter wissenschaftlicher Praxis einhergehen). Vor allem aber ist dies sicherlich mit der Ubiquität und der öffentlichen Zugänglichkeit vor allem von Plattformvideos zu begründen, wodurch, anders als bei *Ethnographischen Videos*, ein Erfordernis der Nachnutzbar- und damit Zugänglichmachung entfällt. Schließlich liegt auch kein besonderes forschungspolitisches Interesse vor, da, anders als bei der Erzeugung von *Ethnographischen Videos*, bei der Produktion von *Vernakularen Videos* keine staatlich zur Verfügung gestellten und in der Forschungscommunity umkämpften Fördermittel eingesetzt werden.

Ethnographische Videos im wissenschaftlichen Film und in der Nachnutzung finden sich in Kapitel IV dieses Sammelbands:

- *Katharina Miko-Schefzig* analysiert in ihrem Beitrag die unterschiedlichen Herangehensweisen des soziologischen Films und der *Videographie* (Tuma/Schnettler/Knoblauch 2013) und welche Synergien oder Begrenzungen sich aus der Kombination dieser beiden Methoden ergeben können. Die zentrale Frage ist, inwiefern Filmtechniken neue Wege zur Wissensproduktion eröffnen können und wie die visuelle Kommunikation in der Soziologie weiterentwickelt werden kann. Die Daten stammen aus soziologischen Filmprojekten.
- *Martin Gruber und Michaela Rizzolli* widmen sich *Ethnographischen Videos*, die im Rahmen eines ethnographischen Dokumentarfilms erhoben wurden, der im Jahr 2015 in Kamerun gedreht wurde. Die Aufnahmen umfassen ethnographische Beobachtungen und Interviews. Der Text setzt einen Schwerpunkt auf die Archivierung und Nachnutzung ethnographischer Filmdaten. Die Rohdaten wurden für das FDZ Qualiservice aufbereitet und zur Nachnutzung archiviert. Die Autor*innen diskutieren die Herausforderungen bei diesem Prozess, bei dem spezielle Anforderungen an den Datenschutz, die Segmentierung und Annotation der Daten berücksichtigt werden mussten.
- *May Jehle* erläutert ihre bildungshistorische Forschung anhand der Nachnutzung von *Ethnographischen Videos* aus pädagogischen Fortbildungskontexten und der Unterrichtsforschung. Der Text befasst sich mit der Untersuchung von Unterrichtsvideographien aus der DDR und der Bundesrepublik, die für bildungshistorische Forschung genutzt werden, um die Entwicklung spezifischer Unterrichtspraktiken und deren Transformation im Kontext politischer und pädagogischer Reformen zu analysieren.
- *Anna Hamer, Jana Helbig und Michael Urban* setzen sich methodologisch mit der Qualitativen Sekundäranalyse von Unterrichtsvideos auseinander. Dabei fokussieren sie auf die De- und Rekontextualisierung von Forschungsdaten. Der Text behandelt die methodologischen und forschungspraktischen Herausforderungen der Verwendung von *Ethnographischen Videos* aus Unterrichtssituationen für Sekundärforschungen, insbesondere die Bedeutung des Umgangs mit Kontextinformationen bei der Sekundäranalyse.
- *René Wilke* schließlich widmet sich den Innovationspotenzialen durch die Sekundäranalyse sowohl für die qualitative Forschung als auch als Mittel und Gegenstand der qualitativen Methodenlehre. Vor dem Hintergrund der Digitalisierung werden dabei insbesondere FDZ als Ressource beleuchtet. Anhand von Studien und Publikationen wird gezeigt, welche Rolle die Sekundäranalyse bereits jetzt für die Forschung und Methodenausbildung spielt. An Beispielen wird erläutert, welche konkreten Möglichkeiten die Sekundäranalyse bietet.

Literatur

Andorfer, Peter (2015): Forschen und Forschungsdaten in den Geisteswissenschaften. DARIAH-DE Working Papers. Göttingen: DARIAH-DE.

Apel, Jochen (2011): Daten und Phänomene: Ein Beitrag zur wissenschaftstheoretischen Realismusdebatte. Berlin: De Gruyter.

Bateson, Gregory/Mead, Margaret (1942): Balinese Character: A Photographic Analysis. New York: New York Academy of Science.

Birdwhistell, Ray L. (1952): Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture. Washington, DC: Department of State, Foreign Service Institute.

Bogen, James/Woodward, James (1992): Observations, Theories and the Evolution of the Human Spirit. In: Philosophy of Science, 59, 4, S. 590–611.

Bohnsack, Ralf/Fritzsche, Bettina/Wagner-Willi Monika (Hrsg.) (2015): Dokumentarische Video- und Filminterpretation. Methodologie und Forschungspraxis. 2., durchgesehene Auflage. Opladen: Barbara Budrich.

Bull, Peter (1985): Body Movement and Emphasis in Speech. In: Journal of Nonverbal Behavior 9, 3, S. 169–187.

Daston, Lorain/Galison, Peter (2007): Objektivität. Frankfurt am Main: Suhrkamp.

DFG (2015): Leitlinien zum Umgang mit Forschungsdaten. https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf (17.10.2024).

Dinkelacker, Jörg/Herrle, Matthias (2009): Erziehungswissenschaftliche Videographie. Eine Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Ekman, Paul/Friesen, Wallace (1969): The Repertoire of Nonverbal Behavior: Categories, Origins, Usage and Coding. In: Semiotica, 1, S. 49–98.

Erickson, Frederick/Shultz, Jeffrey (1982): The Counselor as Gatekeeper: Social Interactions in Interviews. New York: Academic Press, S. xiv + 263.

Flick, Uwe (2011): Triangulation. Eine Einführung. 3., aktualisierte Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

Goodwin, Charles (1981): Conversational Organization: Interaction Between Speakers and Hearers. New York: Academic Press.

Grimshaw, Allen D. (1982): Comprehensive Discourse Analysis: An Instance of Professional Peer Interaction. In: Language in Society, 11, 1, S. 15–47.

Gross, Peter (1985): Theorie und sozialwissenschaftliche Daten. Unveröffentlichtes Manuskript (DE-SAK-LUCK-F1.084-01). Sozialwissenschaftliche Archiv Konstanz (SAK): Universität Konstanz.

Heath, Christian (1997): The Analysis of Activities in Face to Face Interaction Using Video. In: Silverman, David (Hrsg.): Qualitative Research. Theory, Method and Practice. Los Angeles: Sage, S. 183–200.

Heath, Christian (1986): Body Movement and Speech in Medical Interaction. Cambridge: Cambridge University Press.

Heath, Christian/Hindmarsh, John/Luff, Paul (2010): Video in Qualitative Research. Los Angeles: Sage.

Hoyningen-Huene, Paul (1987): Context of Discovery and Context of Justification. In: Studies in History and Philosophy of Science. 18, S. 501–515.

Kirk, Jerome/Miller, Marc L. (1986): Reliability and Validity in Qualitative Research. London.

Kissmann, Ulrike T. (2014): Die Sozialität des Visuellen. Fundierung der hermeneutischen Videoanalyse und materiale Untersuchungen. Weilerswist: Velbrück Wissenschaft.

Kissmann, Ulrike T. (2009): Video Interaction Analysis: Methods and Methodology. Frankfurt am Main: Peter Lang.

Knoblauch, Hubert (2021): Reflexive Methodology and the Empirical Theory of Science. In: Historical Social Research, 46, 2, S. 59–79.

Knoblauch, Hubert (2017): Die kommunikative Konstruktion der Wirklichkeit. Wiesbaden: Springer VS.

Knoblauch, Hubert (2004): Video-Interaktionsanalyse. In: Sozialer Sinn, 1, S. 123–139.

Knoblauch, Hubert (2003): Qualitative Religionsforschung. Religionsethnographie in der eigenen Gesellschaft. Paderborn: Ferdinand Schöningh (UTB).

Knoblauch, Hubert (2001): Fokussierte Ethnographie. In: Sozialer Sinn, 1, S. 123–141.

Knoblauch, Hubert/Schnettler, Bernt/Raab, Jürgen/Soeffner, Hans-Georg (Hrsg.) (2006): Video Analysis. Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology. Frankfurt am Main: Peter Lang.
Knoblauch, Hubert/Tuma, René/Schnettler, Bernt (2014): Videography: Introduction to Interpretive Video-Analysis of Social Situations. Frankfurt am Main: Peter Lang.
Knoblauch, Hubert/Vollmer, Theresa (2019): Ethnographie. In: Baur, Nina/Blasius, Jörg (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 599–617.
Knorr-Cetina, Karin (1984): Die Fabrikation von Erkenntnis. Frankfurt am Main: Suhrkamp.
Liszkowski, Ulf/Carpenter, Malinda/Henning, Anne/Striano, Tricia/Tomasello, Michael (2004): Twelve-month-olds Point to Share Attention and Interest. In: Dev Sci, 7, 3, S. 297–307.
Luckmann, Thomas (2006): Some Remarks on Scores in Multimodal Sequential Analysis. In: Knoblauch, Hubert/Schnettler, Bernt/Raab, Jürgen/Soeffner, Hans-Georg (Hrsg.): Video Analysis. Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology. Frankfurt am Main: Peter Lang, S. 29–34.
Luckmann, Thomas (1980): Philosophie, Sozialwissenschaft und Alltagsleben. In: Luckmann, Thomas: Lebenswelt & Gesellschaft. Grundstrukturen und geschichtliche Wandlungen. Paderborn: Schöningh, S. 9–55.
Meier zu Verl, Christian (2018): Eine wissenschaftssoziologische Studie zur methodologischen Praxis der Ethnografie. Stuttgart: Metzler.
Mittelstrass, Jürgen (2024): Enzyklopädie Philosophie und Wissenschaftstheorie. 8 Bde. 2. Auflage. Stuttgart: Metzler.
Mohn, Bina E. (2022): Kamera-Ethnographie: Schauen, Sehen und Wissen filmisch gestalten. In: Geimer, Alexander/Heinze, Carsten/Winter, Rainer (Hrsg.): Handbuch Filmsoziologie. Wiesbaden: Springer VS, S. 591–611.
Moritz, Christine (2018): »Well, it depends …«: Die mannigfaltigen Formen der Videoanalyse in der Qualitativen Sozialforschung. Eine Annäherung. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 3–37.
Moritz, Christine/Corsten, Michael (2018): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS.
Nassauer, Anne/Legewie, Nicolas (2022): Video Data Analysis. How to Use 21 st Century Video in the Social Sciences. Los Angeles: Sage.
Peltzer, Anja/Keppler, Angela (2015): Die soziologische Film- und Fernsehanalyse: Eine Einführung, Berlin: De Gruyter.
Pittenger, Robert E./Hockett, Charles F./Danehy, John J. (1960): The First Five Minutes: A Sample of Microscopic Interview Analysis. Oxford: Paul Martineau.
Raab, Jürgen (2002): ›Der schönste Tag des Lebens‹ und seine Überhöhung in einem eigenwilligen Medium. Videoanalyse und sozialwissenschaftliche Hermeneutik am Beispiel eines professionellen Hochzeitsvideofilms. In: Sozialer Sinn, 3, S. 469–495.
Raab, Jürge/Stanisavljevic, Marija (2018): Wissenssoziologische Videohermeneutik. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 57–71.
Reichert, Ramón (2007): Im Kino der Humanwissenschaften. Studien zur Medialisierung wissenschaftlichen Wissens. Bielefeld: transcript.
Reichertz, Jo (2013): Gemeinsam interpretieren. Die Gruppeninterpretation als kommunikativer Prozess. Wiesbaden: VS Verlag für Sozialwissenschaften.
Sacks, Harvey/Schegloff, Emanuel/Jefferson, Gail (1974): A Simplest Systematics for the Organization of Turn-taking in Conversation. In: Schenkein, Jim (Hrsg.): Studies in the Organization of Conversational Interaction. New York, S. 5–56.
Scheflen, Albert E. (1965): Quasi-courtship Behavior in Psychotherapy. In: Psychiatry: Journal for the Study of Interpersonal Processes, 28, 3, S. 245–257.
Thiel, Thomas (2003): Film und Videotechnik in der Psychologie. Eine erkenntnistheoretische Analyse mit Jean Piaget, Anwendungsbeispiele aus der Kleinkindforschung und ein historischer Rückblick auf Kurt Lewin und Arnold Gsell. In: Keller, Heidi (Hrsg.): Handbuch der Kleinkindforschung. 3., korrigierte und überarbeitete Auflage. Bern, Göttingen, Toronto, Seattle: Hans Huber, S. 649–708.

Tomasello, Michael/Carpenter, Malinda/Call, Josep/Behne, Tanya/Moll/Henrike (2005): Understanding and Sharing Intentions: The Origins of Cultural Cognition. In: Behavioral and Brain Sciences, 28, 5, S. 675–691.
Tuma, René (2016): Videoprofis im Alltag. Die kommunikative Vielfalt der Videoanalyse. Wiesbaden: Springer.
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.
Wilke, René (2022): Wissenschaft kommuniziert. Eine wissenssoziologische Gattungsanalyse des akademischen Group Talks am Beispiel der Computational Neuroscience. Wiesbaden: Springer.

Daten über Daten über Daten

Das wissenschaftssoziologische Projekt »Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten« an der Universität Konstanz (1978–1982)

Christian Meyer und Christian Meier zu Verl

Zusammenfassung: Der Aufsatz rekonstruiert die wissenschafts- und erkenntnistheoretischen, methodologischen und methodischen Arbeiten des Forschungsprojekts »Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten« (kurz »Daten über Daten«-Projekt), das von 1978 bis 1982 unter der Leitung von Thomas Luckmann und Peter Gross an der Universität Konstanz durchgeführt wurde. Das wissenschaftssoziologische »Pionierprojekt« (Jörg Bergmann), das als Parallelaktion zu den zeitgleichen »Laboratory Studies« gelten kann, untersuchte vor dem Hintergrund phänomenologischer Theorien (Husserl, Schütz), wie alltägliche Beschreibungen und sozialwissenschaftliche Daten in Relation zueinander entstehen. Ziel war es, den Prozess der Datenkonstruktion in den Sozialwissenschaften offenzulegen und dabei zu zeigen, dass Daten keine neutralen, »gefundenen« Objekte sind, sondern durch komplexe alltägliche und wissenschaftliche Praktiken entstehen. Um parallel auch ein interdisziplinäres Notations- und Transkriptionssystem zur adäquaten Beschreibung sozialer Interaktionen zu entwickeln, wurden im Projekt verschiedene Interaktionen audiovisuell aufgezeichnet und analysiert. Aufzeichnung und Analyse wurden wiederum aufgezeichnet und analysiert, um Erhebungs- und Analysekontexte der Datenkonstitution und -prozessierung sowie Szenarien der Datenverwendung, abstrahiert vom Gegenstand, selbst empirisch zu erforschen. Das Projekt legte einen wichtigen Grundstein für die Etablierung und Professionalisierung qualitativer Methoden im deutschsprachigen Raum und die Entwicklung neuer methodischer Verfahren in den Sozialwissenschaften.

Schlüsselwörter: Thomas Luckmann, Wissenschaftssoziologie, Wissenschaftstheorie, Erkenntnistheorie, qualitative Methoden, Notation, Transkription, »Protodaten«

1. Einleitung

Die methodologische Selbstreflexion in den Sozialwissenschaften, wie sie der Band weiterführt, hat eine lange Tradition. Nicht nur theoretische Diskussionen und empirie-praktische Erfahrungsbestände, sondern auch medientechnische Innovationen wie die Erfindung des Tonaufnahmegeräts oder des Smartphones hatten regelmäßig großen Einfluss auf praktische sozialwissenschaftliche Forschungsprozesse und stellten die Wissenschaft immer wieder vor neue theoretische Herausforderungen, etwa, zu bestimmen, was überhaupt als wissenschaftliches Datum begriffen werden kann.

Ein bisher noch (zu) wenig dokumentiertes und diskutiertes wissenschaftssoziologisches Forschungsprojekt, um das es im Text gehen soll, ist das von der Fritz-Thyssen-Stiftung finanzierte »Pionierprojekt« (Bergmann 2016) »Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten«, dessen Kurzformel unter den Beteiligten »Daten über Daten« lautete (vgl. auch Tuma et al. 2013, S. 28; Meier zu Verl 2018, S. 3–9). Das Projekt wurde von Thomas Luckmann in Zusammenarbeit mit Peter Gross von 1978 bis 1982 an der Universität Konstanz durchgeführt. Es kann als Teil der innovativen wissenschaftssoziologischen Bewegung der 1970er und 1980er Jahre gelten, als in mehreren Parallelaktionen damit begonnen wurde, nach soziologischen Erklärungen für wissenschaftliche Ideen und Ideenentwicklungen zu suchen. Das Ziel dieser Bewegung war, soziale Erklärungen nicht nur für fehlerhafte Überzeugungen und Annahmen zu liefern, sondern soziale Faktoren als Einflussfaktoren für alle, d. h. auch wissenschaftlich korrekte Überzeugungen, zu identifizieren. Dabei wurden auch außerhalb des menschlichen Bewusstseins und der menschlichen Rationalität stehende Faktoren und Prozesse berücksichtigt. Das Luckmann'sche Projekt wurde sehr viel weniger rezipiert als die zeitgenössischen Parallelaktionen[1] von Knorr-Cetina (1981), Latour und Woolgar (1979) und Lynch (1985), die als »Laboratory Studies« bekannt wurden,

1 Wir verwenden diesen Begriff hier, um mehrere zeitgleiche, aber voneinander unabhängige wissenschaftssoziologische Versuche, die Lebensweltfundiertheit von Wissenschaft herauszuarbeiten, zu fassen. Luckmann selbst verwendet den Begriff der »Parallelaktion«, der ursprünglich von Robert Musil stammt, in den Projektunterlagen zur Beschreibung der simultanen Erforschung von Alltagsbeschreibungen sozialer Situationen und der Entwicklung einer Notationspartitur, die aus seiner Sicht beide aufeinander bezogen bleiben sollen (Luckmann/Gross 1977, S. 207). Dieser Gedanke ist Teil seines allgemeinen Interesses an der Analyse der Konstitution von Zeichensystemen in der menschlichen Lebenswelt. Hier sieht er ebenfalls die Notwendigkeit einer Parallelaktion (1980b, S. 100): Erstens müsse eine phänomenologische Aufdeckung des Fundierungsverhältnisses zwischen subjektiven Bewusstseinsleistungen und gesellschaftlichen Kommunikationssystemen geleistet werden. Zweitens müsse die empirische Bestimmtheit des Einzelbewusstseins durch ein geschichtlich konkretes gesellschaftliches System der Wirklichkeitskonstruktion und Realitätsvermittlung in den Blick genommen werden. Es handelt sich also um eine Parallelaktion der

und hatte auch andere theoretische Referenzen. Knorr-Cetina, Latour, Woolgar und Lynch stellten theoretische Bezüge zu Ludwik Fleck, Thomas Kuhn, Émile Durkheim, Marcel Mauss und dem späten Ludwig Wittgenstein her und verpflichteten sich methodisch der Ethnografie (bisweilen gepaart mit einem kulturanthropologischen Perspektivenrelativismus, einer Methodenskepsis Feyerabend'scher Art oder einer ethnomethodologischen »Radikalreflexivität«). Demgegenüber orientierte sich das Luckmann'sche Projekt an der Phänomenologie Edmund Husserls und Alfred Schützens und war methodisch nicht nur an neuen Potenzialen der Audio- und Videografie interessiert, wie sie in sozialpsychologischen, linguistischen und interaktionssoziologischen Forschungen Einzug hielten, sondern setzte diese Instrumente auch selbst in der empirischen Forschung ein.

Auch das Erkenntnisinteresse war leicht anders gelagert. Die Autor:innen der »Laboratory Studies« forschten in naturwissenschaftlichen Laboren mit dem Ziel, ethnografisch zu rekonstruieren, wie naturwissenschaftliche Erkenntnisse dort unter vielfältigen situativen und sozialen Einflüssen als Fakten entstehen. Luckmann war stärker an sozialwissenschaftlichen Forschungen interessiert und wollte deren »Daten« ganz im Sinne der neuen Wissenssoziologie (Berger/Luckmann 1967) als gesellschaftliche Konstruktionen untersuchen. Sein Interesse galt den sozialen Prozessen, in denen Daten aus vorwissenschaftlichen Erfahrungs- und Handlungsgegebenheiten im situativ eingebetteten kommunikativen Handeln entstehen und später transformiert und in sozialwissenschaftliche Beweisführungen überführt – z. B. theoretisch und begrifflich überformt oder durch Numerisierung und Metrisierung reduziert und quantifizierbar gemacht – werden. Die hierbei verwendeten Präparationstechniken wie etwa die Verschriftung lebensweltlicher Ereignisse sollten in ihrer Bedeutung für die Entstehung der modernen, empirisch und intersubjektiv kontrolliert arbeitenden Sozialwissenschaften expliziert werden.[2] Das exemplarische Interesse galt dabei der zeitgenössischen Interaktionsforschung in der Linguistik, Sozialpsychologie, Ethnologie und Soziologie, die mit Aufzeichnungsgeräten Situationen sozialer Interaktion aufzeichneten und analysierten.

Der erste Fragekomplex des seinerseits in intensiver interdisziplinärer Kooperation zwischen Soziologie, Linguistik, Psychologie und hermeneutischen Fachgebieten operierenden Projekts bezog sich somit auf den Zusammenhang – Unterschiede, aber auch Kontinuitäten – zwischen alltäglichen und fachwissenschaftlichen Erlebnissen und Beschreibungen von Interaktionssituationen

phänomenologischen und der historisch-empirischen Forschung, deren Ergebnisse stets grundsätzlich zur wechselseitigen Deckung gebracht werden müssen.

2 Garfinkel nennt die vergeblichen Versuche der Tilgung lebensweltlicher Spuren aus wissenschaftlichen Daten, Beschreibungen und Erkenntnissen »Entindexikalisierung« und widmet ihnen ein eigenes »Breaching Experiment«, Latour (1993, S. 19–21) nennt sie mit Bachelard (1988, S. 18) »Reinigungsarbeit«, die charakteristisch sei für die »Moderne«.

(Beschreibungsproblem). Der zweite Fragekomplex adressierte die Transformation von alltäglichen und fachwissenschaftlichen Beschreibungen in sozialwissenschaftliche Daten und die Möglichkeit der Rückbindung dieser Daten an die Erlebnissituation (Datenkonstitutionsproblem).

Das nicht unehrgeizige angestrebte Resultat des Projekts war die Entwicklung einer theoretisch weitgehend neutralen, im Abstraktionsgrad kontrollierbaren Beschreibungssprache (»Notation«) für die integrierte Dokumentation und Darstellung des Gesamtverhaltens in sozialen Situationen. Dieses Notationssystem sollte die unmittelbaren kommunikativen Erfahrungs- und Handlungsgegebenheiten erfassen und konservieren, aus denen in weiteren Stufen der Objektivierung und Formalisierung schließlich sozialwissenschaftliche Daten gebildet werden. Zudem sollte eine Typologie der in den Sozialwissenschaften verwandten Daten und Datenerzeugungsverfahren entwickelt werden. Eine erste, noch hemdsärmelige Typologie unterschied zwischen »A«-, »B«- und »C«-Daten: Antworten bei Befragungen und Interviews, Beobachtungen und ihre Protokolle und »C/Konstruktionen« – Transkriptionen von aufgezeichneten sozialen Situationen. Das Projekt »Daten über Daten« widmete sich den Datentypen »C«.

Die später erstellte Typologie sozialwissenschaftlicher Daten wurde hinsichtlich unterschiedlicher Kriterien von Daten entwickelt: In Hinblick auf die »Abbildqualität« der Daten (von der Ereignisfixierung in der Aufzeichnung über technisch-apparative Transformationen bis zu sinnrekonstruierenden Umschreibungen), ihre mediale Realisierung (Film- und Tonaufnahmen, Verschriftungen und anderer Kodierungen, analoge und digitale Darstellungen) und mit Bezug auf die Kooperationsform von Datenproduzent und Datenlieferant.

Luckmanns Projekt ist somit insofern für den Band von Interesse, als es Erhebungskontexte, Analysekontexte und Nutzungsszenarien der Arbeit mit »Daten« nicht anhand von Forschungen zu anderen Themengebieten reflektierte, sondern zum genuinen Gegenstand der empirischen wissenschaftssoziologischen Forschung machte: Erhebungskontexte und Analysekontexte ebenso wie Nutzungsszenarien sind im eigentlichen Sinne, so die Annahme von Luckmanns Projekt, jeweils Situationen der sozialen Konstruktion von Daten und wissenschaftlichen Erkenntnissen. Dem werden wir im Folgenden weiter nachgehen. Zuerst, indem wir etwas detaillierter die theoretischen Hintergründe beleuchten, an denen sich das Projekt orientierte und in deren Bereich es sich Fortschritte versprach. Daran anschließend werden wir forschungspraktische Erfahrungen, auf die sich das Projekt bezog, sowie Vorgehensweise, Probleme und Erkenntnisse in der Projektarbeit selbst skizzieren, die auch Aussagen über Grundprobleme und Grundfragen in Bezug auf die Erhebungskontexte, Analysekontexte und Nutzungsszenarien der Arbeit mit dem Konstrukt »Daten« erlauben.

2. Theoretische Hintergründe

Das Projekt »Daten über Daten« schöpft aus theoretischen Quellen, die Luckmann bereits in seiner formativen Zeit an der New Yorker New School kennengelernt und an denen er in der ersten Hälfte der 1970er Jahre erneut gründlich gearbeitet hatte. In dieser Zeit arbeitete Luckmann intensiv an der Edition des ersten Bandes der nachgelassenen »Strukturen der Lebenswelt« von Alfred Schütz (1975). Noch stärker wurden diese Hintergründe jedoch theoretisch und programmatisch verdichtet in Luckmanns 1973 erschienenen Aufsatz »Philosophy, Science, and Everyday Life«.[3] In ihm stellt Luckmann fest, dass die von Husserl 1936 diagnostizierte »Krise der Wissenschaften« weiterhin ungelöst ist. Husserl identifiziert diese Krise mit der Feststellung, dass die modernen Wissenschaften ihre eigenen Grundlagen nach wie vor nicht geklärt hätten. Wie Husserl ([1936] 1976, S. 142) sagt: »Die Welt als Lebenswelt hat schon vorwissenschaftlich die ›gleichen‹ Strukturen, als welche die objektiven Wissenschaften, in eins mit ihrer (durch die Tradition der Jahrhunderte zur Selbstverständlichkeit gewordenen) Substruktion einer ›an sich‹ seienden, in ›Wahrheiten an sich‹ bestimmten Welt, als apriorische Strukturen voraussetzen und systematisch in apriorischen Wissenschaften entfalten, in Wissenschaften vom Logos, von den universalen methodischen Normen, an welche jede Erkenntnis der ›an sich objektiv‹ seienden Welt sich binden muss«.

Für Luckmann geht es vor diesem Hintergrund vor allem um den nach wie vor fehlenden »archimedischen Punkt, von dem aus ein Verstehen des Universums *und* ein Verstehen dieses Verstehens in *einem* Zuge geleistet werden könnten« (Luckmann 1980, S. 11–12; Herv. i. O.). In seinem Text rekonstruiert Luckmann die Genese und Zuspitzung der Krise und präsentiert einen Lösungsvorschlag in vier Gedankenschritten.[4]

(1) In der frühen Neuzeit entsteht mit der galileischen und newtonschen Wissenschaft eine »neue Kosmologie«, in der zwischen primären und sekundären Eigenschaften von Dingen unterschieden wird (Galilei) und Kausalbeziehungen zwischen einzelnen – insbesondere den primären – Eigenschaften eruiert werden (Newton). Primär sind die geometrischen und kinematischen Eigenschaften, die wahrgenommene Gegenstände messbar besitzen. Als sekundär gelten demgegenüber diejenigen Eigenschaften, die wir – die wahrnehmenden Menschen – den Gegenständen auf der Grundlage unserer sinnlichen Wahrnehmungen und Vorstellungen zuschreiben, die aber weder geometrisch noch kinematisch sind und sich deshalb einer (reduktiven) mechanistischen Beschreibung entziehen.

3 Die deutsche Komplettübersetzung dieses Textes erschien 1980. Auf sie werden wir uns im Folgenden beziehen.

4 Vgl. hierzu auch Meyer (2021).

Die primären Qualitäten lassen sich in einer mathematischen Weltbeschreibung (*mathesis universalis*) exakt ausdrücken, wobei *Zahlen* und *Zahlenverhältnisse* die Vergleichbarkeit zwischen den beschriebenen Gegenständen garantieren. In der neuen Kosmologie besteht also die eigentliche, objektive Wirklichkeit allein aus den primären Qualitäten, denn nur was mathematisch ausgedrückt werden kann, gilt nunmehr als wahrhaft, objektiv und absolut wirklich. Die sekundären Eigenschaften hingegen und alle damit zusammenhängenden Fragestellungen – wozu alle soziologischen Probleme zählen – fallen aus dem Bereich dieser sicheren Erkenntnis heraus.

(2) Bei der Akzeptanz des galileischen Modells entsteht daher die Herausforderung, in diesem neuen Schema auch den Menschen unterzubringen. Luckmann skizziert zwei unterschiedliche Versuche, aus dieser Situation zweier Welten – einer Welt der primären, mathematisierbarer Qualitäten und einer Welt subjektiver Sinneswahrnehmungen – in theoretisch und methodisch konsistenter Form der Sozialwissenschaft einen Ort zuzuordnen. Die beiden Versuche bestehen darin, (a) entweder den Menschen und sein subjektives Innenleben aus der neuen Kosmologie begründet auszuschließen oder (b) das Menschenbild zu rekonfigurieren, sodass der Mensch und seine sinnlichen Vorstellungen – mithin die sekundären Qualitäten – ebenfalls auf primäre Eigenschaften reduziert und so mathematisierbar gemacht werden können. Die erste Lösung (a) ist die von Descartes vorgeschlagene Unterscheidung von *res cogitans* und *res extensa*. Ihr zufolge ist der Mensch bzw. seine Seele nicht zurückführbar auf primäre Qualitäten und daher nicht mathematisierbar. Aus dieser Perspektive kann es nachgerade keine Wissenschaft (im strengen, galileischen Sinne) vom Menschen geben. Die zweite Lösung (b) geht demgegenüber davon aus, dass der Mensch Teil der Natur und damit reduzierbar auf primäre Qualitäten ist. Luckmann (1980, S. 25) zufolge beharren sie auf der Annahme, »das Universum sei zwar trügerisch, aber nichtsdestoweniger vollständig erkennbar; die dem vorwissenschaftlichen Menschen – einer minderwertigen, ›subjektiven‹ Spezies, die nur allzu schnell durch sekundäre Qualitäten den klaren Blick verliert – gegebenen Erscheinungen tragen in sich verborgen eine Struktur ›objektiver‹ primärer Qualitäten«. In dieser Tradition verortet Luckmann (1980, S. 27) alle Ansätze, die (vergeblich) versuchten, »die ›primären‹ Qualitäten des Menschen als eines sozialen, politischen, historischen Wesens dadurch zu bestimmen, dass man ihn in ein Trieb- oder Instinktbündel verwandelte, in einen *homo oeconomicus, sociologicus*, in einen Spielstrategen, in ein Persönlichkeits-Subsystem eines ›Handlungssystems‹ usw.« Zudem stellt sich hier die epistemologische Frage nach der Verknüpfung von Erlebnisqualitäten (Qualia) und Messdaten (epistemological gap).

(3) Das Problem der Sozialwissenschaften liegt selbst in der Nachkriegszeit noch darin, dass sie diese Dualität akzeptiert hat und sich keinen dritten

Weg denken konnte. So entstand im besten Fall eine friedliche Koexistenz, im schlimmsten Fall ein Wettkampf oder gar Krieg zweier Paradigmen, die wahlweise als Wissenschaften und Künste (Aristoteles), Naturwissenschaft und Geisteswissenschaft (Dilthey), nomothetische und idiographische Wissenschaften (Windelband) oder *science* und *humanities* (Snow) bezeichnet wurden. Die statistisch arbeitende Methodologie der Sozialwissenschaften, die ein *messendes* Vorgehen für sich in Anspruch nimmt, ist dabei, so Luckmann (1980, S. 24), eindeutig »ein Spross galileischer Naturwissenschaft«. Wenngleich sie die Messbarkeit sozialer Gegenstände unterstellt, seien es jedoch keine primären Eigenschaften, sondern »Assoziationen ›sekundärer‹ Qualitäten […], die eigentlich ›gemessen‹ werden – so, als repräsentierten sie unmittelbar die ›primären Qualitäten‹« (Luckmann 1980, S. 47). Die vermeintlichen Messungen seien grundlegend von sprachlichen, darunter unausweichlich alltagssprachlichen Prinzipien geprägt, die bei der Zurechnung von Ziffern oder Zahlen zu sprachlich konstituierten Objekteigenschaften und Ereignissen zur Anwendung kommen. Diese Kritik wurde ausführlich von Cicourel (1964) formuliert, auf den sich Luckmann (1980, S. 47) bezieht. Aus dieser Tatsache resultiere die spezifische Krise der Sozialwissenschaften der 1960er und 1970er Jahre, als quantitative Methoden die universitäre Lehre und disziplinäre Diskussion dominierten und die Etablierung qualitativer Methoden – die freilich ebenso wenig in der Lage sind, das Grundproblem, das Luckmann anspricht, aufzulösen – erst in den Anfängen war.

(4) Luckmann schlägt zur Lösung dieser Krise vor, die Herstellung von Vergleichbarkeit zwischen den Gegenständen nicht mittels einer aus Zahlen und Zahlenrelationen bestehenden Formalsprache, sondern mithilfe der Protosoziologie vorzunehmen. Er revidiert also keineswegs, sondern erneuert vielmehr das kosmologische Ziel einer *mathesis universalis* auch für das menschliche soziale Verhalten auf allen Ebenen. Eine solche der sozialen Welt angemessene *mathesis universalis* müsse allerdings sowohl von der galileischen Kosmologie als auch der leibnizschen Formallogik *unabhängig* sein und statt von der formalsprachlichen Reduktion und Numerisierung ihrer Gegenstände von der Prämisse der *epistemologischen Reflexivität* der Wissenschaften vom menschlichen Verhalten ausgehen (Luckmann 1980, S. 30–33).

Die Sozialwissenschaften, die sich der Beschreibung und Erklärung der sozialen Konstruktionen sozialer Wirklichkeit widmen, müssen laut Luckmann ein Programm der Formalisierung entwickeln, das den konstitutiven Strukturen des Alltagslebens – die eben selbst den raffiniertesten statistischen Methoden unausweichlich zugrunde liegen – angemessen ist. Das Programm einer *mathesis universalis* der sozialen Wirklichkeit kann daher nur in einer Phänomenologie der elementaren, invarianten und universalen Grundstrukturen des menschlichen Alltagslebens bestehen. Nur so ergebe sich auch eine befriedigende Lösung für das Problem der Vergleichbarkeit historischer Daten.

Um dies zu leisten, muss die Suche nach Formalisierungsmöglichkeiten vorangetrieben werden. Diese müssen dem Kriterium der »Adäquanz« gerecht werden, das laut Schütz (1972), auf den sich Luckmann hier bezieht, besagt, dass die Modelle der Soziologie, die den subjektiv gemeinten Sinn der beteiligten Akteure interpretieren, mit den von diesen verwendeten Deutungen kongruent sein müssen. Denn die sozialen Akteure der Alltagswelt verwenden bereits selbst Interpretationen, die Schütz »Konstruktionen erster Ordnung« nennt. Die wissenschaftlichen Interpretationen – etwa Modelle rationalen Handelns – schließen unweigerlich an diese Konstruktionen der Akteure an. Es handelt sich also um »Konstruktionen zweiter Ordnung«. Adäquanz bedeutet daher: »Jeder Ausdruck, der in einem wissenschaftlichen System benutzt wird, das sich auf menschliches Handeln bezieht, muss so konstruiert werden, dass eine innerhalb der Lebenswelt durch einen individuell Handelnden geleistete Handlung, und zwar so, wie es die typische Konstruktion anzeigt, sowohl für den Handelnden selbst als auch für seine Mitmenschen vernünftig und verstehbar ist« (Schütz 1972, S. 47). Die Erfüllung dieses Kriteriums verbürgt die Konsistenz der Konstruktionen der Sozialwissenschaftler:innen mit den Konstruktionen, die von der sozialen Wirklichkeit im Alltagsdenken gebildet werden. Auch alle sozialwissenschaftlichen Kausalitätsunterstellungen basieren letztlich – ob explizit oder implizit – auf der Adäquanz der Sinnkonstruktionen.

Beim Wechsel von Konstruktionen erster zu Konstruktionen zweiter Ordnung erfolgt nicht nur eine Abstraktion, sondern es erfolgt auch ein Wechsel der »Einstellung« der sich damit befassenden Person. Während die handelnde Person in der Alltagswelt Konstruktionen erster Ordnung nutzt, um pragmatische Ziele in der Welt zu verfolgen, ist die wissenschaftlich arbeitende Person bei der Produktion von Konstruktionen zweiter Ordnung von pragmatischen Zwängen entlastet. Die Alltagsperson nimmt die Welt, wie sie ihr begegnet, als gegeben hin, während die wissenschaftlich arbeitende Person auch Zweifel kultiviert. Dieser Einstellungswechsel kommt auch bei der Konstitution von Daten zum Tragen und wurde entsprechend im Projekt in den empirischen Fragenkatalog aufgenommen und gefragt: Welche Veränderung entsteht durch die Betrachtung einer sozialen Szene mit wissenschaftlicher Einstellung gegenüber einer Alltagseinstellung? Was wird als (lebensweltliche) Basis beibehalten, was verworfen, ersetzt, ergänzt?

Aus den hier skizzierten Hintergründen heraus müssen aus Luckmanns Sicht sozialwissenschaftliche Formalisierungen – hierzu zählen auch Prozesse der Datafizierung – stets an die bereits vorhandene protosozialwissenschaftliche Sprache des Alltags anschließen. Eine strenge phänomenologische Beschreibung der Strukturen der Lebenswelt stellt also eine Formalisierung von Aussagen über menschliches Verhalten dar, wie sie bereits umgangssprachlich artikuliert werden. Diese Verbindung gilt es zu bewahren.

Luckmann (1980, S. 51) identifiziert zwei Beschreibungsebenen für die in dieser Form positionierte »Messung« von Aspekten menschlicher Handlungen oder deren Objektivationen:

a) Die Ebene der invarianten Strukturen, die den Typisierungen der sozialen Wirklichkeit zugrunde liegen und damit die soziale Wirklichkeit mitkonstituieren;
b) die invarianten Strukturen, die der Versprachlichung geschichtlich konkreter Typisierungen menschlichen Handelns in der menschlichen Erfahrung zugrunde liegen.[5]

Als Beispiele für die allgemeinen Strukturen subjektiver Orientierung und subjektiven Handelns in der Welt des Alltagslebens führt Luckmann (1980, S. 54) den erlebten Raum, die gelebte Zeit, die Elementarstruktur von *face*-to-face-Situationen, die Ebenen der Anonymität, den biographisch-historischen Charakter jeglicher Erfahrung und die gelebte Intersubjektivität der Kommunikation im Alltagsleben an.

Ziel der Luckmann'schen Universalienforschung ist mithin die Schaffung einer allgemeinen Matrix für die innerhalb der Einzelwissenschaften geleisteten empirischen Analysen, die sich mit der Erfassung und Erklärung konkreter historisch spezifischer Geschehnisse im Alltagsleben beschäftigen. Das Ziel dieses Zugangs ist es, eine »generelle, auf die Ebene menschlichen Handelns zugeschnittene Matrix für Aussagen über menschliches Verhalten zu erstellen, die in historisch jeweils konkreten Umgangssprachen formuliert sind« (Luckmann 1980, S. 50). Gelingt dies, dann können die Universalstrukturen des Alltagslebens als Metasprache und *tertium comparationis* für die je historischen Sprachen dienen, in denen sozialwissenschaftliche Daten über menschliches Handeln vorinterpretiert vorliegen.

Diese Ambition lag auch dem Projekt »Daten über Daten« zugrunde. Mit ihm wurde eine überdisziplinäre phänomenologische Analyse der Konstruktion und Verwendung von »Daten« angestrebt, die die zwei naiven Realismen der galileischen Wissenschaften[6] vermeidet und eine Matrix erzeugt, die in den Einzelwissenschaften genutzt werden kann, um die Analyse von Daten auf eine stabile Grundlage zu stellen.

Exemplarisch sollte dies im Projekt zunächst am Beispiel der Überführung aufgezeichneter sozialer Geschehnisse in Texte erfolgen, d. h. Protokolle bzw.

5 In diesem Zusammenhang stellt Luckmann (1980, S. 52) auch einen Vergleich mit der seinerzeit erst in den Anfängen befindlichen Erforschung und Identifikation von Sprachuniversalien her.

6 Das heißt die Unterstellung zweier klarer Korrespondenzen zwischen (a) Welt und Geist (Wahrnehmung als Abbildung der Welt) und (b) Geist und Sprache (sprachlicher Ausdruck als Abbildung mentaler Zustände).

Transkripte. Ziel war die Entwicklung einer Beschreibungssprache zur Transkription sozialer Interaktionssituationen, die den referierten, sehr hohen Ansprüchen genügen kann. Die Projektmitarbeiter:innen stellten dabei fest, dass wissenschaftliche Beschreibungen und Vercodungen von Interaktionssituationen mittels theoretisch begründeter, konventionalisierter Zeichensysteme und Transformationsprozeduren bislang auf zwei vermeintlich natürlich gelernten Notations- und Transformationsweisen sozialer Wirklichkeit beruhen: auf der umgangssprachlichen Beschreibung und der standardsprachlichen Verschriftung. Diese ließen sie daher als unproblematisch verwandte »Prototypen« der wissenschaftlichen Ereignisbeschreibung und -vercodung gelten.

Hinzu kommt, dass in umgangssprachlichen Beschreibungen die Sachverhalte und Ereignisse, die in ihnen jeweils beschrieben werden, ebenfalls gewissermaßen als »Daten« behandelt werden – zwar nicht als »wissenschaftliche« Daten, aber als »Proto«- oder »Ethno-Daten«. Alltagsakteure betreiben entsprechend in gewissem Sinne selbst Datenerhebung und -analyse. Dies umfasst sogar, so der Projektbericht, dass die Frage nach der Reliabilität, Validität und Objektivität, die für die Qualität wissenschaftlicher Beschreibungen von so zentraler Bedeutung ist, in modifizierter Form immer auch für die Produzent:innen und Rezipient:innen umgangssprachlicher Beschreibungen von Relevanz ist.

Jedoch klammern diese Formen der »Protodatafizierung« andererseits auch vieles aus: Im alltäglichen Erleben werden etwa Differenzen zwischen unterschiedlichen Sinneseindrücken (von Geschehenem, Gehörtem, Gesprochenem), unterschiedlichen zeitlichen Strukturen und Phasenverläufen von den Alltagsakteuren quasi-automatisch integriert. In der wissenschaftlichen Untersuchung unterschiedlicher kommunikativer Modi (sprachlich, nicht-sprachlich, parasprachlich) und entsprechend differenzierter Zeitlichkeiten, Bewusstseinsspannungen und Affektionen müssen diese für die Herstellung einer »Gesamtpartitur« jedoch künstlich mittels komplizierter Verfahren isoliert und anschließend wieder harmonisiert werden.

Ein wissenschaftlich nutzbares Notationssystem steht vor der Herausforderung, einerseits die einzelnen relevanten Merkmale, Modalitäten und Kanäle des Geschehens genau zu konservieren, aber andererseits bei dieser Aufspaltung des Geschehens nicht dessen Gesamtgestalt, Harmonie und Integration zu zerschlagen. Im Projekt wurde dies im Anschluss an Georg Simmel als Problem der »Sinnschwelle« gefasst (Bergmann 1985). Jedes soziale Geschehen kann zwar gedacht werden als aufgebaut aus immer kleiner werdenden Geschehensmomenten (zeitlich) und -komponenten (modal). Unterschreitet deren Dokumentation jedoch eine bestimmte Schwelle, dann kann die Rekonstruktion des sozialen Phänomens unmöglich werden, da die Reintegration der Einzelmomente und -komponenten zum Ganzen scheitert und »das Phänomen verloren geht«. Bei Unterschreiten der »Schwelle der Zerkleinerung«, wie Simmel sie nannte, »verflüchtigt sich der Individualcharakter einer Gegebenheit, durch den sie zeitlich

fixierbar [wird] […], durch diejenige Auseinanderlegung und Spezialisierung ihrer Elemente, die doch als wachsende Exaktheit und Erkenntnis der Dinge ›wie sie wirklich gewesen sind‹, gilt. […] Mit der Kenntnis jeder Muskelzuckung jedes Soldaten würde uns jene einheitliche Lebendigkeit des ganzen Ereignisses, die Anfang und Ende seines zeitlichen Bildes verbindet, verlorengehen; das historische Element muß so groß bleiben, daß sein Inhalt Individualität behält. […] Die geschichtliche Erkenntnis bewegt sich also in einen dauernden Kompromiß zwischen der Aufstellung ausgedehnter Einheitsgebilde, deren Kontinuität zwar die Form des Geschehens nachbildet, aber nicht mit der Einzelheit realer Anschauungen zu erfüllen ist – und diesen letzteren, die im wissenschaftlichen Ideal nur je einen chronologischen Punkt bezeichnen und gerade dadurch dieses Ideal der Stetigkeit des realen Geschehens entrücken« (Simmel 1916, S. 28–30) (vgl. Corsten i. d. B.).

Im Projekt wurde dieses Problem ebenso wie alle anderen in diesem Abschnitt angesprochenen Fragen nicht theoretisch behandelt, sondern empirisiert. Es wurden zum Zweck der methodischen Dekomposition kommunikativer Vorgänge zunächst Film- und Tonbandaufzeichnungen von verschiedenen Gesprächssituationen angefertigt. Auf dieses Material wurde dann eine Reihe von Beschreibungs- und Analyseverfahren angewandt.

Das Projekt »Daten über Daten« wurde in einer Zeit umgesetzt, als sich Aufzeichnungsgeräte auch in der sozialwissenschaftlichen Forschung durchsetzten und adressierte proaktiv damit verbundene methodische Herausforderungen. Denn Aufzeichnungen ermöglichen eine Granularität der Rekonstruktion von sozialen Situationen, die für die Akteure selbst, die sich in der Flüchtigkeit der Situation befinden, nicht erreichbar ist. Ein weiterer Fragekomplex des Projekts bezog sich daher auf die Frage, welche kommunikativen Modalitäten überhaupt von Bedeutung für die Akteure sind, und welche von ihnen als irrelevant angesehen und behandelt werden. Während etwa die schriftliche Protokollierung sozialer Situationen Selektionen und Relevanzsetzungen der Beobachter:innen beinhalten, fallen solche Selektionen (sieht man von der Kameraeinstellung ab) bei Aufzeichnungsgeräten aus: die in den Fokus genommene Szene kann immer wieder und mit immer neuen Fragestellungen angeschaut werden. Daher stellt sich die theoretische Frage, ob die Kenntnisnahme bewusst wahrgenommener Aspekte der Situation bereits erschöpfend die Konstruktionen erster Ordnung seitens der Akteure begründen oder ob es auch flüchtige, »gesehene, aber unbemerkt« gebliebene Informationen gibt, die ebenfalls die Konstruktionen erster Ordnung informieren. Aus diesem Grund arbeitete das Projekt stets multi-methodisch, indem es selbst die quantitative und qualitative Auswertung der Aufzeichnungen mit Interviews mit den Akteur:innen und Analyst:innen der Aufzeichnungen verschränkte. Dies bezog sich sowohl auf phonetische Parameter, die auditiv-phänomenale Wahrnehmungen erfasste, als auch auf nonverbales Verhalten, bei dessen Analyse allein für den Bereich der Körperbewegungen 52

Dimensionen für jede:n Interaktionspartner:in entwickelt wurden, die erst daran anschließend mit mimisch-gestischem Verhalten in Beziehung gesetzt wurden, und die interaktive Organisation, welche die Mechanismen betraf, mittels derer das Gespräch von den am Gespräch Beteiligten als ein soziales Unternehmen hervorgebracht, abgewickelt und benutzt wurde. Insbesondere mithilfe von Interviews wurde das sinnhafte Erleben der Akteur:innen wie Beobachter:innen erforscht, etwa in Bezug auf reziprok unterstelltes Verständnis und Motive der Akteur:innen. Hier fanden die Projektmitarbeiter:innen heraus, dass die Künstlichkeit der Situation interessanterweise – und entgegen den Erwartungen – in hohem Maße geeignet war, die Leistungen der Akteur:innen, soziale Wirklichkeit zu gestalten, hervortreten zu lassen.

Schließlich bezog sich ein weiterer Forschungsbereich auf die umgangssprachlichen Beschreibungen und anderen Formen von Konstruktionen erster Ordnung, also die »Proto«- bzw. »Ethno«-Daten. Dabei wurden zwei strukturelle Ebenen untersucht: zum einen die für solche »rekonstruktiven Beschreibungen« typischen Muster (z. B. narrative Strukturen); zum anderen ihre durch die dialogische Produktion im Interview bedingten Merkmale. Aus diesem Projektbereich ging thematisch später das ebenfalls von der Thyssen-Stiftung finanzierte Forschungsprojekt über rekonstruktive Gattungen der Kommunikation von Luckmann und Jörg Bergmann hervor.

Ziel war es, all diese Modalitäten in einer umfassenden, alle Verhaltensdimensionen übergreifenden Notationspartitur zu integrieren, das unterschiedliche Granularitäten erlauben sollte, allerdings letztlich in Testversionen steckenblieb.

Während der Projektlaufzeit entstand im Jahr 1978 ein intensiver regelmäßiger Austausch mit Ulrich Oevermann, den Luckmann aus seiner Zeit an der Universität Frankfurt (1965–1970) gut kannte, und dessen Arbeitsgruppe. Beide trafen sich ein bis zweimal im Jahr an wechselnden Orten, um methodologische Diskussionen zu führen und in »Datensitzungen« gemeinsam Daten zu analysieren. Die objektive Hermeneutik war selbst in der Arbeit an Datenmaterial aus mit Tonband aufgezeichneten Interaktionen in Familien entstanden (Oevermann 2013, S. 89) und konnte daher reichhaltige Erfahrungen zu den methodologischen Fragen und Problemen des Projekts einbringen.

In zahlreichen Texten hat Oevermann seine Perspektive auf das »Datenkonstitutionsproblem« erläutert. Aus seiner Sicht besitzen soziale Gegenstände (Situationen, Erzeugnisse etc.) eine »Ausdrucksgestalt«, die im »Gesamt an Daten, in denen sich die erfahrbare Welt der Sozial-, Geistes- und Kulturwissenschaften präsentiert und streng methodisch – im Unterschied zu: praktisch – zugänglich wird, in denen also die sinnstrukturierte menschliche Praxis in allen ihren Ausprägungen erforschbar wird« (Oevermann 2002, S. 3). Ausdrucksgestalten tragen eine Sinn- und Bedeutungsstrukturiertheit in symbolischer oder gar sprachlicher und damit erst intersubjektiver und allgemein sozialer Weise in sich. Ihnen muss kein Sinn durch Interpretation zugeschrieben

werden, sondern es muss der in ihnen sich Ausdruck verschaffende Sinn rekonstruiert werden. Um analysierbar zu werden, müssen Ausdrucksgestalten als Protokolle (materiale Dokumente) sozialer Wirklichkeit und somit als »Daten« konserviert werden.

Dabei soll die beforschte Lebenspraxis möglichst unverfälscht konserviert werden. Oevermann unterscheidet zwischen technisch aufgezeichneten und beschreibenden bzw. gestalteten Protokollen, zwischen naturwüchsiger und inszeniert protokollierter Wirklichkeit und zwischen Fremdprotokollierung und Eigenprotokollierung (Oevermann 2002, S. 84–88). Die jeweils ersteren Varianten hält er als Datengrundlage für geeigneter, da die protokollierte Wirklichkeit hier noch nicht »in eine Wirklichkeit des Protokollierenden verwandelt« (Oevermann 2000, S. 84) wurde. Bei der Transformation sozialer Lebenspraxis in ein Protokoll vollziehen sich dennoch Transformationen: Das Protokoll repräsentiert zeitlich eine andere, von der Zeitlichkeit der enthaltenen Lebenspraxis losgelöste Realität. Es gilt daher, im Protokoll die Lebenspraxis in ihrer Sequenzialität zu bewahren. Nach Oevermann sind »alle Erscheinungsformen von humaner Praxis durch Sequenziertheit strukturiert bzw. konstituiert« (Oevermann 2000, S. 64), d. h. jede Handlung baut auf der vorherigen auf und bezieht sich auf diese. Nur wenn dies im Protokoll nachvollziehbar bleibt, kann die Lebenspraxis kontrolliert vom Handlungsdruck entlastet analysiert werden. Auch die Transkription erzeugt Veränderungen: Zwar ist die Schriftsprache ein quasi naturwüchsiges Notationssystem, aber Video-Aufzeichnungen sind kaum vollständig zu transkribieren. Wegen der Implizitheit vieler aufgezeichneter Vorgänge muss den Forscher:innen bei der Analyse eine Explikation der relevanten Präsuppositionen und Implikationen der Sinnzusammenhänge gelingen, die in den vorliegenden Ausdrucksgestalten und damit zusammenhängenden Erfahrungsgehalten enthalten sind. Auch Oevermann (2000, S. 95) sieht hier eine Gefahr der »Zerstörung« der Ausdrucksgestalt durch Transkription, die sowohl im Prozess der Datenerhebung als auch im Prozess der Datenauswertung erfolgen kann. Die Güte von Daten bemisst sich also auch in der objektiven Hermeneutik daran, wie gut sie die spezifischen Wirklichkeiten, die in Ausdrucksgestalten objektiviert sind, in Protokollen konservieren. Für die Güte von Protokollen ist methodologisch ausschlaggebend, dass die Protokollierung möglichst frei von Einflüssen der späteren Datenauswertung bleibt und die Protokolle daher immer wieder mit verschiedenen Fragen adressiert werden können. Je besser dies gelingt, desto größer ist die falsifikatorische Relevanz und die intersubjektive Gültigkeit der Daten. Für eine gültige Datenauswertung ist entscheidend, den pragmatischen Kontext zu berücksichtigen und einzubeziehen, in dem das auszuwertende Datum entstanden ist. Aus Oevermanns Sicht können diese drei Kategorien der Ausdrucksgestalt, des Texts und Protokolls dazu beitragen, eine »vereinheitlichende Methodologie der Datenerhebung und Datenauswertung« zu etablieren, mit der »die gegenstandsabhängigen Unterschiede zwischen den

Einzeldisziplinen in den Psycho-, Sozial-, Geistes- und Kulturwissenschaften« aufgehoben werden (2002, S. 5). Hier liegt er somit sehr nah an Luckmanns Ambition einer *mathesis universalis* für die Sozialwissenschaften.

3. Das Problem der Beschreibung sozialer Wirklichkeit – Die Entwicklung eines gesamtsozialwissenschaftlichen Notations- und Transkriptionsverfahrens

Das zentrale empirische Interesse des »Daten über Daten«-Projekts bestand in einem erweiterten Verständnis der beiden Probleme der »Konstitution von Daten« und der »sozialwissenschaftlichen Beschreibung«. Zur Entwicklung einer wissenschaftstheoretisch zunächst neutralen Sprache zur Beschreibung sozialer Wirklichkeit, die Schützens Postulat der Adäquanz gerecht wird, kooperierten im Projekt die Disziplinen Soziologie, Linguistik (insbesondere Phonetik) und Psychologie sowie hermeneutische Fachgebiete, um zu verstehen, wie seitens der Sozialwissenschaften soziale Wirklichkeit in wissenschaftliche Beschreibungen dieser Wirklichkeit transformiert wird und welche Rolle dabei alltägliche Formen des Beschreibens spielen. Die genaue Kenntnis dieses Transformationsprozesses sollte durch die Projektarbeit empirisch beobachtet und methodisch produktiv für die Entwicklung eines umfassenden und interdisziplinären Notations- und Transkriptionsverfahrens genutzt werden.

Die methodologischen Erkenntnisse über das Problem der Konstitution sozialwissenschaftlicher Daten sollten für die Entwicklung einer Beschreibungs- oder auch Notationssprache reflektiert werden, sodass auch eine historisch-vergleichende Forschung im Sinne der von Luckmann avisierten sozialwissenschaftlichen *mathesis universalis* möglich würde. Mit Blick auf die in den 1970er Jahren bereits bestehenden Verfahren der Notation und Transkription stellte bereits der Projektantrag fest, dass diese Verfahren die soziale Wirklichkeit entweder zu weit reduzieren oder zu stark detaillieren. Beide Varianten unterschreiten dadurch die oben diskutierte *Sinnschwelle*, was es den Forscher:innen unmöglich macht, die Handlungen und Relevanzen der Akteur:innen lebensweltlich nachzuvollziehen und sozialwissenschaftlich adäquat zu verstehen und zu erklären (u. a. Luckmann/Gross 1977; Luckmann 1979, 2003). Mit dem Projekt sollte demgegenüber ein Beschreibungsverfahren entwickelt werden, das Konstruktionen erster Ordnung in Konstruktionen zweiter Ordnung überführt, ohne die mitgeführten Konstitutionsbedingungen der Konstruktionen erster Ordnung zu entfernen.[7]

7 Ähnliche Überlegungen finden sich auch bei Garfinkel ([1958–63] 2019), der das Verhältnis beider Konstruktionsordnungen als *Lebenswelt Pair* beschreibt, ohne jedoch eine Beschreibungssprache, die dem Adäquanz-Postulat folgen kann, zu entwickeln oder auch nur entwickeln zu wollen.

Wie wollten nun Luckmann und seine Mitarbeiter:innen diese neue interdisziplinäre Beschreibungssprache entwickeln? Das Pionierprojekt widmete sich entlang der bereits genannten provisorischen Typologie sozialwissenschaftlicher Daten vor allem den »C«-Daten, die in Form von auditiven und audiovisuellen Aufzeichnungen soziale Wirklichkeit registriert und anschließend wissenschaftlich zugänglich gemacht, mit Sinn belegt und reproduzierbar gemacht werden müssen. Insofern reduzieren diese transformierenden Verfahren der Notation und Transkription nicht nur soziale Wirklichkeit durch ihre mediale Form der Verschriftlichung in Bezug auf bestimmte Erlebnismomente (z. B. multimodale, temporale und pragmatische Momente), sondern sie fügen ihr auch stets etwas Neues hinzu: eine Stabilisierung, Objektivierung, Entzeitlichung und Sinnaddition. Diese *Gewinne* und *Verluste* der sozialwissenschaftlichen Beschreibung sozialer Wirklichkeit sollten bei der Entwicklung eines umfassenden Notations- und Transkriptionsverfahrens als sozialwissenschaftliche Transformationen des empirischen Materials methodologisch reflektiert und methodisch produktiv einbezogen werden (Luckmann/Gross 1977). Dem Anspruch nach wurde das zu entwickelnde Verfahren dabei als »Gesamtpartitur« bezeichnet, die neben einer adäquaten Beschreibung der sozialen Wirklichkeit sozialwissenschaftliche Forschung in den Bereichen der Soziologie, Linguistik, Psychologie und hermeneutischen Fachgebiete ermöglichen sollte. Eine solche, überaus anspruchsvolle Gesamtpartitur wurde durch das Projekt jedoch nicht entwickelt, weshalb Luckmann (u. a. 2003) immer wieder von einem Scheitern des Projekts sprach.

Um die Entwicklung der Gesamtpartitur empirisch auf eine solide Grundlage zu stellen, wurde ein Setting gewählt, in dem von den Projektmitarbeiter:innen mit Versuchspersonen alltagsnahe Gespräche über unterschiedliche Themen geführt wurde. Insgesamt zeichnete das Projekt vier solcher Gespräche auf. Im Anschluss an die Gespräche wurden einige Versuchspersonen zudem interviewt, um deren retrospektiv lebensweltlichen Beschreibungen in Form von offenen Interviews wiederum per Audioaufzeichnungen zu registrieren und zu untersuchen. Diese umfassend registrierten Aufzeichnungen bildeten die Grundlage für die Erprobung verschiedener Beschreibungsverfahren durch die einzelnen Projektmitarbeiter:innen. Dabei experimentierten sie mit bereits bestehenden Verfahren der Notation und Transkription wie der Konversationsanalyse, der phonetischen Beschreibung, der Notation des Sprechausdrucks und der Notation nonverbalen Verhaltens (u. a. Luckmann 1979; Gross 1979; Jorns 1979; Winkler 1979, 1981; Dünker 1979). In diesem Prozess wurden nicht nur die einzelnen Verfahren reflektiert und am eigenen Material weiterentwickelt, sondern auch erste Überlegungen einer multimodalen und interdisziplinären Verknüpfung von Notations- und Transkriptionssystemen formuliert (Winkler 1979b). Eine Integration des

Transkriptionsverfahrens der Konversationsanalyse mit der phonetischen Transkription und nonverbalen Notationsverfahren wurde über die Interpretation von Sprechhandlungen versucht. Hierzu wurden die interpretativen Erkenntnisse des einen Verfahrens in Relation zu den Erkenntnissen der beiden anderen Verfahren gesetzt, sodass deren interpretative Übereinstimmungen und Unterschiede empirisch beobachtet werden konnten (Winkler 1979b, S. 463–464). Damit setzten die Arbeiten des Projekts zunächst auf der Ebene der Analyse und Interpretation an, um mögliche Integrationspotenziale auszuloten. Projektarbeiten, die sich auf die grundlegende Ebene der Beschreibung und einer Integration verschiedener Beschreibungssprachen beziehen, waren zwar vorgesehen, wurden aber durch die Projektmitarbeiter:innen nicht systematisch durchgeführt.

Im Folgenden möchten wir exemplarisch das experimentelle Design vorstellen, auf deren Grundlage die Datenkonstitutions- und Beschreibungsprobleme der Sozialwissenschaften empirisch beobachtbar gemacht, methodologisch reflektiert und methodisch gelöst werden sollten. Ähnlich zu Bateson's »Natural History of an Interview«, auf das das »Daten über Daten«-Projekt jedoch nicht rekurriert, wurde in einem Filmstudio in Konstanz eine »Ursituation« mit Videokameras und separaten Mikrophonen aufgezeichnet. Dazu wurden vier verschiedene *face-to-face*-Situationen (je ca. zehn Minuten) mit wechselnden Versuchspersonen produziert. Jeder Durchlauf wurde aus drei unterschiedlichen Perspektiven von den Videokameras und von zwei Richtmikrophonen aufgezeichnet (siehe Abb. 1).

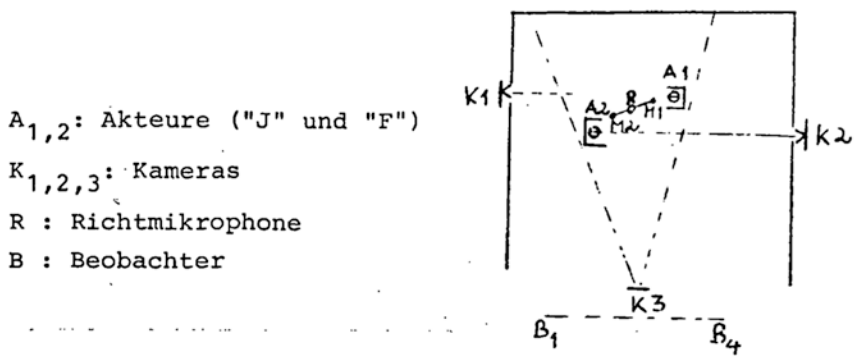


Abbildung 1: Unveröffentlichter Bericht für die Fritz Thyssen Stiftung 1978/79, S. 1 (DE-SAK-LUCK-F1.068-01)

Die Versuchspersonen saßen sich während des Durchlaufs etwas seitlich versetzt gegenüber, sodass die zwei Kameras jeweils eine sitzende Person und eine dritte Kamera die Gesamtsituation filmen konnte (siehe Abb. 2).

Abbildung 2: Halbtotale Kameraperspektive der vier »Ursituationen« (projektintern auch als take a, b, c und d bezeichnet)

Die Aufzeichnungen wurden von den einzelnen Projektmitarbeiter:innen mit unterschiedlichen Verfahren beschrieben, analysiert und interpretiert. Jörg Bergmann und Susanne Uhmann waren für die konversationsanalytische Beschreibung, Ulrich Jorns und Jeanette Bussi-Düncker für die nonverbale Beschreibung, Peter Winkler für die phonetische Beschreibung sowie Babis Karakalos und Jürgen Müller für die hermeneutische Untersuchung zuständig.

Die Arbeiten zur Entwicklung einer Gesamtpartitur widmeten sich entsprechend vier Dimensionen der sozialwissenschaftlichen Beschreibung menschlicher sozialer Interaktion: verbal-sprachlich-dialogischer, stimmlicher und bewegungs- und körpersprachlicher Dimensionen sowie dem sinnhaften Erleben der Akteur:innen. Die einzelnen Beschreibungen im Projekt wurden also nicht zum Zwecke der Beantwortung eigener thematischer Fragestellungen angefertigt, sondern waren darauf gerichtet, die verschiedenen Prozesse der Datenkonstitution und die beschreibungssprachliche Adäquanz von konservierten Daten offenzulegen. Dies bedeutete auch, dass die experimentelle Anlage des Projekts nicht direkt den Relevanzen sozialwissenschaftlicher Forschungsarbeiten folgt, die die Konstruktionen erster Ordnung untersuchen. Vielmehr macht sie die Konstruktionen zweiter Ordnung zum Gegenstand, die im eigenen Projekt und nicht etwa einem anderen Projekt produziert wurden. Diese selbstreferentielle experimentelle Anordnung des

Projekts führte zu verschiedenen Freiheiten, aber auch Problemen bei der Entwicklung einer Gesamtpartitur.

Dabei ist methodologisch bemerkenswert, dass die Entwicklung des Verfahrens nicht an spezifischen Forschungsgegenständen, sondern an einem allgemeinen, die interdisziplinären und multiparadigmatischen Sozialwissenschaften umfassenden Gegenstand ausgerichtet war. Das Projekt hatte neben seinen methodologischen und methodischen Interessen keine sozialwissenschaftliche Forschungsfrage zur Untersuchung eines spezifischen Phänomens. Die von Forschungsfragen befreite Entwicklung von Beschreibungsverfahren kann verschiedene Verfahren methodisch anwenden und methodologisch in Beziehung zueinander setzen, ohne eine am Gegenstand ausgerichtete Fragestellung berücksichtigen zu müssen. Diese Befreiung der untersuchten Konstruktionen zweiter Ordnung von spezifischen Relevanzen führt jedoch auch zu methodologischen Problemen, da die Entwicklung von Beschreibungsverfahren normalerweise an Forschungsfragen und -gegenständen ausgerichtet ist. Eine experimentelle Entwicklung kann zwar Verfahren kontrolliert testen, sie jedoch nur bedingt weiterentwickeln, ohne sich auf konkrete Phänomene und Fragestellungen zu beziehen (Bergmann 1991). Dieser methodisch relevante Phänomenbezug bei der Entwicklung von Methoden fehlte dem »Daten über Daten«-Projekt.

Die endgültige, in einem Produkt verkörperte Entwicklung eines allgemeinen sozialwissenschaftlichen und zugleich auch forschungsökonomischen Notations- und Transkriptionsverfahrens gelang dem Projekt nicht. Dennoch wurden wichtige Erkenntnisse erarbeitet, die die Entwicklung nachfolgender Verfahren angestoßen und beeinflusst haben. Auch wenn die Arbeiten an dem im Projektantrag vorgesehenen Sammelband über den »Alltag der Daten der Wissenschaftler« nicht abgeschlossen wurden, wurden im Rahmen der Projektarbeit relevante Erkenntnisse als Aufsätze veröffentlicht (Bergmann 1981, 1985; Gross 1981; Winkler 1981; Auer/Uhmann 1982; Luckmann 2003). Zudem haben einige Mitarbeiter:innen des Projekts in anschließenden Forschungskontexten an der Entwicklung nachfolgender Verfahren wie dem Gesprächsanalytischen Transkriptionssystem (GAT) mitgewirkt (Selting et al. 1998; Selting et al. 2009; Uhmann pers. Komm.). Allerdings realisiert GAT nicht den methodologischen Anspruch einer im Schütz'schen Sinne »adäquaten« Beschreibungssprache für die gesamten interdisziplinären Sozialwissenschaften, wie es Luckmann mit seiner *mathesis universalis* avisierte.

4. Fazit

Das Projekt »Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten« – landläufig »Daten über Daten« –, das Thomas Luckmann und Peter Gross unter Mitarbeit

zahlreicher weiterer Personen zwischen 1978 und 1982 an der Universität Konstanz durchgeführt haben, war für die damalige Zeit überaus innovativ und stand den Parallelaktionen der *Laboratory Studies* in nichts nach. Thomas Luckmann hat in vielen Interviews stets bedauert, dass es zu keinem wirklichen Abschluss gekommen sei, worunter er vermutlich die Publikation des zentralen Ergebnisbandes verstand, von dem uns (im Luckmann-Nachlass des Sozialwissenschaftlichen Archivs Konstanz) der Entwurf, die Gliederung sowie die Skizzen und sogar Endfassungen mehrerer Kapitel vorliegen. Woran dies lag, ist bis heute nicht vollkommen klar. Neben der Tatsache, dass es überaus ambitioniert war, war ein weiterer Grund sicherlich, dass Luckmanns Arbeitskraft stark von der Arbeit am zweiten Band der »Strukturen der Lebenswelt« (Schütz/Luckmann 1984) beansprucht war, dass Peter Gross bereits 1979 auf eine Professur nach Bamberg wechselte, und daher nur noch unter erschwerten Bedingungen an der Projektleitung partizipieren konnte, oder dass weitere Mitarbeiter:innen unter den prekären Bedingungen des Wissenschaftsbetriebs an ihren eigenen wissenschaftlichen Karrieren arbeiten mussten (Jörg Bergmann arbeitete z. B. an seiner Dissertation und Habilitation), nicht nachhaltig an die Universität Konstanz gebunden werden konnten oder ganz aus der Wissenschaft ausschieden.

Ob Luckmanns Diagnose überhaupt zutrifft, ist eine andere Frage. Denn das Projekt hinterlässt uns zahlreiche wertvolle Einsichten. Insbesondere erinnert es uns daran, Metaphern gegenüber skeptisch zu bleiben, die eine Fiktion der Objektförmigkeit von Daten suggerieren und sie als stabile Einheiten darstellen, die man – um nur einige der zentralen Metaphern zu erwähnen – »sammelt«, »erhebt«, »analysiert«, »interpretiert« oder »nutzt«. In solchen Metaphern zeigt sich die lebensweltliche Verankerung wissenschaftlichen Arbeitens, die Husserl (1936) und später eben erneut Luckmann (1973, 1980) diagnostiziert hat. Allerdings hat das Projekt dies nicht nur theoretisch hergeleitet und proklamiert, sondern vielmehr empirisch gezeigt, dass diese Aktivitäten stets daran beteiligt sind, Daten überhaupt erst als solche vermeintlichen Objekte zu konstituieren. Wie dies in langwierigen und aufwändigen transformatorischen, transkriptiven, analytischen und interpretativen Aktivitäten erfolgt, wurde im Projekt dokumentiert, auch wenn in dieser Frage keine synthetisierenden Erkenntnisse entstanden sind.

Auch wurden die Schwierigkeiten, Daten auf eine Weise zu konservieren, die neutral ist und die Daten zugleich aus vielen interdisziplinären Perspektiven rekonstruierbar hält, zwar demonstriert. Jedoch wurden zunächst auch hier keine endgültigen Ergebnisse produziert. Ganz sicher wurden im Projekt aber die Grundlagen gelegt für das Transkriptionssystem GAT, das es in seiner heutigen Form sicherlich nicht gegeben hätte, wenn Jörg Bergmann, Susanne Günthner und Susanne Uhmann nicht im »Daten über Daten«-Projekt und in anderen Kontexten als Mitarbeiter:innen für und mit Luckmann gearbeitet hätten.

Das Projekt »Daten über Daten« hat darüber hinaus ein Habitat und Umfeld geschaffen, das der Entwicklung und Verbreitung der qualitativen Methoden im gesamten deutschsprachigen Raum förderlich war. Der Aufstieg und die feste institutionelle Verankerung der qualitativen Methoden in den 80er und 90er Jahren ist ohne das Projekt nicht denkbar, auch weil es die wissenschafts- und erkenntnistheoretische Klärung und Schärfung vorangetrieben hat, die nicht nur in die neue Wissenssoziologie mit ihren verschiedenen Methodologisierungen (z. B. Gattungsanalyse, wissenssoziologische Hermeneutik) Eingang fanden, sondern auch in zwar verwandte, aber doch zu unterscheidende Methoden wie die objektive Hermeneutik, die deutschsprachige und skandinavische Konversationsanalyse und interaktionale Linguistik oder die dokumentarische Methode von Ralf Bohnsack. Der nachhaltige Wert des Projekts liegt also darin, eine Diskussion angestoßen und geführt zu haben, die produktive Wirkungen auf die wissenschafts- und erkenntnistheoretischen, methodologischen und methodenpraktischen Diskussionen im deutschsprachigen Raum und darüber hinaus hatte, insbesondere was die Vielfalt der sozialwissenschaftlichen Denk- und Arbeitsweisen anbetrifft.

Literatur

Auer, Josef C. P./Uhmann, Susanne (1980): Aspekte der konversationellen Organisation von Bewertungen. In: Deutsche Sprache 1, S. 1–32.

Bachelard, Gaston (1988): Der neue wissenschaftliche Geist. Frankfurt am Main: Suhrkamp.

Berger, Peter L./Luckmann, Thomas (1967): Die gesellschaftliche Konstruktion der Wirklichkeit. Frankfurt am Main: Fischer.

Bergmann, Jörg R. (1981): Frage und Frageparaphrase. Aspekte der redezuginternen und sequenziellen Organisation eines Äußerungsformats. In: Winkler, Peter (Hrsg.): Methoden der Analyse von Face-to-Face-Situationen. Stuttgart: Metzler, S. 128–142.

Bergmann, Jörg R. (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. Aufzeichnungen als Daten der interpretativen Soziologie. In: Bonß, Wolfgang/Hartmann, Heinz (Hrsg.): Entzauberte Wissenschaft. Zur Relativität und Geltung soziologischer Forschung. Göttingen: Schwarz, S. 299–320.

Bergmann, Jörg R. (1991): Deskriptive Praktiken als Gegenstand und Methode der Ethnomethodologie. In: Graumann, Carl F./Herzog, Max (Hrsg.): Sinn und Erfahrung. Phänomenologische Methoden in den Humanwissenschaften. Heidelberg: Asanger, S. 86–102.

Bergmann, Jörg (2016): Nachruf Thomas Luckmann (1927–2016). In: Zeitschrift für Soziologie 45, H. 4, S. 298–304.

Cicourel, Aaron V. (1964): Method and Measurement in Sociology. New York: Free Press.

Dünker, Jeannette (1979): Mimischer Affektausdruck und sprachliche Kodierung. Zürich: Universität Zürich.

Garfinkel, Harold (2019): Parsons' Primer. Stuttgart: Metzler.

Gross, Peter (1979): Gesprochenes Verschriften und Miteinanderreden beschreiben. Einige Unterschiede und Probleme. In: Zeitschrift für Semiotik 1, S. 153–159.

Gross, Peter (1981): Ist die Sozialwissenschaft eine Textwissenschaft? In: Winkler, Peter (Hrsg.): Methoden der Analyse von Face-to-Face-Situationen. Stuttgart: Metzler, S. 143–167.

Husserl, Edmund (1976): Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie (Hua VI).

Jorns, Ulrich (1979): Kodierung und Sinnzuschreibung bei der Notation nichtverbaler Phänomene dargestellt an Beispielen von Kopfhaltungen und Gesichtsbewegungen. In: Zeitschrift für Semiotik 1, S. 225–249.
Knorr Cetina, Karin (1981): The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science. Oxford: Pergamon Press.
Latour, Bruno (1993): We Have Never Been Modern. Cambridge, Mass.: Harvard University Press.
Latour, Bruno/Woolgar, Steve (1979): Laboratory Life. The Social Construction of Scientific Facts. London: Sage.
Luckmann, Thomas (1973): Philosophy, Science, and Everyday Life. In: Natanson, Maurice (Hrsg.): Phenomenology and the Social Sciences: Evanston: Northwestern University Press, S. 143–185.
Luckmann, Thomas (1979): Verhaltenspartituren: Notation und Transkription. In: Zeitschrift für Semiotik 1, S. 149–151.
Luckmann, Thomas (1980): Philosophie, Sozialwissenschaft und Alltagsleben. In: Luckmann, Thomas: Lebenswelt und Gesellschaft. Grundstrukturen und geschichtliche Wandlungen. Paderborn: Schöningh, S. 9–55.
Luckmann, Thomas (1980b): Aspekte einer Theorie der Sozialkommunikation. In: Luckmann, Thomas: Lebenswelt und Gesellschaft. Grundstrukturen und geschichtliche Wandlungen. Paderborn: Schöningh, S. 93–121.
Luckmann, Thomas (2003): Von der alltäglichen Erfahrung zum sozialwissenschaftlichen Datum. In: Srubar, Ilja/Vaitkus, Steven (Hrsg.): Phänomenologie und soziale Wirklichkeit. Entwicklungen und Arbeitsweisen. Opladen: Leske + Budrich, S. 13–26.
Luckmann, Thomas/Gross, Peter (1977): Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Konstitution sozialwissenschaftlicher Daten. In: Bielefeld, Hans-Ulrich/Hess-Lüttich, Ernest W. B./Lundt, André (Hrsg.): Soziolingusitik und Empirie. Beiträge zu Problemen der Corpusgewinnung und -auswertung. Wiesbaden: Athenaion, S. 198–207.
Lynch, Michael (1985): Art and Artifact in Laboratory Science. A Study of Shop Work and Shop Talk in Research Laboratory. London: Routledge & Kegan Paul.
Meier zu Verl, Christian (2018): Daten-Karrieren und epistemische Materialität. Eine wissenschaftssoziologische Studie zur methodologischen Praxis der Ethnografie. Stuttgart: Metzler.
Meyer, Christian (2021): Leib und Zeit. Zwei anthropologische Universalien in protosoziologischem Zugriff. In: Dreher, Jochen (Hrsg.): Mathesis universalis – Die aktuelle Relevanz der »Strukturen der Lebenswelt«. Wiesbaden: Springer VS, S. 113–142.
Oevermann, Ulrich (2000): Die Methode der Fallrekonstruktion in der Grundlagenforschung sowie der klinischen und pädagogischen Praxis. In: Kraimer, Klaus (Hrsg.): Die Fallrekonstruktion. Sinnverstehen in der sozialwissenschaftlichen Forschung. Frankfurt am Main: Suhrkamp, S. 58–156.
Oevermann, Ulrich (2002): Klinische Soziologie auf der Basis der Methodologie der objektiven Hermeneutik – Manifest der objektiv hermeneutischen Sozialforschung. Frankfurt am Main. https://nbn-resolving.org/urn:nbn:de:hebis:30-5409.
Oevermann, Ulrich (2013): Objektive Hermeneutik als Methodologie der Erfahrungswissenschaften von der sinnstrukturierten Welt. In: Langer, Phil C. et al. (Hrsg.): Reflexive Wissensproduktion, Frankfurter Beiträge zur Soziologie und Sozialpsychologie. Wiesbaden, S. 69–98.
Schütz, Alfred (1972): Das Problem der Rationalität in der sozialen Welt. In: Schütz, Alfred: Gesammelte Aufsätze II. Studien zur soziologischen Theorie. Den Haag: Martinus Nijhoff, S. 22–50.
Schütz, Alfred/Luckmann, Thomas (1975): Strukturen der Lebenswelt. Neuwied: Luchterhand.
Schütz, Alfred/Luckmann, Thomas (1984): Strukturen der Lebenswelt. Frankfurt am Main: Suhrkamp.
Selting, Margret/Auer, Peter/Barden, Brigit/Bergmann, Jörg/Couper-Kuhlen, Elizabeth/Günthner, Susanne/Meier, Christoph/Quasthoff, Uta/Schlobinski, Peter/Uhmann, Susanne (1998): Gesprächsanalytisches Transkriptionssystem (GAT). In: Linguistische Berichte 173, S. 91–122.
Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar/Bergmann, Jörg/Bergmann, Pia/Birkner, Karin/Couper-Kuhlen, Elizabeth/Deppermann, Arnulf/Gilles, Peter/Günthner, Susanne/Hartung, Martin/Kern, Friederike/Mertzlufft, Christine/Meyer, Christian/Morek, Miriam/Oberzaucher, Frank/Peters, Jörg/Quasthoff, Uta/Schütte, Wilfried/Stukenbrock, Anja/Uhmann, Susanne (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion 10, S. 353–402.

Simmel, Georg (1916): Das Problem der historischen Zeit. In: Liebert, Arthur (Hrsg.): Philosophische Vorträge. Veröffentlicht von der Kantgesellschaft (Nr. 12). Berlin: Reuther & Reichard.
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.
Winkler, Peter (1979): Notationen des Sprechausdrucks. In: Zeitschrift für Semiotik 1, S. 211–224.
Winkler, Peter (1981): Anwendungen phonetischer Methoden für die Analyse von Face-to-Face Situationen. In: Winkler, Peter (Hrsg.): Methoden der Analyse von Face-to-Face-Situationen. Stuttgart: Metzler, S. 9–46.

Die interpretative Analyse gefilmter Praxis

Verheißungen und Fallstricke

Michael Corsten

Zusammenfassung: Der Beitrag sondiert die Forschungslage zur Qualitativen Videoanalyse aus praxistheoretischer Perspektive. Dabei zeigt sich, dass Differenzen in der sozialontologischen Auffassung von Praxis Konsequenzen haben für die methodologische Reflexion sowie für die methodische Durchführung von Videoanalysen (Abschnitt 2). Zudem lassen sich Divergenzen konstatieren, die in Bezug auf den Datencharakter von audiovisuellen Aufzeichnungen vorliegen und die Frage nach deren adäquater Aufbereitung tangieren (Abschnitt 3). Der Beitrag geht weiter darauf ein, dass in neueren Studien der Qualitativen Videoanalyse nicht nur konventionell zwischen einer Makro-, Meso- und Mikroebene der Beobachtung unterschieden, sondern ein mittels audiovisueller Aufzeichnung beobachtbarer Nanobereich des Sozialen als Zielgröße der Analyse ausgezeichnet wird (Abschnitt 4). Die praxistheoretische Reflexion der Qualitativen Videoforschung plädiert abschließend für eine systematischere Beachtung von kritischen Entscheidungspunkten im Prozess der Forschung, die aufgrund der zuvor dargestellten sozialontologischen Differenzen und methodischen Varianten im Umgang mit der Datenaufbereitung entstehen, und deren transparentere Adressierung bei der Darstellung und dem Vergleich von Forschungsergebnissen (Abschnitt 5).

Schlüsselwörter: Vollzugswirklichkeit der Praxis, Ordnungsebenen der Praxis, Transkription oder Ethnographie, Mikro- oder Nanoanalyse, methodologische Konsequenzen, Entscheidungspunkte des Forschungsprozesses

1. Einleitung

Von der technischen Möglichkeit soziale Wirklichkeit speziell durch Videokameras aufzuzeichnen, sind Verheißungen ausgelöst worden, die auch in der Qualitativen Sozialforschung in den letzten rund 20 Jahren sehr stark angenommen wurden und zu einer Flut von methodischen Ansätzen, empirischen Untersuchungen und Publikationen geführt haben. Das Datenmaterial der audiovisuell aufgezeichneten sozialen Prozesse verspricht die Chance, soziale Wirklichkeit im unmittelbaren Vollzug zu registrieren, zu dokumentieren und für zukünftige Forschung zu konservieren. Die Wirklichkeit sozialer Praxis muss nicht mehr

umständlich aus zweiter Hand erfragt werden oder durch Feldforscher:innen beobachtet und mit entsprechenden Fehlerquellen protokolliert werden, sondern lässt sich anhand des Wiederabspielens des aufgezeichneten Materials wieder und wieder beobachten. Auch der Replikation von Aufbereitung und Analyse eines stets wieder verfügbaren Datenmaterials scheinen kaum Grenzen gesetzt.

2. Praxis als symbolisch und materiell produzierte Vollzugswirklichkeit

Gleichwohl gilt es in theoretischer, genauer sozialontologischer Hinsicht genauer zu klären, was anhand von Kameras und Mikrofonen aufgezeichnet werden soll, um nachzuvollziehen, was wie aufgezeichnet werden kann. Dazu gilt es aus soziologisch interpretativer Perspektive zu klären, was unter Praxis verstanden werden soll. Hierbei zeigt sich bald, dass in der neueren Soziologie unterschiedliche Auffassungen von Praxis existieren, die Auswirkungen auf die Beurteilungen der Chancen und Risiken (vor allem Fehlschlüsse) der Analyse audiovisuell aufgezeichneter Praxis besitzen. Dazu wird hier zunächst Verständnisweisen sozialer Praxis nachgegangen, die mehr oder weniger stark von deren symbolischen Strukturierung ausgehen und auf unterschiedliche Weise Einzelhandlungen oder Tätigkeiten fokussieren, den Zusammenhang zwischen symbolischer Geltung und materieller Konstitution (Körperlichkeit, Stofflichkeit, Dinge) verschieden interpretieren sowie die Zeitlichkeit und Prozessualität der Praxis auf differente Weise berücksichtigen. Die Unterschiede in diesen drei Hinsichten haben erhebliche Konsequenzen für die Frage, was in Videos gesehen (und gehört) werden kann, ob und wie das Gesehene versprachlicht werden soll, und was als Vollzug der Praxis registriert und wie sequenzialisiert werden kann. Hinzu kommt, dass den Praxisauffassungen unterschiedliche Verständnisse von (sozialer) Ordnung unterliegen, die wiederum Folgen dafür haben, wie der Umfang, die Reichweite und der Vernetzungsgrad sozialer Wirklichkeit eingeschätzt wird.

2.1 Verständnisweisen sozialer Praxis

Vollzugswirklichkeit: Wenn Menschen gemeinsam (in Wechselwirkung miteinander) Praxis ausüben, welche Art von Wirklichkeit wird dann vollzogen? Die Soziologie oder die Theorie des Handelns hat uns hier eine Reihe von Begriffen hinterlassen, wie Verhalten, Handeln, Tun oder Tätigkeit, sowie für diese Arten des Agierens verschieden komplexe Teileinheiten (Verhaltens- oder Handlungszüge, basic actions, Praktiken) und Gesamtheiten (wie Verhaltens- oder Handlungsmuster, Tätigkeiten, Praxis) postuliert.

Im Begriff »Vollzugswirklichkeit« (Bergmann 1985) wird der zeitliche Aspekt der Prozessualität der Praxis betont. Das Prozesshafte der sozialen Wirklichkeit wurde dabei zuerst von Garfinkel (1967) kritisch der strukturfunktionalen Theorieperspektive von Talcott Parsons entgegengestellt. Die soziale Realität gilt dabei als Reihe von Ereignispunkten, deren Fortsetzung als kontingent anzusehen ist. Diese als ethnomethodologisch bezeichnete Perspektive weist gerade in der zeitlichen Hinsicht starke Parallelen zur Konversationsanalyse (Sacks/Jefferson 1992) auf, insofern dort von einer Folge von Turns (Interaktions- bzw. Konversationszügen) ausgegangen wird. Tun bzw. soziale Praxis wäre in dieser für die Interpretative Soziologie maßgeblichen Forschungsrichtung eine Folge von Handlungs-, Interaktions- oder Praxiszügen.

Instantaner Realismus: Mit Artur Danto (1973) gesprochen handelt es sich bei der ethnomethodologisch-konversationsanalytischen Perspektive um die Variante eines instantanen Realismus'. Konstitutiv für soziale Wirklichkeit ist der einzelne Handlungszug (›turn‹), der ereignishafter Natur ist und mit dem Moment des Erscheinens wiederum auch vergeht. Praxis wird instantan – von Augenblick zu Augenblick – vollzogen. Wirklichkeit muss daher re-aktualisiert werden. Zeitlich konstante Strukturbildungen der sozialen Wirklichkeit sind somit gegen den aus zeitlichen Gründen von Moment zu Moment entstehenden Verfall oder Entzug von Wirklichkeit aufrechtzuerhalten. Die Praxis kann hier mit Techniken der Erinnerung bzw. mithilfe eines sozialen bzw. kollektiven Gedächtnisses Struktur aufbauen, wobei die Erinnerung bzw. der Gedächtnisabruf stets ebenso ereignishaft aus einer Gegenwart heraus erfolgt. Gesellschaftliche Praxis erfolgt so immer in »Gesellschaften der Gegenwart« (Nassehi 2011). Gerade daher spielt die Möglichkeit der Aufzeichnungen und der darauf beruhenden Konservanz von Daten über die Wirklichkeit in der Ethnomethodologie eine so große Rolle (siehe hier Abschnitt 3).

Tun und Tätigkeit: Eine davon abweichende Praxistheorie lässt sich der Psychologie Leontievs entnehmen, genauer seiner Unterscheidung zwischen Einzelakt und Tätigkeit, und wird auch in der Sozialisationstheorie von Bronfenbrenner (1981) berücksichtigt. Demnach seien Akte (Einzelhandlungen) »molekular«, während Tätigkeiten als »molar« (»moljarnaja edinica«; Leontiev 2012, S. 79) aufzufassen sind. Während Moleküle somit Kleinstverbindungen von Reinstoffen bezeichnen, wie z. B. Wasser als H_2O, so wird mit »Molarität« auf die Zusammensetzung eines Stoffgemischs (die Luft, das Blut) verwiesen, die aus mehreren Stoffen besteht. Übertragen auf menschliche Praxis werden *Akte* als Einzelhandlungen oder Handlungszüge bestimmt, wie ein Tanzschritt oder das Aufkleben einer Briefmarke. Insofern können dann zusammenhängende Handlungen wie Tanzen oder Briefe schreiben (inklusive des Versendens, Korrespondieren) als *Tätigkeiten* verstanden werden.

Methodische Konsequenz I: Wenn hier die Praxisauffassungen der Ethnomethodologie (Garfinkel, Sacks) und Sozialökologie (Bronfenbrenner, Leontiev)

gegenübergestellt werden, dann nicht primär, um einen grundlagentheoretischen Streit aufzumachen, sondern um nach den unterschiedlichen Konsequenzen dieser Perspektiven für die Analyse der in audiovisuellen Records aufgezeichneten (gefilmten) Praxis zu fragen. Aus der ethnomethodologischen Perspektive mag dann die Aufzeichnung eines Hebens des linken Beins erstmal nur das Heben eines Beins dokumentieren und noch keinen Tanzschritt darstellen, der sich eben erst unter Voraussetzung des praktischen Rahmens des Tanzens erschließen lassen würde. Das Heben des Beins wäre eine körperlich realisierte und materialisierte Handlung, die bemerkt oder nicht bemerkt, gesehen oder nicht gesehen, erinnert oder vergessen werden kann. Sie hätte als solche noch keine (soziale) Bedeutung als Geste, als erwartbarer Zug im Rahmen einer Praxis.

Symbolische Geltung und materielle Beschaffenheit: Hier gelangen wir zu einer weiteren Bifurkation. Soll Praxis als immer schon sozial und symbolisch verstanden werden, oder bedarf es – wie z. B. aus der Perspektive des New Materialism – eines veränderten Verständnisses des Zusammenhangs zwischen Praxis, Materialität und Bedeutung? Mit Latour (2008), Stengers (2008) oder Barad (2012) gesprochen hängt davon ab, wen wir zur Gesellschaft zählen oder nicht. Aus einer eher klassischen oder überkommenen Perspektive der Interpretativen Soziologie lautete die Frage von Latour, Stengers oder Barad dann anders: Wen oder was zählen wir auf welche Weise (inwiefern) zur Gesellschaft? Oder konkreter: Inwiefern gehören menschliche Akteure, Tiere, Dinge, Techniken/Werkzeuge usf. zur Gesellschaft? Oder wie ergibt der Zusammenhang von menschlichen Akteuren, Tieren, Dingen, Werkzeugen usf. Gesellschaft? Während Latours »Akteur-Netzwerk-Theorie«, Stengers »Spekulativer Konstruktivismus« oder Barads »Agentieller Realismus« zu einer Gleichstellung von Akteuren, Tieren, Dingen als Aktanten tendieren, unterscheiden traditionelle Theorien der symbolischen Praxis (Ethnomethodologie, Sprachpragmatik, oder auch Bourdieu 1990) einen Vorrang institutioneller Fakten, und deren Abhängigkeit von sozialen Akteuren bzw. Akteurskonstellationen (wie Klassen bei Bourdieu 1987).

Mit Searle (1995) kann daher die Frage gestellt werden, wie sich »institutionelle Fakten« erzeugen lassen, und wer mit dem »we make it the case that S has power« gemeint sein kann. Searles Argument wäre dabei, dass lediglich humane Akteure die Zurechnung leisten können, dass jemand oder etwas über Macht verfügt, insbesondere dann, wenn es sich um legitime Macht bzw. Autorität handeln soll.

Methodische Konsequenz II: Auch diese grundlagentheoretische Differenz ist nicht irrelevant für sehr konkrete Entscheidungen, die im Rahmen von Videoanalyse getroffen werden müssen. Wie etwa sind bauliche, territoriale Bedingungen zu berücksichtigen, wenn es um die interpretative Analyse von gefilmten Praktiken geht, wie etwa Paniken (Duisburger Love Parade) oder Gewalteskalationen von Fußballfans, wie wir im vierten Abschnitt am Beispiel der Studie von Reichertz/Keysers (2018) nochmals näher betrachten?

Bedeutungen, Bewegungen, Bilder: Vor dem Hintergrund der hier kurz und lose skizzierten grundlagentheoretischen Klärungen des Praxisbegriffes sind konkreter für die Analyse gefilmter Praxis Entscheidungen zu treffen, die das Verhältnis von Bedeutungen, Bildern (Ausschnitte, Turns) oder Bewegungen (Zusammenhänge von Bildern, Bewegtbilder) betreffen. Das besondere Problem der interpretativen Analyse gefilmter Praxis ist somit der »Bedeutsamkeitszusammenhang« (Mannheim 1980) von »Bewegtbildern« (Moritz/Corsten 2019). Denn in dieser Perspektive sind Bewegtbilder nicht bloße Aneinanderreihungen von Bildern bzw. Ausschnitten, sondern Bewegungen. Bewegungen beziehen sich dabei auf ein Tun von Akteuren in der gefilmten Zeit, wie auch auf die Veränderung der Kontextbedingungen des gefilmten Raums.

Zeitbewusstsein und Bedeutsamkeitszusammenhang: Es handelt sich dabei um ein bekanntes Problem aus der Phänomenologie der Lebenswelt (Husserl 1928; Schütz/Luckmann 1979; Bergson 1908). Wie wird der Flug eines Vogels, der Tanz eines Paars, das Hören einer Melodie oder Sprechen eines Satzes registriert bzw. beobachtet, wenn es sich dabei nicht nur um die bloße Abfolge von Flügelschlägen, Schritten, Tönen oder Lauten handelt? Mit Bergson gesprochen geht es um die »kinematographische Funktion«, die im Bedeutsamkeitszusammenhang zwischen den Flügelschlägen oder den Lauten besteht, und sie zum Flug oder zum Sprechakt machen.

Methodische Konsequenz III: Analysiert man Videos/Bewegtbilder somit von Bild zu Bild, dann geht der Zusammenhang der Bewegung, sprich der gefilmten Praktik und/oder Praxis verloren. Und auch, wenn man wie Reichertz (2018) die Analyse auf einzelne Moves umstellt, bleibt die Frage nach dem Zusammenhang der Moves (Schritte, Flügelschläge etc.) bestehen.

2.2 Ordnungen

Wir gelangen damit zur Frage nach dem Bedeutsamkeitszusammenhang und womöglich der symbolischen Ordnung, die einer Praxis als Rahmen (Goffman 1974) oder »Background« (Searle 1995) stets schon mitgegeben ist.

Order at all points: In einem vordergründigen Widerspruch zu der Annahme, dass die soziale Praxis als Vollzugswirklichkeit von Moment zu Moment zerfällt und wieder re-aktualisiert wird, steht die Annahme, dass es zu jedem Zeitpunkt soziale Ordnung gibt. Methodisch bedeutet dies, dass jede Momentaufnahme, also jeder Turn, jeder aufgezeichnete Wirklichkeitsausschnitt dahingehend ›gelesen‹ oder rekonstruiert werden kann, dass auch hier Ordnung vorliegt. Es käme somit zu einer Strukturreproduktion an jedem Ereignispunkt der Vollzugswirklichkeit. So wie etwa in der Systemtheorie angenommen wird, dass ein soziales System über ereignisförmige basale Operationen erschlossen werden kann (das Wirtschaftssystem beispielsweise über die Operation der Zahlung), so weist aus

der Perspektive der Ethnomethodologie jede symbolische Operation Bedeutungsstrukturen auf. Die können daran abgelesen werden, wie die ereignishafte Praxis vollzogen wird, indem die einzelnen Turns nur genau genug beobachtet und in ihrer spezifischen Kombination rekonstruiert werden. Dieser ethnomethodologischen (und konversationsanalytischen) Grundidee folgen in der neueren Interpretativen Soziologie beispielsweise auch die Objektive Hermeneutik und die Dokumentarische Methode.[1]

Order over all points: Gleichwohl ist die Idee der ›Order at all points‹ zeittheoretisch und damit auch sozialontologisch limitiert. Dies lässt sich anhand der Differenz einer ›Order at all points‹ und einer ›Order over all points‹ explizieren. Zu jedem Punkt meint nicht das gleiche (oder dasselbe) wie über alle Punkte. Hierin wird nochmals die zeittheoretische Differenz einer Perspektive, die eine Menge von Ereignissen als Abfolge deutet, und einer Auffassung, die eine Menge aufeinanderfolgender Ereignisse als zusammenhängendes Tun (bzw. Tätigkeit) erfasst, deutlich. Der Zusammenhang, der *über alle Ereignisse* (Turns, Praktiken, Handlungszüge) *hinweg* besteht, verweist auf etwas anderes als die *zu jedem Zeitpunkt* in den Ereignissen erkennbare Ordnung. Während letztere auf eine strukturhomologe Reproduktion von aufeinanderfolgenden Ereignissen hinauslaufen muss, kann eine Ordnung über alle Ereignisse auch heterogene, d. h. in ihrer operativen Strukturierung andersartige Handlungs-, Interaktions- oder Kommunikationszüge, Praktiken in Zusammenhang setzen. So wie eine Schulstunde nicht aus einer operativen Grundpraktik (z. B. ›unterrichten‹), die von Moment zu Moment strukturhomolog iteriert wird, bestehen muss, sondern verschiedenartige Elemente (wie Begrüßung, Einführung, Orientierung, Aufgabenstellungen, Lösungswege sondieren, Lösungen auszeichnen, Rekapitulieren, Abschließen und Verabschieden) dramaturgisch verbindet, und insofern Unterricht als Zusammenhang erzeugt, ergeben sich für eine videoanalytische Interpretation/Rekonstruktion einer Unterrichtsstunde eben genau auch diese beiden methodischen Grundhaltungen.

Order over all cases: Ein letzter Aspekt der Ordnung betrifft die Frage nach der Reichweite ihrer Geltung. Es mag zwar eine allgemeine Eigenschaft sozialer Praxis sein, dass sie an jedem Punkt oder über alle Punkte Ordnung aufweist, nur ergibt sich daraus noch nicht, ob es sich dabei immer um dieselbe Ordnung oder um verschiedene Ordnungen handelt, und welche Reichweite der Geltung diesen Ordnungen jeweils zukommt. Dieses Problem wird i. d. R. von der Quantitativen Sozialforschung adressiert, wenn es darum geht, ob die an einem bestimmten Fall gewonnenen Mustereigenschaften auch für andere Fälle gelten bzw. für welche anderen.

1 Wobei Garfinkel (1967) die Ethnomethodologie selbst als Weiterentwicklung von Mannheims ursprünglicher Idee der Dokumentarischen Methode ansah.

Fassen wir die Differenz der drei Perspektiven auf soziale Ordnung nochmals an einem typischen Beispiel der Analyse von Unterrichtsvideos (z. B. Hecht 2009) zusammen. Aus der Sicht einer interpretativen Analyse sozialer Praxis (hier des Unterrichts) kann bereits an der Untersuchung eines einzelnen Moments, z. B. eine Geste des Aufzeigens eines Schülers bzw. einer Schülerin, soziale Ordnung erkannt werden, z. B. ungeduldige Angst von der Lehrperson übersehen zu werden, oder ambitionierte Disziplin. Im ersten Fall würde sich die Ungeduld in unruhigen, hektisch kreisenden Bewegungen des Arms bemerkbar machen können, im zweiten etwa an einem stabil durchgestreckten Arm und einer eindeutigen Geste der Hand (z. B. ausgestreckter Zeigefinger). Die Anzeichen von Ordnung an einem bestimmten Punkt der Praxis verweisen in diesem Fall auch auf eine Ordnung über alle Punkte der Unterrichtspraxis. Denn sowohl das ungeduldige als auch das disziplinierte Aufzeigen antizipieren die Dramaturgie des Unterrichts, erfassen mindestens intuitiv die Rolle (die Funktion, den Beitrag) des Meldens im Rahmen des Unterrichtsgeschehens als gemeinsam vollzogener Tätigkeit, die weit über das Moment des Aufzeigens hinausweist. Gleichwohl gelten die Formen der Aufzeigegeste nicht notwendig für alle Fälle des Aufzeigens einer:eines bestimmten Schülerin bzw. Schülers und auch nicht für alle Formen des Unterrichts.

Aus den drei möglichen interpretativ-analytischen Perspektiven auf die potenzielle soziale Ordnung einer gefilmten Praxis ergibt sich die forschungspraktische und methodische Frage, ob in qualitativen Videountersuchungen allein eine der Perspektiven ausreicht, oder ob Forschungsergebnisse erst dann hinreichend begründet sind, sofern sie die drei genannten Perspektiven zueinander in Beziehung setzen können.

3. Aufzeichnung und Beschreibung von Prozessen in der Zeit

Bisher haben wir uns mit Problemen der Interpretation von gefilmter Praxis beschäftigt, die sich als methodische Konsequenzen unterschiedlicher sozialontologischer Auffassungen von sozialer Praxis und sozialer Ordnung ergeben können. Dabei haben wir unterstellt, dass uns Videodaten unmittelbaren Zugang zur aufgezeichneten Praxis bieten. Diese Voraussetzung erweist sich jedoch als unvorsichtig.

3.1 Taping the world (Original und Reproduktion)

Wir hatten oben bereits mit Jörg Bergmann auf die »Flüchtigkeit sozialer Wirklichkeit« hingewiesen. An diesem Charakter sozialer Wirklichkeit können auch Aufzeichnungen nichts ändern. Auch wenn wir durch eine audiovisuelle

Aufzeichnung wieder und wieder hören können, was zu einem bestimmten Zeitpunkt und an einem bestimmten Ort gesagt wurde, und sehen können, was Akteure damals und dort getan haben, so bleibt die gefilmte Praxis dennoch vergangen. Die Aufzeichnung bewahrt die Wirklichkeit nicht in dem Sinne auf, wie wir z. B. ein Gepäckstück aufbewahren. Während der Inhalt des Gepäcks i. d. R. nach der Aufbewahrung noch Bestand hat, gilt dies für die aufgezeichnete Vollzugswirklichkeit nicht. Damit hängt noch ein weiterer Punkt zusammen. Während man bei der Gepäckaufbewahrung später i. d. R. noch alle Sachen beisammen hat, zeichnen die in der Videoforschung eingesetzten Aufnahmegeräte eben nur Ton- und Filmspuren, d. h. auditive und visuelle wieder wahrnehmbare Spuren auf.

Aus diesen Gründen ist Harvey Sacks Wendung vom »Taping the World« zu stark, es wird nicht die ganze Welt aufgezeichnet und die Aufzeichnung vermag es nicht, den Gegenwartscharakter des Aufgezeichneten wieder herzustellen. Die Aufzeichnung bleibt eine nur ausschnitthafte, selektive Sammlung von Spuren. Wie in der Archäologie besteht bei der interpretativen Videoanalyse für die Forschenden nur die Möglichkeit, die gegebenen Fundstücke mit den fehlenden Stücken spekulativ (oder gedankenexperimentell) zusammenzusetzen.

Auch wenn somit durch die Aufzeichnung Momente der Welt registriert wurden, so eröffnet auch hier letztlich nur der Weg der Rekonstruktion die Chance, den fragmentarischen Charakter des Aufgezeichneten zu überwinden. Um welche Art von Datenmaterial es sich bei audiovisuellen Aufzeichnungen handelt, wird daher fraglich. Das betrifft vor allem das Problem, ob die audiovisuellen Aufzeichnungen direkt als Daten analysiert werden können, oder ob Schritte der Datenaufbereitung erforderlich sind und wieso.

3.2 Transkription und Partitur

Da sich die Qualitative Videoanalyse i. d. R. mit einzelnen Fällen befasst und diese auch sehr genau Sequenz für Sequenz analysiert, liegt es nahe, dass die audiovisuelle Aufzeichnung selbst als relevantes Datum aufgefasst wird. Allerdings ist sowohl für den Fluss von Bewegtbildern bzw. der Geschehnisse, die damit festgehalten wurden, als auch für die prozesshafte soziale Praxis selbst zu konstatieren, dass die aufgezeichneten Momentaufnahmen an den Videos beobachtenden Forscher:innen ebenso vorüberziehen wie die Eindrücke einer gelebten Praxis an den daran Beteiligten.

Die Datenaufbereitung sollte daher dem Nachvollzug des »kinematographischen« Zusammenhangs der Bilder und Worte dienen, auf die Bergsons Rede zielte. Im Fall des aufgezeichneten Sprechens kann sich die Transkription auf die einspurige Linearität der aufeinanderfolgenden Worte fokussieren, deren Linie gewissermaßen als Leitgerüst verwenden. Zwar ergibt sich bei feineren Transkriptionen das Problem der Überlappungen zwischen Sprechenden. Allerdings

kann sich die Transkription von Gesprächen auf die geltende Sprechregel verlassen, dass zur selben Zeit nur einer spricht (und der andere zuhört).

Audiovisuelle Aufzeichnungen ergeben nun das Problem, dass neben den sprachlich rekonstruierbaren Lauten weitere, durch unterschiedliche Sinnesorgane und Beobachterstandpunkte wahrnehmbare Praxisdimensionen existieren, die Bewegungen der Personen vor der Kamera, womöglich die Kamerabewegung selbst sowie filmtechnische Aspekte wie Einstellungsgröße, Einstellungswinkel und Kameraausschnitt. Audiovisuelle Aufzeichnungen verweisen auf die Multimodalität sozialer Praxis. Die Transkription nimmt die Form einer Partitur an, die »Feldpartitur« (Moritz 2009) sozialer Praxis (vgl. Meyer/Meier zu Verl; Niediek; Vollmer i. d. B.).

Ein wichtiger Bezugsrahmen bleibt dabei die chronometrische Zeit, deren minütlicher, sekündlicher oder millisekündlicher Ablauf den transkribierten Lauten, Worten, fotografierten Inhalten, Kameraeinstellungen usw. eine Achse verleiht, auf der sie nacheinander und übereinander abgetragen werden können. Die Informationen, die für einen Moment, also auf das gleichzeitig Ablaufende bezogen sind, nehmen erheblich zu. Allerdings stellt sich die Frage, ob sich die zu einem Moment als Muster oder Kombination beobachtbaren Informationen als Bewegung in der Zeit tatsächlich weiterverfolgen lassen oder ob nicht – wie in den Husserl'schen Beispielen vom Vogelflug oder von der Melodie – die Beobachtung des Fortgangs einer Praxis durch die Fokussierung auf eine Wahrnehmungs- bzw. Praxisdimension (z. B. der Flug des Vogels oder das Spiel des führenden Instruments) gesichert wird. Dann ergibt sich aber wiederum aufgrund der Multimodalität das Problem, welcher Spur gefolgt werden soll.

Bleiben wir im Bild der Partitur, so bedürfte es für eine Interpretation der gefilmten Praxis sowohl eines Orchesters (zur Besetzung aller Spuren/notierten Instrumente) als auch eines Dirigenten, der die verschiedenen Spuren/Instrumente zusammenführt. In der tatsächlichen Forschungspraxis treffen wir allerdings häufig auf eine Überforderung der Interpretationsgruppen, die sich in der Vielfalt der Details in synchroner und diachroner Perspektive verlieren. Das Geschehen als Ganzes – der Vogelflug, die Aufführung eines Musikstücks, eine Radsportetappe, eine Unterrichtsstunde usw. – geht verloren im Labyrinth der instantan dekomponierten Wirklichkeiten (vgl. Meyer/Meier zu Verl i. d. B.).

3.3 Ethnographie und Beschreibung

Es ließe sich ein anderer Weg einschlagen, um von den Rohdaten der audiovisuellen Aufzeichnungen zu einer Aufbereitung eines konservierten hör- und sichtbaren Geschehens zu gelangen, und zwar der ethnographische Weg der »dichten Beschreibung« der gefilmten Praxis. Ziel wäre dabei eine mehr oder weniger

angemessene Versprachlichung des Gesehenen und Gehörten in der Aufzeichnung. Dies wäre dann keine Transkription und auch keine Partitur, sondern ein Protokoll oder spezifischer: ein »Feldprotokoll«. Auch wenn es viel »Verdrießliches über [so, M. C.] ein beliebtes Instrument« wie das Feldprotokoll zu berichten gibt (vgl. Reichertz 1992), so böte immerhin die nach wie vor existierende audiovisuelle Aufzeichnung einen Referenzpunkt, an dem sich die Feldprotokolle abgleichen ließen. Man mag dies für aufwändig halten; aufwändiger als die Erstellung einer Partitur aus den mehrdimensionalen Aufzeichnungen ist es sicherlich nicht.

Was das ethnographische Protokoll gegenüber der Transkription als Partitur zum Vorzug hat, ist der Umstand, dass hier die Versprachlichung unmittelbarer als Interpretation erkennbar bleibt, während sie bei der Partitur hinter den sterilen Notations- bzw. Transkriptionssystemen zu verschwinden droht.

Solche ethnographischen Protokollierungen aus audiovisuell aufgezeichnetem Datenmaterial sollten ähnlich wie in Bohnsacks Vorschlägen zum Vorgehen einer Dokumentarischen Methode zwischen einer möglichst registrierend verfahrenden Beschreibung und einer reflektierend rekonstruierenden Interpretation der Beschreibungen unterscheiden, und diese aufeinander folgend durchführen. Ergeben sich im zweiten Schritt der interpretativen Rekonstruktion Mehrdeutigkeiten oder zeigen sich Lücken in der Beschreibung, so wäre dies Anlass auf das Rohmaterial der Aufzeichnung selbst wieder zurückzugehen.

Ziel wäre dabei den »Bedeutsamkeitszusammenhang« zu erschließen, der Karl Mannheim in seiner ursprünglichen Fassung der »Dokumentarischen Methode« als »Dokumentsinn« eines registrierten sozialen Geschehens vorschwebte.

4. Skopie(n): Brennweiten und Brennpunkte der interpretativen Analyse

In den beiden vorigen Abschnitten ist an einigen Stellen von der unterschiedlichen Reichweite der interpretativen Analyse gefilmter Praxis die Rede gewesen. Damit einher geht die Frage, ob die Qualitative Videoanalyse einen bestimmten perspektivischen Zugang zur sozialen Realität privilegiert. Einfacher gefragt: Favorisiert Qualitative Videoanalyse eine Makro- oder Mikrosoziologie, oder gar – wie es heute prominent heißt – eine »Nanosoziologie«?

Üblicherweise werden qualitative Verfahren der Hermeneutik, des Symbolischen Interaktionismus, der Ethnomethodologie und Soziolinguistik der Mikrosoziologie zugerechnet. Daher ist danach zu fragen, ob die Möglichkeiten der Videoanalyse noch genauer in die millisekündlichen Abläufe des Geschehens einzudringen, rechtfertigen, eine Nanosoziologie der sozialen Wirklichkeit anzustreben.

4.1 Von Makro bis Nano – und wie zurück?

Wie bereits weiter oben erwähnt, folgt die Qualitative Videoanalyse einer Haltung, die von Autoren wie Harold Garfinkel der sogenannten »Grand Theory« (damals repräsentiert durch Talcott Parsons) entgegengesetzt wurde. Große Theorien, die sich für die übergreifenden Ordnungsleistungen interessieren, die durch soziale Strukturen erzeugt würden, tendierten zu einer Überbetonung der Stabilität und Reproduktion von sozialer Ordnung. Demgegenüber untersuchen Ethnomethodologie oder Symbolischer Interaktionismus die fragilen Operationen und Prozesse, in denen die Akteure im Kleinen, im Mikrobereich des Sozialen, Ordnungen erzeugen, ohne von der Stabilität der Ordnung ausgehen zu können bzw. diese als empirisch gegeben vorauszusetzen.

Von Augenblick zu Augenblick ist soziale Praxis darauf angewiesen, Formen, Beziehungen oder Konstellationen des Sozialen zu aktualisieren und im Umlauf zu halten. Soziale Strukturen sind nicht aus sich selbst heraus von Bestand, sie werden in der Praxis der Akteure stets aufs Neue vollzogen und reorganisiert. Daraus ergeben sich zwei mögliche Fragen: (1) Wie groß bzw. klein, lang oder kurz ist der Augenblick, in dem die Ordnung des Sozialen stets wieder hergestellt wird? (2) Wie hängen die kleinen Momente der Praxis mit den großen, zeitlich übergreifenden sozialen Einheiten zusammen?

Bezogen auf die beiden Fragen erweisen sich die gängige Makro-Meso-Mikro-Differenzierung der Soziologie kaum als tauglich. Sicher können sie als heuristisch eingenommene Perspektiven dazu beitragen, unterschiedlich weitreichende Formen gesellschaftlicher Relationen zu entdecken. Aber jenseits eines konkreten Untersuchungsgegenstands erweisen sie sich eher als leere Abstraktionen.

Dennoch lässt sich in Bezug auf die erste Frage darüber nachdenken, wie klein der Ausschnitt einer aufgezeichneten sozialen Praxis sein kann, um noch sinnvoll analysiert werden zu können. Kann es so etwas wie eine Mindestdauer für ein zeitliches Intervall der intersubjektiv geteilten Aufmerksamkeitsspanne für soziale Praxis geben? Wie lang ist etwa der Moment, in dem aus einer Ansammlung von Personen ein fokussiertes Zusammentreffen wird?

Im Grunde genommen hängt die Antwort auf die erste Frage zu einem relevanten Teil von der Beantwortung der zweiten Frage ab. Denn es kommt dabei auf den Grad der Kopplung zwischen Einheiten an, die mit je unterschiedlicher Reichweite auf sich selbst bezogen sind und zudem miteinander relationiert sein können. Je nachdem beispielsweise, ob ein Unterricht als Frontalunterricht zwischen der Lehrperson und dem dispersen Publikum der Schüler:innen oder als Gruppenarbeitsphase organisiert ist, wechselt das Verhältnis der Relationen, die sich hier auf Makro-, Meso- oder Mikroebene abspielen und gegenseitig stützen können. Die abstrakte Differenzierung dient daher als Heuristik, die konzeptionell für den untersuchten Gegenstand sensibilisieren soll, und damit gleichwohl zur Entwicklung methodischer Kriterien anregt, die zu erfüllen sind, um das

Zusammenspiel sozialer Ordnungsbildung von Prozessen mit unterschiedlicher sozialer Reichweite rekonstruieren zu können. Dazu widmen wir uns im nächsten Abschnitt einem Beispiel aus der Forschungsliteratur, das auf der Basis der bisherigen Überlegungen reflektierend reinterpretiert wird.

Zuvor soll allerdings noch auf eine andere heuristische Differenzierung hingewiesen werden, mit der sich die unterschiedlichen Reichweiten von Praxiskreisläufen nachvollziehen lassen. Gemeint ist die aus der kommunikativen Gattungsanalyse stammende Unterscheidung zwischen Außenstruktur, Binnencodierung und situativ vermittelter Realisierung (Günthner/Knoblauch 1994) von Kommunikation, die hier im weiten Sinn als symbolische Praxis verstanden wird. Die an einzelnen ›Points‹ eines Praxiskreislauf sichtbaren ›Turns‹, Beiträge oder Praktiken stellen die Ereignisse der situativen Realisierung sozialer Praxis dar. Zur Ebene der Binnencodierung stehen sie in einem formativen (präfigurativen) Verhältnis, da Praxis einer symbolischen Codierung durch inkorporierte Wahrnehmungs- und Verhaltensschemata, Skripte und/oder Wertmuster unterliegt. Praxis ist insofern inhärent (von innen her) durch ein ihr verfügbares Wissens- und Bedeutungsrepertoire ko-konstituiert. Zur Außenstruktur weiterer Bereiche der sozialen Welt steht Praxis in einem funktionalen Verhältnis.

Übertragen auf das im nächsten Abschnitt diskutierte Beispiel der Untersuchung von Gewalteskalation im Kontext von Fußballevents lassen sich Binnencodierung, Außenstruktur und situative Realisierungsebene folgendermaßen veranschaulichen: Auf der situativen Realisierungsebene zeigen sich jeweils verschiedenartige Beiträge, die aufgrund der Binnencodierung der Praxis als dem Fußballevent typische Verhaltens- oder Handlungszüge zugerechnet werden können, z. B. Tragen von Emblemen, Trikots, Fahnen, Schals usw., die Personen als Fan klassifizierbar machen, das Ausführen typischer Gesten (z. B. Aufstehen und Heben der Arme zur Produktion einer ›Welle‹ durch die Fans), Skandieren von Parolen oder Singen von Vereinshymnen, usf.).

Dazu gehören zudem situativ nur vereinzelt (gelegentlich) realisierte Akte wie Beschimpfungen oder sarkastische Parolen gegenüber den Fans der gegnerischen Clubs oder auch der Polizei, sowie Gestik, Mimik oder Bewegungen, die sich als Anzeichen einer aggressiven Bereitschaft zur Konfrontation lesen lassen. Aber auch Ereignisse der kollektiven Ergriffenheit von Momenten des Spiels oder der eigenen Darbietungen (Efferveszenz).

Als funktional erweist sich die Fanpraxis für die Außenstruktur der feldspezifischen Öffentlichkeit des Unterhaltungssystems Sport, sie trägt spezifisch zur emotionalen Herstellung eines Massenevents bei und steigert dabei gegebenenfalls dessen Wert im ökonomischen Feld der Gesellschaft insgesamt. Feldspezifisch codierte und situativ realisierte Praxen stehen daher in polykontexturalen Funktionsbezügen zu Bereichen der gesellschaftlichen Außenstruktur. Forschungsmethodisch hat dies zur Konsequenz, dass eine Aufzeichnung situativ realisierter Praxis nicht ohne Rekurs auf geteilte Bedeutung zur Rekonstruktion

der Binnencodierung und auf Weltwissen zur Funktionsbestimmung für die Außenstruktur interpretiert werden kann.

4.2 Beispiel: Die Studie »Emotion. Eskalation. Gewalt.«

Zur Auswahl des Beispiels: Die Studie »Emotion. Eskalation. Gewalt.« (Reichertz/Keysers 2018) wurde deshalb gewählt, weil sie in mehreren Hinsichten als exemplarisch für die neuere qualitative Videoforschung in den Kultur- und Sozialwissenschaften angesehen werden kann. Zudem weist sie in ihrem Zugang einige Überschneidungen mit dem hier zu veranschaulichenden praxistheoretischen Ansatz der Videoanalyse auf, die ich in Grundzügen kurz skizzieren möchte. Erstens folgt die Studie konzeptionell der rahmenden Idee des »Kommunikativen Konstruktivismus« (Keller/Knoblauch/Reichertz 2013), die mit der hier vorgestellten Perspektive folgende Merkmale teilt. Sie geht von der Vorstellung der sozialen Realität als prozesshafte Vollzugswirklichkeit aus, die symbolisch praktisch bzw. kommunikativ konstituiert wird. Gleichwohl werden materielle bzw. physische Eigenschaften der sozialen Praxis berücksichtigt, wie z. B. die Aspekte der Leibgebundenheit und Körpervermitteltheit von Praktiken, Wahrnehmungen und Kommunikation sowie der Bedeutung von Territorialität als physische Begrenzung und regulative Dimension des Sozialverhaltens. Dies überträgt die gewählte Studie auch überzeugend auf den von ihr allgemein untersuchten Gegenstand, die Bedeutung von Emotionalität in Massenveranstaltungen und deren potenzielle Folgen für die Ausübung und Eskalation von Gewalt in Massen. Zweitens berücksichtigt speziell Reichertz (2018) in seiner theoretischen Reflexion der Erträge der Studie ausdrücklich Sachverhalte, die sich als »order at all points«, als »order over all points« und als »order over all cases« verstehen lassen. Hierbei kennzeichnet er als praxisspezifischen Ereignispunkt der Analyse audiovisueller Aufzeichnungen den »move«, der sich eben von »Bildausschnitt« bzw. »still« als Grundeinheit der Dateninterpretation unterscheide (Reichertz 2019, S. 109). Basal geht es insofern um die interpretative Rekonstruktion der »order at all moves«. Gleichwohl berücksichtigt Reichertz (2018, S. 218 ff.) mit der Dimension der »trajectory« auch den Sachverhalt des »Zum-Fußball-Gehens« als eine »order over all moves«. Bemerkenswert ist zudem die Hinzunahme von Verfahren der automatisierten Bewegtbildanalyse (des Computersehens) zur algorithmusgestützten Detektion von ordnungstypischen versus abweichenden Bewegungen der Masse in »regions of interest« (Reichertz 2018; Horn et al. 2018, S. 133; Horn et al. 2019), die sich als Suche nach »order over all cases« begreifen lässt. Und nicht zuletzt orientiert sich Reichertz (2018, S. 243) sowie anschließende Einzelstudien (Moritz 2019, S. 336) an der heuristischen Differenzierung zwischen Makro-, Meso- und Mikroperspektiven nebst der Einführung einer Nanoperspektive der Analyse videographierter Praxis.

Systematische Methodik der Videoanalyse: Verfahrenslogisch postuliert Reichertz (2019) die von ihm eingeführte »Kunstlehre der wissenssoziologisch-hermeneutischen Videointerpretation«, die er als eigenständig und systematisch ausgearbeitet ansieht, und mit dem Transkriptionssystem HANOS (HAndlungsorientiertes NOtationsSystem) ausgestattet ist. Die auf Handlungen fokussierte Methodik weist daher wiederum Parallelen zu einer praxistheoretischen Videoanalyse auf und betont dabei eine für sie zentrale Unterscheidung von Handlungsebenen, nämlich der »Handlung der Kamera« und der »Handlung vor der Kamera«. Dabei hat laut Reichertz (2019) die »Handlung der Kamera«, also der praktische Umgang mit dem Aufzeichnungsgerät, sogar methodisch-analytische Priorität.

Die Auszeichnung der Nanoperspektive: Sowohl bei Reichertz (2018) als auch bei Moritz (2019) wird die Nanoperspektive besonders hervorgehoben. Reichertz folgt einer konventionellen soziologischen Heuristik, wenn er (a) der Makroperspektive übergreifende sozialstrukturelle Ordnungs-, Differenzierungs- und Funktionslogik zurechnet, (b) auf der Mesoebene soziale Phänomene wie Organisationen, Institutionen oder Brauchtum verortet und (c) der Mikroebene Sachverhalte der alltäglichen Verhaltensabstimmung zuweist (Reichertz 2018, S. 288). Davon wird die Nanoperspektive definitorisch abgehoben: »Mit Nanobereich ist dagegen die Ausdrucksebene sozialer Interaktion gemeint, bei der die einzelnen bedeutungstragenden Einheiten entweder von so kurzer Dauer sind oder aber sich in minimalen Veränderungen (irgendwas stimmt nicht, ist minimal aus dem Takt) zeigen, die entweder für die normale wissenschaftliche Beobachtung (mittels Augen und Ohr) nicht wahrnehmbar, aber auf jeden Fall kaum erinnerbar und damit auch nicht für die Analyse verfügbar ist« (Reichertz 2018, S. 288). Er hält die »Beachtung und Vermessung des Nanobereichs […] für die Sozialforschung von enormer Bedeutung« (ebd.). Aus praxistheoretischer und wahrscheinlich auch aus ethnomethodologischer Perspektive ist die von Reichertz hier vorgenommene Dramatisierung des Unterschieds zwischen einem Mikrobereich der alltäglichen Abstimmung und einem Nanobereich, der sich durch Flüchtigkeit (von kurzer Dauer, minimal wahrnehmbar und kaum erinnerbar) auszeichnet, irreführend. Denn schließlich war es das Verdienst von Ethnomethodologie und Symbolischem Interaktionismus genau auf flüchtige Ereignishaftigkeit der sozialen Praktiken im Mikrobereich alltäglicher Interaktionen aufmerksam zu machen und darauf mit einer Methodik der genauen Sequenzanalyse von Praktiken (Turns, Acts, Moves etc.) zu antworten. Das bedeutet: Reichertz fügt einer ereignisgenau verfahrenden Methodologie nichts systematisch Neues hinzu. Es ließe sich allenfalls darüber streiten, ob durch die Möglichkeit audiovisueller Aufzeichnung Beobachtungschancen hinzutreten, noch kürzere Einheiten zu fassen, wobei sich aus ethnomethodologischer Perspektive fragen lässt, ob dies dann von den Akteuren im Alltag selbst noch registriert werden und daher für den Vollzug sozialer Praxis überhaupt relevant sein kann. Daher sollte kritisch nach

dem Ertrag der Nanoperspektive für die empirische (datengetriebene) Gegenstandsanalyse gefragt werden. Nun zum Beispiel:

Ethnographische Beschreibung welcher Praxis: Worin besteht nun der Ertrag der aufwändigen Studie und wie wurde er generiert? Reichertz (2018) verweist auf ein wichtiges methodisches Problem hinsichtlich des Zugangs zur untersuchten Praxis, der Gewalteskalation bei Massenveranstaltungen, speziell zwischen Fußballfangruppen und anderen Akteuren (z. B. andere Fangruppen, Polizei). Für die Videographie bestehe hier das Problem, dass in solchen Gruppen Filmaufnahmen durch Externe verpönt seien. Dadurch kam es in der Studie »Emotion. Eskalation. Gewalt« zu einer Verschiebung der untersuchten Praxis. Nicht die Gewaltpraxis selbst, sondern der »Imagefilm«, also die Selbstbewerbung der Fangruppen, in denen auch zusammengeschnittene Gewaltszenen vorkamen, wurde untersucht (vgl. Meißner 2018). Sicher ist auch der Imagefilm als kommunikative Handlung eine Form von Praxis, eine Handlung der Kamera, wie Reichertz sagen würde, die von jemandem praktiziert werden muss und als soziale Praxis sui generis Momente der Ordnungsbildung aufweist. Insbesondere dann jedoch, wenn in solchen Imagefilmen die »montierende Kamera« (Reichertz 2019) dominiert, wird es schwer aufgrund des Filmmaterials auf die darin vermittelte Handlung bzw. Praxis vor der Kamera zurückzuschließen. Insofern wurden dem Material der auf YouTube auffindbaren Fanvideos weitere Datenquellen zur Seite gestellt: Expert:inneninterviews mit Polizeipersonal, Feldbeobachtungen sowie auch kartographisches Material, um die in den Videos gezeigten Szenen als faktisches Territorium rekonstruieren zu können.

Eine solche ethno- und kartographische ›Begleitforschung‹ (Bettmann 2018) ist notwendig, da nur so das räumlich-physische Dispositiv nachvollziehbar wird, in das die zusammengeschnittenen Sequenzen der gefilmten Praxis eingebettet sind, und aus dem sie sich in ihrer Prozessdynamik entfalten. Diese Ethnographie bezieht sich auf einen Imagefilm einer Fangruppe von Fortuna Düsseldorf, die sich auf Begebenheiten im Umfeld eines Spiels beim MSV Duisburg zugetragen haben. Leider wird die Rekonstruktion der »Nanobereiche« der Entstehung von »Ereignissen« im Rahmen so gefilmter Praxis in der von Reichertz und Keysers publizierten Studie an einem anderen Fall, einem Imagevideo zum Spiel zwischen Fortuna Düsseldorf und Hansa Rostock rekonstruiert (Pellner 2018). Die in Bettmann (2018) vorgenommene Ethnographie für das Spiel zwischen Düsseldorf und Duisburg wird allerdings in einer anderen Publikation auf der Basis der »Feldpartitur« und mit Bezug auf den Nano-Bereich des Sozialen durch Moritz (2019) gespiegelt.

Die im Prinzip sinnvoll angelegte und in ihrem theoretischen Ertrag in Reichertz (2018) luzide und differenziert beschriebene Untersuchung verdeutlicht jedoch nach wie vor bestehende Schwächen in der Umsetzung einer praxisanalytischen Perspektive in der aktuellen Qualitativen Videoforschung. Die punktuell interessanten Ergebnisse zur Emergenz der gefilmten Ereignisse (Pellner 2018),

die sehr nützlichen Rekonstruktionen des räumlichen Dispositivs in Bettmann (2018), die Interpretation der im Imagefilm sich dokumentierenden Stilisierung als widerständige und solidarische Fangemeinschaft (Meißner 2018) und die theoretische Gesamtdeutung durch Reichertz (2018) bleiben nur lose miteinander verkoppelt, sodass sich der »Bedeutsamkeitszusammenhang« (Mannheim 1980) der Praxis nicht präzise aus dem analysierten empirischen Material herleiten lässt.

5. Entscheidungspunkte einer interpretativen Analyse gefilmter Praxis

Warum ist das so? Zunächst ließen sich die beschriebenen Mängel auf ein allzu hohes Ziel zurückführen, das gerade von der Verheißung ausgeht, mit Videoanalysen könnten Nanobereiche des Sozialen in ihren eigenartigen und bisher nicht erforschten Emergenzbedingungen freigelegt werden. Mir scheinen die Probleme jedoch eher damit zusammenzuhängen, dass es der Qualitativen Videoforschung an praktikablen Faustregeln mangelt, anhand derer die gesicherten Erträge der Untersuchungen in unterschiedlichen Dimensionen, die sich in Filmmaterial manifestieren und die Resultate der unterschiedlichen Verfahren, die ihnen ergänzend zur Seite gestellt werden müssen, sortiert und vergleichend gegenübergestellt werden können, um ein zusammenhängendes Bild der Empirie außenstehenden Forscher:innen nachvollziehbar zu machen. Folgende Punkte der Entscheidung im Forschungsprozess sollten daher stets in Zwischenrunden der Untersuchung zur vergleichenden Bilanzierung herangezogen werden:

(a) Die zentrale Entscheidung über den Gegenstandsfokus einer Untersuchung von audiovisuellen Aufzeichnungen sollte immer wieder in solchen Runden zum Abgleich mit der vollzogenen Forschung und deren Ergebnisse in Erinnerung gerufen werden: Welche Praxis soll primär untersucht werden? Am Beispiel der Studie von Reichertz und Keysers (2018) hätte stärker beachtet werden können, ob die gefilmte Gewalteskalation in Massen oder die Produktion eines Imagefilms im Zentrum stehen soll. Dabei ist nicht grundsätzlich ausgeschlossen, dass auch das Material der Imagefilme zur Rekonstruktion von Gewaltexzessen beitragen kann, nur müsste dazu der analytische Fokus konsequent beibehalten werden.
(b) Wichtig ist zudem die Verhältnisbestimmung der in Videos dokumentierten symbolischen Praxis zu ihrer materiell-physischen Einbettung, im Beispiel: das Verhältnis der symbolischen Codierungen der Bilder im Imagefilm und den ethno- und kartographisch ermittelten Erkenntnissen über das räumliche Dispositiv, in dem die symbolische Praxis eingebettet ist. In methodisch-interpretativer Hinsicht stellt sich dabei auch jeweils die Frage, welches Deu-

tungswissen zur Entschlüsselung der symbolischen Praxis und welches Weltwissen zur Rekonstruktion der räumlich-physischen Ausstattung des Praxiskontexts erforderlich sind.

(c) Ganz besonders wichtig ist, gerade im Hinblick auf die Auszeichnung einer mit Videoanalyse verbundenen Nanoskopie, in welcher Perspektive die Entstehung von Ordnung untersucht werden soll. Nur so kann die Vielfalt der verschiedenartigen Informationen, die sich in audiovisuellen Aufzeichnungen synchron in sehr heterogenen Dimensionen und diachron in der Zeit dokumentieren, angemessen aufeinander bezogen werden.

(d) Hierbei gilt es zu rekonstruieren, in welche Praxisumläufe mit je welcher Reichweite der Rückkoppelung die untersuchte Praxis eingebunden ist sowie welche räumlichen und zeitlichen Reichweiten der Praxis mit ihr einhergehen. Nur so können die Momente der Praxis in ihrer Rückbezüglichkeit einerseits unabhängig voneinander untersucht und daran anschließend in ihrem Zusammenspiel als »Bedeutungszusammenhang« sozialer Praxis rekonstruiert werden. Kurzum: In der Qualitativen Videoanalyse gibt es nach wie vor noch viel zu tun, nicht zuletzt in praxisanalytischer Absicht.

Literatur

Barad, Karen (2012): Agentieller Realismus. Berlin: Suhrkamp.

Bergmann, Jörg (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. Aufzeichnungen als Daten der interpretativen Soziologie. In: Bonß, Wolfgang/Hartmann, Heinz (Hrsg.): Entzauberte Wissenschaft. Sonderband 3 der Sozialen Welt. Göttingen: Otto Schwartz, S. 299–320.

Bergson, Henri (1908): L'Evolution creatrice. Paris: F. Adlon.

Bettmann, Richard (2018): Memo vom Fußballspiel Fortuna Düsseldorf vs. MSV Duisburg. In: Reichertz, Jo/Keysers, Verena (Hrsg.): Emotion. Eskalation. Gewalt. Weinheim und Basel: Beltz Juventa, S. 62–70.

Bourdieu, Pierre (1987): Sozialer Sinn. Frankfurt am Main: Suhrkamp.

Bourdieu, Pierre (1990): Was heißt sprechen. Wien: Braumüller.

Corsten, Michael (2020): Zur symbolischen Geltungsweise sozialer Praxis. In: Zeitschrift für Soziologie 49, H. 1, S. 11–27.

Danto, Artur (1973): Analytische Philosophie der Geschichte. Frankfurt am Main: Suhrkamp.

Garfinkel, Harold (1967): Studies in Ethnomethodology, Englewood-Cliffs: Prentice-Hall.

Goffman, Erving (1974): Frame analysis: An essay on the organization of experience. Cambridge: Harvard University Press.

Günthner, Susanne/Knoblauch, Hubert (1994): Gattungsanalyse. In: Hitzler, Ronald/Honer, Anne (Hrsg.): Sozialwissenschaftliche Hermeneutik. Opladen: Leske + Budrich, S. 281–307.

Hecht, Michael (2009): Das Heben und Senken eines Armes. Ablauf einer ethnomethodologischen Konversationsanalyse von Videos aus deutschen und kanadischen Schulen. In: Corsten, Michael/Krug, Melanie/Moritz, Christine (Hrsg.): Videographie praktizieren. Wiesbaden: Springer, S. 119–137.

Horn, Daniela/Houben, Sebastian/Schöner, Gerald (2018): Erste Ansätze zur automatischen Erkennung von Gruppenverhalten mithilfe des Computersehens. In: Reichertz, Jo/Keysers, Verena (Hrsg.): Emotion. Eskalation. Gewalt. Weinheim und Basel: Beltz Juventa, S. 130–147.

Horn, Daniela/Ibisch, André/Tschentner, Marc (2019): Automatisierte Videoanalyse. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer, S. 445–455.

Husserl, Edmund (1928): Vorlesungen zur Phänomenologie des inneren Zeitbewußtseins. Tübingen: Max Niemeyer.
Keller, Reiner/Knoblauch, Hubert/Reichertz, Jo (2013): Kommunikativer Konstruktivismus. Theoretische und empirische Arbeiten zu einem neuen wissenssoziologischen Ansatz. Wiesbaden: Springer.
Latour, Bruno (2008): Wir sind nie modern gewesen. Versuch einer symmetrischen Anthropologie. Frankfurt am Main: Suhrkamp.
Leontiev, Aleksej N. (2012): Tätigkeit. Bewusstsein. Persönlichkeit. Berlin: Lehmanns.
Mannheim, Karl (1980): Strukturen des Denkens. Frankfurt am Main: Suhrkamp.
Meißner, Joanna (2018): Auf geht's Fortuna: Kämpfen und Siegen! Der Imagefilm »Derby-Chaos in Duisburg« der Fortuna-Düsseldorf-Fanszene als kommunikative Handlung. In: Reichertz, Jo/Keysers, Verena (Hrsg.): Emotion. Eskalation. Gewalt. Weinheim und Basel: Beltz Juventa, S. 119–128.
Moritz, Christine (2009): Die Feldpartitur. Wiesbaden: Springer.
Moritz, Christine (2019): »Massenfieber« – Methodische Annäherungen an emotionale Intensität unter dem Aspekt der Nano-Kommunikation in großen Menschenansammlungen. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer, S. 331–364.
Moritz, Christine/Corsten, Michael (Hrsg.) (2019): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer.
Nassehi, Armin (2011): Gesellschaft der Gegenwarten. Berlin: Suhrkamp.
Pellner, Lara (2018): Fußball. Eskalation. Gewalt. Rekonstruktion eines Ereignisses. In: Reichertz, Jo/Keysers, Verena (Hrsg.): Emotion. Eskalation. Gewalt. Weinheim und Basel: Beltz Juventa, S. 72–90.
Reichertz, Jo (1992): Aufklärungsarbeit. Stuttgart: Enke.
Reichertz, Jo (2018): Masse, Kommunikation, Eskalation, Gewalt. Versuch, einen sozialen Prozess zu beschreiben. In: Reichertz, Jo/Keysers, Verena (Hrsg.): Emotion. Eskalation. Gewalt. Weinheim und Basel: Beltz Juventa, S. 203–348.
Reichertz, Jo (2019): Die Kunstlehre der wissenssoziologisch-hermeneutischen Videointerpretation. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer, S. 101–119.
Reichertz, Jo/Keysers, Verena (Hrsg.) (2018): Emotion. Eskalation. Gewalt. Weinheim und Basel: Beltz Juventa.
Sacks, Harvey/Jefferson, Gail (1992): Lectures on Conversation. Oxford: Blackwell.
Schütz, Alfred/Luckmann, Thomas (1979): Strukturen der Lebenswelt. Band 1. Frankfurt am Main: Suhrkamp.
Searle, John (1995): The Construction of Social Reality. New York: Free Press.
Stengers, Isabell (2008): Spekulativer Konstruktivismus. Berlin: Merve.

II. Videographische und medienanalytische Ansätze mit Vernakularen Videos natürlicher sozialer Situationen und inszenierter Inhalte

1. Videographische Vernakulare Videos

Ausbruch und Verharmlosung von Gewalt

Kontextualisierungen eines YouTube-Videos

Ulrike T. Kissmann

Zusammenfassung: Mithilfe der Videohermeneutik wird eine spontan-authentische Live-Aufnahme des Videoportals YouTube ausgewertet. Auf dem Video ist ein eskalierender Konflikt zwischen einer weißen Ladendiebin und einem schwarzen Drogerieangestellten in den USA zu sehen. Die Analyseergebnisse machen verschiedene Ebenen der Kontextualisierung des YouTube-Videos sichtbar. Dabei handelt es sich einerseits um die Rahmung des Videos bei der Bereitstellung auf dem Videoportal. Der Titel des Videos »Lady goes crazy after she shoplifts« suggeriert, dass die Ladendiebin nicht vorsätzlich handelt. Im Gegensatz dazu kann die Feinanalyse zeigen, dass die Ladendiebin sehr bewusst handelt und absichtlich die Situation von Täter und Opfer umkehrt. Der Beitrag präsentiert die situationsübergreifenden Kontexte, auf die die Ladendiebin in der Interaktion Bezug nimmt. Diese Kontextualisierungen ermöglichen ihr, dem schwarzen Angestellten den Status des Straftäters zuzuschreiben gemäß dem rassistischen Stereotyp, dass schwarze Männer Verbrechen begehen.

Schlüsselwörter: Mikroanalyse von Gewalt, Rassismus, Videohermeneutik, Multimodalität, Zugehörigkeitsmarkierungen

1. Einleitung

Aktuelle qualitative Videoanalysen von Gewalt nutzen vermehrt spontan-authentische Live-Aufnahmen, die mit Smartphone-Kameras angefertigt und in Videoportalen öffentlich zugänglich gemacht wurden (vgl. z. B. Braunsmann/Wagner 2018; Hoebel 2022; Weenink/Tuma/van Bruchem 2022). Im Gegensatz zu sogenannten »pranks« (engl. für Streiche) werden sie spontan und authentisch, also ohne Absprachen mit den gefilmten Personen, aufgenommen. Für die Wissenschaft ergeben sich daraus neue Möglichkeiten der Datennutzung, weil gewaltförmige Handlungen sonst nur schwer für die Analyse zugänglich sind (vgl. Meyer/Grauert/Oberzaucher i. d. B.). Gleichzeitig beinhaltet dieses Format audiovisueller Daten spezifische Erhebungs-, Analyse- und Nutzungsbedingungen, die bei der Untersuchung von Gewalthandlungen und der Veröffentlichung zu berücksichtigen sind (vgl. z. B. Knoblauch/Wilke 2018). In diesem Beitrag wird ein YouTube-Video analysiert, das 2017 in den USA aufgenommen und unter

dem Titel »Lady goes crazy after she shoplifts« auf das Videoportal hochgeladen wurde. In Deutschland würde seine Veröffentlichung dagegen einen Rechtsverstoß darstellen, weil die Persönlichkeitsrechte der abgebildeten Personen gefährdet wären. Für die dauerhafte Archivierung des Weblinks in der digitalen Bibliothek »Internet Archive« im Rahmen dieses Beitrags wurde deshalb das Video selbst nicht gespeichert.[1] Auf diese Weise lässt sich die Existenz der URL belegen, ohne jedoch durch die Weiterverbreitung des Videos die Persönlichkeitsrechte der gefilmten Personen zu verletzen. Der Artikel geht der Frage nach, wie die Erhebungskontexte des YouTube-Videos in den Analysekontexten relevant werden. Dafür werden die Verhaltensweisen der filmenden und der gefilmten Personen sowie ihr gegenseitiger Bezug aufeinander rekonstruiert. Welche Sinnzusammenhänge werden während und nach den gewaltförmigen Handlungen hergestellt? Gibt es situationsübergreifende Kontexte, auf die in der Interaktion verwiesen werden?

Gegenwärtig werden Kontextualisierungen von Gewalt verstärkt in der Soziologie analysiert und besprochen. Viele Diskussionen schließen an die Denkfigur einer dritten Instanz an und bezeichnen damit nicht nur Beobachter*innen von Gewalt, sondern auch symbolische Ordnungen und andere Sinnzusammenhänge, die den Ausbruch von Gewalt bezeugen und kommunikativ zum Ausdruck bringen (vgl. z. B. Lindemann 2017; Coenen/Tuma 2022; Koloma Beck 2011; Reemtsma 2008). Diese Ansätze gehen über eine bloße Situationsanalyse von Gewalt hinaus, weil gewaltförmige Handlungen nicht nur als unmittelbare körperliche Reaktion auf eine Situation konzipiert werden wie oftmals in der Tradition von Randall Collins (2008) (vgl. z. B. Weenink 2014; Lindegaard et al. 2017). Stattdessen analysieren die Ersteren den gesamten Verlauf der Interaktion in seiner sequenziellen Abfolge und beziehen die Kontexte mit ein, auf die die Interaktionspartner*innen in der Auseinandersetzung selbst oder danach verweisen. In dem Beitrag werden zwei verschiedene Ebenen der Kontextualisierung des YouTube-Videos vorgestellt, die die Handlung jeweils unterschiedlich rahmen und ihr eine diametral entgegengesetzte Bedeutung verleihen. *Erstens* verleiht die Person mit den Urheberrechten dem Video eine Rahmung, indem sie es auf das Videoportal hochlädt und mit dem Titel »Lady goes crazy after she shoplifts« versieht. Der Titel suggeriert, dass die Ladendiebin irrational ist und nicht bewusst handelt. Die hermeneutische Feinanalyse fokussiert dagegen die Kontextualisierungen der teilnehmenden Personen während der spontan-authentischen Live-Aufnahme. Sie zeigt, wie der Konflikt zwischen einer weißen Ladendiebin und einem schwarzen Drogerieangestellten entsteht und präsentiert die Umstände, die zu einem körperlichen Angriff führen. Obwohl der Titel des Videos von einer Ladendiebin spricht, die durchdreht, wird bei näherer Betrachtung des

1 https://web.archive.org/web/20240505074438/https://www.youtube.com/watch?v=PFdFO-6XIyCk.

Handlungsverlaufs deutlich, dass sie vorsätzlich handelt. In diesem Beitrag werden *zweitens* die Kontextualisierungen der Teilnehmenden beschrieben, die es ermöglichen, dass die Ladendiebin ihre Position als Straftäterin infrage stellen und die Situation umkehren kann. Durch Hinzunahme der Erhebungskontexte in die Analyse der audiovisuellen Daten wird deutlich, wie alle Beteiligten zu einer gemeinschaftlichen Deutung der Situation beitragen. Ihr konzertiertes Verhalten belegt, dass die weiße Ladendiebin übergeordnete Kontexte jenseits der Situation mobilisiert, um den schwarzen Angestellten als Kriminellen darzustellen, entsprechend dem rassistischen Vorurteil, dass schwarze Männer Verbrechen begehen.

Im nächsten Abschnitt wird in die Videohermeneutik eingeführt und ihre spezifische Vorgehensweise erläutert. Der dritte Abschnitt legt den Schwerpunkt auf die Erhebungskontexte und wie sie für die Analyse relevant werden. Gegenstand der Untersuchung sind die Kameraführung und ihr Beitrag zum Wendepunkt im Konflikt. Im vierten Abschnitt werden die situationsübergreifenden Kontextualisierungen präsentiert, auf die die Ladendiebin verweist, um das Verhältnis von Täter und Opfer umzukehren. Schließlich wird dadurch deutlich, wo die Gewalt in der Interaktion beginnt.

2. Zur Methode der Videohermeneutik

Die soziologische Hermeneutik basierte ursprünglich auf der Untersuchung von sprachlicher Verständigung (vgl. z. B. Soeffner 1984; Oevermann 1983). Erst seit einigen Jahren werden auch Bilder und Videos für die Analyse von Kommunikation und Interaktion genutzt (vgl. z. B. Raab 2008; Reichertz/Englert 2011; Kissmann 2014; Müller/Soeffner 2018). Das Hauptmerkmal der Videohermeneutik stellt die Feinanalyse von videoaufgezeichneten Interaktionen dar. Darin wird eine ausgewählte Sequenz des Videos in 5-Sekunden Schritten sequenziell, also Schritt für Schritt analysiert. In jedem 5-Sekunden-Segment werden Gedankenexperimente durchgeführt, um Handlungsoptionen zu entwickeln, die möglicherweise im folgenden Segment auftreten könnten. Auf diese Weise ist es möglich, in jedem neuen Segment Interpretationsmöglichkeiten auszuschließen und einen Punkt zu erreichen, an dem eine einzige Interpretation übrigbleibt. Ein weiteres Merkmal der Videohermeneutik besteht in ihrer multimodalen Vorgehensweise, in der die visuellen und die hörbaren Elemente einer Interaktion separat analysiert werden. Dadurch kann der spezifische Beitrag von Geräusch, Gespräch und körperlichem Verhalten zur Entstehung von gewaltförmigem Verhalten rekonstruiert werden. In der Feinanalyse wird dafür zunächst in 5-Sekunden-Schritten die nonverbale Interaktion ausgewertet, um dann auch das transkribierte Gespräch, einschließlich der Geräusche in 5-Sekunden-Schritten zu analysieren. Anschließend werden die Ergebnisse

zusammengeführt und kontrastierend verglichen, damit der gesamte sequenzielle Ablauf der Interaktion mit seinen hörbaren und visuellen Elementen nachvollzogen werden kann.

Das hier präsentierte Video hat eine Gesamtlänge von 4 Minuten und 22 Sekunden. Es wurde immer wieder durchgesehen, in Zeitlupe und mit Pausen. Auf dieser Grundlage wurde die Entscheidung getroffen, die ersten 30 Sekunden für die Feinanalyse auszuwählen, weil sich zeigte, dass der Anfang konstitutiv für die Begegnung der weißen Ladendiebin und des schwarzen Drogerieangestellten ist. Der Beginn des vorliegenden Filmmaterials entspricht dem Moment, in dem die filmende Person sich für die Begegnung der beiden interessierte und begann, sie mit einer Smartphone-Kamera aufzuzeichnen. Für die hermeneutische Feinanalyse wurde die 30-Sekunden-Sequenz in 5-Sekunden-Segmente unterteilt. Für jedes Segment wurden so viele Deutungen wie möglich mit entsprechenden Gegenlesarten und ihren Folgedeutungen aufgestellt, um sie im folgenden Segment überprüfen zu können. Bei diesen Gedankenexperimenten wird das Wissen über das Folgegeschehen eingeklammert, d. h. die Forschenden tun so, als würden sie nicht den weiteren Verlauf kennen. Durch das sequenzielle Vorgehen werden in jedem neuen Segment die bisherigen Deutungen überprüft und gegebenenfalls ausgeschlossen. Auf diese Weise kommt es zu einer allmählichen Schließung von Deutungen, sodass die finale Interpretation als valide betrachtet werden kann.

3. Erhebungskontexte: Die Kameraführung

Betrachtet man zunächst das allererste Videostandbild, dann offenbart sich ein Blick in das Innere eines Drogeriemarktes. Das Bild ist leicht gekippt und im unteren linken Bereich unscharf. Es sind keine Personen zu sehen. Stattdessen ist die Kamera auf ein Regal mit Drogerieprodukten gerichtet, ohne diese richtig zu fokussieren. In der Analyse der nonverbalen Interaktion wird zunächst die Kontextinformation eingeklammert, dass es sich um einen Ladendiebstahl handelt. Dies unterstützt die interpretierenden Personen dabei, sich auf basale Bewegungen wie z. B. die Kameraführung zu fokussieren. Erst im zweiten Schritt wird die nonverbale Interaktion zusammen mit der Kontextinformation analysiert, dass es sich bei dem Video um einen Konflikt zwischen einer Ladendiebin und einem Drogerieangestellten handelt.[2] Das erste 5-Sekunden-Segment

2 Durch die Einklammerung des Kontexts lassen sich im ersten Schritt basale Körperbewegungen identifizieren, die im Alltag vorreflexiv zugänglich sind und auf praktischem Wissen beruhen. Erst im zweiten Schritt geht die Feinanalyse darüber hinaus und greift auf das theoretische Wissen zurück, was Ladendiebe und Angestellte in einem Fall von

zeigt auf diese Weise, dass sich die Kamera etwas impulsiv und unkoordiniert über das Regal hinausbewegt. Die Aufnahme schwankt und ist schließlich auf zwei Personen gerichtet, eine Person, die als junge weiße Frau gelesen wird, und eine andere Person, die als schwarzer Mann gelesen wird, die sich in der Mitte eines Ladens gegenüberstehen. Sie halten beide eine Tasche fest mit den Händen umklammert. Obwohl die Kamera auf sie gerichtet ist, ist der Fokus nicht vollständig auf sie eingestellt. Die Feinanalyse ergab, dass es sich hierbei um eine spontane Alltagssituation handelt, weil die Kameraführung zögerlich und weder geplant noch professionell ist. Während der Anlass der Aufnahme ursprünglich zufällig war, ergaben die Gedankenexperimente, dass das Video absichtlich fortgeführt wurde, weil das anfängliche starke Schwanken aufhört, nachdem sich die Kamera auf das Paar konzentriert hat. Die Interaktion der beiden hat also das Interesse der filmenden Person geweckt. Gleichwohl bleibt ein leichtes Zittern der Kamera bestehen. Es hört erst auf, nachdem ein Wendepunkt in der Interaktion eintritt. Wie im folgenden Abschnitt ersichtlich wird, verweist der Erhebungskontext, also die Kameraführung auf die Zuspitzung im Konflikt der beiden gefilmten Personen.

Wird die sequenzielle Analyse ohne die Kontextinformation weitergeführt, dass es sich um einen Ladendiebstahl handelt, dann beschränken sich die Gedankenexperimente darauf, dass sich der schwarze Mann und die weiße Frau gegenüberstehen und beide den Griff einer Tasche fest umklammert halten. Der Mann spricht mit der jungen Frau, während sie ihn anschaut. Sie scheint sich für etwas in der Tasche zu interessieren, da sie ihre Hand in die Tasche taucht und versucht, etwas herauszunehmen. Es gelingt ihr jedoch nicht und sie zieht ihre Hand zurück. Von diesem Moment an beginnt die Frau, mit Gewalt an der Tasche zu zerren, um sie in ihren Besitz zu bringen. Der schwarze Mann hält die Tasche fest. Die Feinanalyse ergab, dass er dafür keine körperliche Kraft anwenden muss. Da er viel größer und stärker als die junge weiße Frau ist, kann er aufrecht und mit Gleichgewicht auf beiden Füßen stehen. Im Gegensatz dazu muss die junge Frau ihren ganzen Körper einsetzen, um die Tasche zu ziehen (siehe Abb. 1):

Ladendiebstahl üblicherweise tun. Die beiden Analyseschritte stellen unterschiedliche Formen von Sozialität dar, die in natürlichen Umgebungen miteinander verbunden sind, aber unabhängig voneinander untersucht werden können.

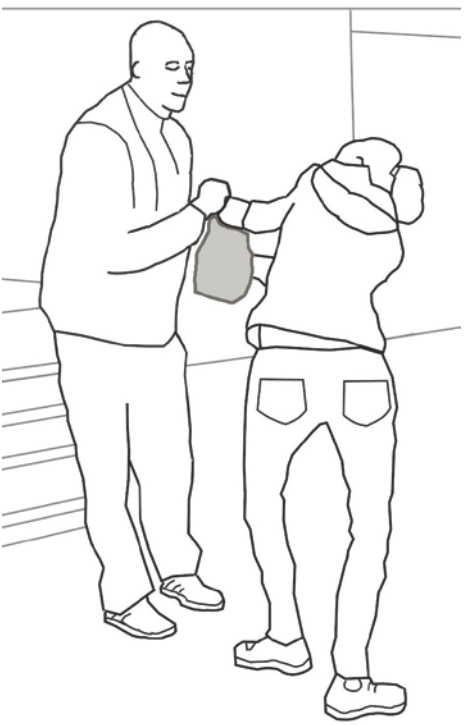


Abbildung 1: Ungleichgewicht der Kräfte zwischen den beiden Personen (Copyright: Autorin)

Außerdem haben die Gedankenexperimente gezeigt, dass der Mann ruhig und geduldig bleibt. Er hat die Situation unter Kontrolle, während die Bemühungen der Frau wirkungslos sind. An diesem Punkt hätte sie aufgeben können und der Konflikt wäre beendet gewesen. Am Ende des dritten Segments kommt es jedoch zu einem plötzlichen Wendepunkt in der Interaktion. Die weiße Frau schafft es, ihre körperliche Schwäche zu kompensieren und durch mündliche Rede die Oberhand zu gewinnen. Bevor sie spricht, stellt sie sich aufrecht und von Angesicht zu Angesicht dem schwarzen Mann gegenüber. In diesem Moment wird die Kamerabewegung gleichmäßig und das leichte Zittern hört auf. Die filmende Person hat den Wendepunkt antizipiert und wartet auf die Aussage der Frau. Der Erhebungskontext, also die Kameraführung verweist auf den spezifischen Bedeutungsgehalt dieser Situation für die Entstehung von Gewalt. In Abschnitt 4 wird gezeigt werden, dass mit der Aussage der Frau die Gewalt beginnt. Auch der weitere Verlauf der nonverbalen Interaktion hebt hervor, dass sie etwas auf den Punkt bringt, das wichtig sein muss, denn der schwarze Mann wendet sofort den Kopf von ihr ab. Die Feinanalyse ergab, dass diese Geste ein Ausdruck starker Missbilligung ist. Aus dem nächsten Abschnitt wird hervorgehen, dass diese basale Körperbewegung die Reaktion auf das rassistische Verhalten der weißen

Frau darstellt. Zunächst wird jedoch die bisherige Interaktion mit der Kontextinformation betrachtet, dass der schwarze Mann ein Drogerieangestellter und die weiße Frau eine Ladendiebin ist, die gerade etwas gestohlen hat.

Die Kontextinformation führt spätestens im dritten Segment zu dem Ergebnis, dass es sich um die Tasche der weißen Ladendiebin handelt, die beide mit den Händen umklammern. Dadurch dass der schwarze Drogerieangestellte die Tasche festhält, verhindert er, dass die Ladendiebin flieht. Der Inhalt der Tasche erweist sich als relevant für die junge Frau. Dies ist der Grund, warum sie nicht ohne diese Tasche fliehen kann. In dem sich abzeichnenden Konflikt verhält sich der schwarze Angestellte nicht invasiv und beschränkt sich strikt auf seine Funktion als Drogeriemitarbeiter, da er weder die Handgelenke noch die Arme der Frau festhält, sondern einfach die Tasche fest umklammert. Während in der bisherigen Feinanalyse das Ungleichgewicht der körperlichen Kräfte vorherrschte, wird mit der Kontextinformation, dass es sich um einen Ladendiebstahl handelt, in der Interpretation verstärkt darauf abgehoben, wer im Recht ist. Der schwarze Angestellte ist aufgrund seiner Funktion als Vertreter des geschädigten Geschäfts im Recht. Die weiße Ladendiebin hingegen sieht sich im Recht, obwohl die Situation ihr keinen Anlass dazu gibt. Indem sie durch ihr Verhalten darauf besteht, dass sie im Recht ist, macht sie deutlich, dass sie die Stellung des schwarzen Mannes nicht anerkennt. Ein Gedankenexperiment führte zu der Interpretation, dass ihr Verhalten als unbedacht und überstürzt angesehen werden könnte. Dies erwies sich jedoch als unzutreffend, da ihr Verhalten in allen Segmenten deutlich zeigt, dass sie ganz bewusst handelt.[3] Die Ergebnisse unterstreichen, dass die Ladendiebin glaubt, einen guten Grund zu haben, die Stellung des schwarzen Angestellten zu delegitimieren und davon auszugehen, dass der Konflikt zu ihren Gunsten gelöst werden kann. Worauf bezieht sich die weiße Frau, um sich im Recht zu sehen?

Elijah Anderson (1999, S. 16 ff.) hat in einer ethnographischen Studie nachdrücklich dargestellt, dass Schwarze in US-amerikanischen Städten nur selten in Geschäften in Wohngegenden beschäftigt sind, die überwiegend aus wohlhabenden und gebildeten Weißen bestehen. Außerdem werden schwarze Männer argwöhnisch betrachtet, wenn sie als Kunden teure Einrichtungen wie z. B. Juweliergeschäfte betreten, entsprechend dem rassistischen Stereotyp, dass schwarze Männer Verbrechen begehen. Obwohl seit seiner Studie diese Milieus zunehmend ethnisch gemischt sind, ist davon auszugehen, dass das oben angegebene rassistische Stereotyp fortbesteht und schwarze Männer von Weißen problemlos als Kriminelle angesehen werden können. Da im vorliegenden Video die Umstehenden in der Drogerie weiß sind, ist es sehr wahrscheinlich, dass das Filmmaterial

3 Die zweckgerichteten Handlungen der Ladendiebin wurden insbesondere daran deutlich, dass sie bewusst ihre Hand in die Tasche taucht, um etwas herauszuziehen. Da dies nicht gelingt, versucht sie in der Folge auf eine andere Weise, ihr Problem zu lösen.

in einer der genannten Wohngegenden mit überwiegend wohlhabenden und gebildeten Weißen aufgenommen wurde.[4] Aus diesem Grund ist die Stellung des schwarzen Drogerieangestellten in einer solchen Gegend nicht durchschnittlich. Infolgedessen ist davon auszugehen, dass die bestehenden ethnischen Beziehungen die weiße Ladendiebin dazu veranlasst haben, sich in der Lage zu sehen, den Status des schwarzen Angestellten zu delegitimieren. In Abschnitt 4 wird präsentiert werden, wie sie am Wendepunkt der Interaktion ihre Zugehörigkeit zur weißen Mehrheitsgesellschaft deutlich macht und diese Kontextualisierung nutzt, um sich einer Strafverfolgung zu entziehen.

Zusammenfassend hat die Feinanalyse der nonverbalen Interaktion gezeigt, dass der Wendepunkt des Konflikts in dem Moment stattfindet, in dem die Kameraführung stabil wird und die weiße Ladendiebin eine Aussage macht. Die filmende Person hat die Wende antizipiert und sieht sich darin bestätigt, die Auseinandersetzung mit der Smartphone-Kamera festzuhalten. Der Erhebungskontext ist daher für die Analyse relevant, weil er auf den besonderen Bedeutungsgehalt dieser Situation verweist. Einerseits kompensiert die weiße Frau ihre körperliche Schwäche und es kommt zu einem Kräftegleichgewicht zwischen den beiden handelnden Personen. Andererseits belässt die weiße Ladendiebin es nicht beim Gleichklang der physischen Kräfte und deutet mit ihrem Verhalten an, dass sie glaubt, im Recht zu sein. Mit ihrer mündlichen Rede lässt die Ladendiebin schließlich keinen Zweifel daran, dass sie im Recht ist.

4. Analysekontexte: Wie die Ladendiebin die Situation umkehrte

Das transkribierte Gespräch weist eine männliche und eine weibliche Stimme auf. Von Anfang an wird die Tasche von den beiden beteiligten Personen als zentraler Bestandteil des Konflikts gesehen. Die als männlich gelesene Person fordert die andere Person auf, »you let the bag go« (siehe Tab. 1: Transkript des Gesprächs). Sie ringen um die Tasche und man hört das Quietschen von Schuhen auf dem Boden. Daraufhin ruft die als weiblich gelesene Person »no«. Die segmentweise Analyse hat gezeigt, dass die männliche Person in dem Konflikt die Oberhand hat, da sie schnell sowie überlegt reagiert und sagt: »I'm not gonna let go«. Darüber hinaus führt die männliche Person mit »we're gonna call the police and deal with this the right way« die Polizei ein. Auf diese Weise macht er deutlich, dass das Gesetz auf seiner Seite ist. Gleichzeitig ist die Haltung der männlichen Person als problemlösungsorientiert zu erkennen, da er durch die Verwendung

4 Es gibt in dem gesamten Video nur eine einzige Situation, in der andere Schwarze zu sehen sind. Bei 2 Minuten und 40 Sekunden betritt eine schwarze Familie mit kleinen Kindern die Drogerie. Vom Eingang aus sehen sie sich den Konflikt an und verschwinden dann aus dem Blickfeld der Kamera in das Innere der Drogerie.

der ersten Person Plural seinen Gegner in die Lösung des Problems einbezieht. Die Analyse des transkribierten Gesprächs hat gezeigt, dass er zusammen mit der Formulierung »deal with this the right way« eine Lösung finden und den Streit beilegen will. Obwohl die männliche Person die Kontrolle über die Situation hat, werden seine Sätze sehr schnell gesprochen. Er macht keine Pausen zwischen den Worten. Es konnte gezeigt werden, dass dies ein Ausdruck seiner Erregung und Empörung ist.

Tabelle 1: Transkript des Gesprächs (A ist die männliche Person und B die weibliche Person)

Transkribiertes Gespräch	
A: You let /(the bag go) Schuhe quietschen auf dem Boden/	
B: /NO↑ almost shouted/ /(2) Schuhe quietschen auf dem Boden/	
but, (go)	
A: I'm=not=gonna=let=go we're=gonna=call=the=police	
and=deal=with=this=the=right=way	
B: That [is	
A: [OKAY ((fragend))	
B: **not** the right way to just (1) to take someone's (2) purse and	
/(unverständliche Wörter) lautes, quietschendes Geräusch/	
A: You=are=**not** gonna=go [anywhere	
B: [**You** (ein unverständliches Wort)	
A: we=gonna ↑call the police	
B: yeah you know (what) [happens to you	
A: [you=shoplifted	
B: when you (insult) someone (2) you are going to fucking jail	
Regeln für die Transkription	
,	kurze Pause
(4)	Länge der Pause in Sekunden
((lachend))	Kommentar der Transkribierenden
/…/	Anfang und Ende des kommentierten Phänomens
not	betont
( )	Inhalt der Äußerung unverständlich
(the bag go)	unsichere Transkription
yes=yes	schneller Anschluss
yes [it was [nothing	gleichzeitiges Sprechen ab ›it‹
OKAY	mit Nachdruck
↑	Stimme wird gehoben
↓	Stimme wird gesenkt

Die weibliche Person will gerade einen neuen Satz mit »that is« beginnen, als die männliche Person sie unterbricht und »okay« fragt. Auch dies ist ein Beleg für seine problemlösende Haltung. Im Gegensatz zu seiner bereits erwähnten Verwendung des Plurals der ersten Person, der die Einbeziehung seines Gegenübers ermöglicht, legt die weibliche Person den Schwerpunkt auf eine generalisierende und abstrakte Argumentation. Sie setzt ihren Satz mit den Worten »not the right way to just to take someone's purse« fort. Damit nimmt sie das Thema Recht und Unrecht dort auf, wo die männliche Person abgesetzt hat. Aber ihre Formulierung ist keineswegs inklusiv. In den Gedankenexperimenten wurde sie vielmehr als herablassend empfunden. Die männliche Person geht nicht auf diese Äußerung ein und wiederholt stattdessen: »You are not gonna go anywhere«. Bei dem Wort »anywhere« unterbricht ihn die weibliche Person und beginnt mit Nachdruck »you« zu sagen. Es wird so artikuliert, als würde sie gleich einen Angriff starten. Betrachtet man die Verschränkung von hörbaren und visuellen Elementen der Interaktion, so stellt man fest, dass das Wort »you« genau in dem Moment geäußert wird, in dem das Gleichgewicht der physischen Kräfte zwischen den Handelnden erreicht ist. An diesem Punkt stellte sich die weiße Frau aufrecht und von Angesicht zu Angesicht dem schwarzen Mann gegenüber (siehe Abb. 2). Sowohl das körperliche Verhalten als auch der Beginn der mündlichen Äußerung sind Ausdruck des oben beschriebenen Wendepunkts in der Interaktion.

Abbildung 2: Als die weiße Frau beginnt, »you« zu sagen, stellt sie sich dem schwarzen Mann aufrecht gegenüber (Copyright: Autorin)

Die anfängliche Erregung und Empörung der männlichen Person wachsen an dieser Stelle, da er auf das nachdrückliche »you« der weiblichen Person mit der Wiederholung des Satzes »we gonna call the police« reagiert. Diesmal erhebt er jedoch seine Stimme nach »gonna« und macht damit seine Anspannung und

Wut deutlich. Bei der Analyse des gesamten Ablaufs der Interaktion mit ihren hörbaren und visuellen Elementen wurde deutlich, dass die Worte der männlichen Person von einer nickenden Geste begleitet werden. Sie ist durchweg zustimmend und unterstreicht seine Ansicht, dass der Streit nun ein Ende finden sollte. Im Kontrast dazu stehen die Ergebnisse der Analyse der nonverbalen Interaktion. Hier wurde im vorhergehenden Abschnitt gezeigt, dass die weiße Frau an dem Wendepunkt in der Interaktion ihre körperliche Schwäche kompensiert und einen wichtigen Punkt vorbringt. Sie ist davon überzeugt, dass der Konflikt zu ihren Gunsten gelöst werden kann und sieht sich im Recht. Darüber hinaus belegen die Ergebnisse der Analyse des transkribierten Gesprächs nun, dass sie sich als jemanden präsentiert, die die Deutungshoheit über (ihre Sicht der) Rechtsnormen hat. Als der Mann seinen Satz mit »the police« beendet hat, antwortet sie sofort »yeah you know what happens to you«. Auf die oben erwähnte emphatische Äußerung »you« folgt also tatsächlich ein verbaler Angriff. Die weibliche Person bedroht ihr Gegenüber. Ein solches Verhalten ist vergleichbar mit Randall Collins' »boasting« (2008, S. 345) (engl. für prahlen). Aus seiner Sicht ist der Inhalt von Prahlerei und Beleidigung nicht von Bedeutung (vgl. Collins 2008, S. 364). In diesem Beitrag wird jedoch argumentiert, dass ihr spezifischer Inhalt in der Interaktion sehr wohl von Relevanz ist. Die Prahlerei der weiblichen Person enthält eine Zugehörigkeitsmarkierung zur weißen Mehrheitsgesellschaft, die ihr den Status zuweist, Rechtsnormen zu klären. Während Anderson (1999, S. 112) Zugehörigkeitsmarkierungen von schwarzen Jugendlichen zur Gegenkultur beschrieben hat, offenbaren die Ergebnisse hier den umgekehrten Fall. Eine weiße Frau nutzt symbolische Kontextualisierungen, um in der Interaktion die Oberhand zu gewinnen und sich der Strafverfolgung zu entziehen.

Während die weibliche Person die Drohung ausspricht, stellt die männliche Person gleichzeitig mit dem Ende seiner Äußerung »you shoplifted« klar, dass sie die Täterin ist, nicht das Opfer. Damit verdeutlicht er das tatsächliche Verhältnis zwischen Täter und Opfer, zwischen Recht und Unrecht, denn bis zu diesem Zeitpunkt hat die weibliche Person so gesprochen, als ob sie im Recht wäre. Die Äußerungen der männlichen Person können als Antizipation dessen, was folgen wird, betrachtet werden. Die weibliche Person kehrt mit ihrer Äußerung »when you insult someone you are going to fucking jail« das Verhältnis zwischen Täter und Opfer, zwischen Recht und Unrecht um. Es handelt sich eindeutig um eine Prahlerei, da sie nicht in der Lage ist, Rechtsnormen zu klären. Stattdessen wird die Prahlerei als Zugehörigkeitsmerkmal zur weißen Mehrheitsgesellschaft verwendet. Die Formulierung »fucking jail« beinhaltet auch eine Beleidigung. Wie bereits erwähnt, bezieht sich die Zugehörigkeitsmarkierung auf die ethnischen Beziehungen, die der weißen Frau den Status zuweisen, im Recht zu sein und über einen schwarzen Mann urteilen zu können. Sie nimmt ihn nicht als Individuum in einem bestimmten Kontext, d. h. als Angestellten der Drogerie wahr. Vielmehr sieht die weiße Frau ihn als Stereotyp

eines schwarzen Mannes. Die Analyse der nonverbalen Interaktion hat bereits deutlich gemacht, dass sie die Position des schwarzen Mannes als Angestellten nicht anerkennt. Die multimodale Analyse des transkribierten Gesprächs bestätigt dieses Ergebnis jetzt, denn die weiße Frau wertet ihr Gegenüber entsprechend dem rassistischen Stereotyp des kriminellen schwarzen Mannes ab. Dies ermöglicht es ihr, die Situation umzukehren und dem schwarzen Mann den Status eines Straftäters zuzuschreiben.

Die Reaktion des Angestellten auf das rassistische Verhalten der Ladendiebin besteht in einer Geste der starken Missbilligung, da er sofort den Kopf von ihr abwendet (siehe Abb. 3). Die Gedankenexperimente haben deutlich gemacht, dass in dieser Geste zum Ausdruck kommt, dass seine Würde nachhaltig verletzt ist. In dem weiteren Verlauf der Sequenz wird dies dadurch belegt, dass die schwarze Person die Gefängnisdrohung der Ladendiebin genau in dem Moment aufgreift, in dem sie ihn bei 1 Minute und 32 Sekunden körperlich angreift. Auf ihren Fußtritt reagiert er mit »I really take you to jail for that«. Dies zeigt, dass der verbale Angriff der Ladendiebin »when you insult someone you are going to fucking jail« bei Sekunde 17 bis 20 eine ernsthafte Verletzung der Würde der schwarzen Person nach sich gezogen hat. Dass die Äußerung einer jungen Frau eine derartige Macht über einen gestandenen Angestellten haben kann, liegt darin begründet, dass auf situationsübergreifende Kontexte Bezug genommen wird, die die Demütigung schwarzer Personen legitimieren. Mit der Äußerung der Ladendiebin werden Zugehörigkeitsmarkierungen zur weißen Mehrheitsgesellschaft mobilisiert, die es ihr ermöglichen, die Situation von Täter und Opfer umzukehren.

Abbildung 3: Nach dem verbalen Angriff dreht der schwarze Angestellte seinen Kopf missbilligend zur Seite (Copyright: Autorin)

In der gesamten Länge des Videos von 4 Minuten und 22 Sekunden ist das Verhalten des schwarzen Angestellten durch die Vermeidung eines möglichen Verstoßes gegen das Recht geprägt. Obwohl seine Würde durch die Ladendiebin verletzt wurde, trägt er an keinem Punkt der Interaktion zur Eskalation des Konflikts bei. Stattdessen bezieht die Ladendiebin die weißen Umstehenden bewusst in die Auseinandersetzung ein. Bei 1 Minute und 19 Sekunden z. B. wendet sich die Ladendiebin einer anderen Kundin zu und beschwert sich, dass der Angestellte ihr wehtut. Doch die Kundin schaut sie nur an und scheint ihr nicht zu glauben. Als die Ladendiebin den Angestellten bei 1 Minute und 32 Sekunden tritt und dieser, wie oben beschrieben, ihre Drohung mit dem Gefängnis aufgreift, schlägt er weder zurück noch tut er ihr in irgendeiner Weise weh. Der schwarze Angestellte hält lediglich die Tasche fest und hindert die Ladendiebin auf diese Weise daran zu fliehen. Nach 1 Minute 34 Sekunden schlägt sie ihm schließlich ins Gesicht. Dem Angestellten gelingt es, den Angriff abzuwehren, indem er die Tasche zu Boden zieht, die immer noch von beiden Parteien festgehalten wird. Die weiße Frau stürzt zu Boden, bleibt aber unverletzt. Während dieser Auseinandersetzung berührt der Angestellte zum ersten Mal ihre Arme, um sie am Boden zu halten. Selbst in dieser Lage versucht die Frau zu kämpfen, aber der Angestellte drückt sie an ihren Schultern zu Boden. Die anderen Umstehenden versuchen, die Ladendiebin zu beruhigen. Bei 1 Minute und 50 Sekunden sagt eine Kundin zu ihr: »Ma'am you need to chill«. Schließlich, bei 3 Minuten und 40 Sekunden, erscheint ein weißer Polizist in der Drogerie. Von diesem Moment an hört die weiße Frau auf, sich zu wehren.

5. Zusammenfassung und Schlussfolgerung

Der Titel des YouTube-Videos »Lady goes crazy after she shoplifts« suggeriert, dass die Ladendiebin nicht vorsätzlich handelt. Damit wird versucht, ihr Verhalten zu entschuldigen. In diesem Beitrag wurde stattdessen das Ziel verfolgt, das Verhalten der Ladendiebin ernst zu nehmen. Die ersten 30 Sekunden des Videos wurden einer hermeneutischen Feinanalyse unterzogen. Es zeigte sich, dass nach 15 Sekunden ein Wendepunkt in der Interaktion eintritt. Zum einen war diese Wende in der nonverbalen Interaktion durch den plötzlichen Wechsel von einem Ungleichgewicht zu einem Kräftegleichgewicht zwischen der Ladendiebin und dem Angestellten gekennzeichnet. Zum anderen machte die Ladendiebin in diesem Moment eine Äußerung, die keinen Zweifel daran ließ, dass sie sich im Recht fühlt. Außerdem hatte die filmende Person den Wendepunkt antizipiert, da in genau diesem Moment das leichte Zittern der Kamera aufhörte und die Kamera ganz scharf gestellt wurde. Die Erhebungskontexte waren damit für die Analysekontexte relevant, weil sie auf den spezifischen Bedeutungsgehalt dieser Situation verwiesen. Die Analyse des transkribierten Gesprächs

hat gezeigt, dass die Ladendiebin das Verhältnis von Täter und Opfer umkehrte und dem schwarzen Angestellten den Status des Straftäters zuwies. Durch sogenanntes »boasting« (Collins 2008, S. 345) gab die Ladendiebin vor, die Deutungshoheit über ihre Auffassung von Rechtsnormen zu haben. Im Gegensatz zu Randall Collins (2008) hatte der Inhalt der Prahlerei jedoch einen Einfluss auf den Verlauf des Konflikts. Es wurde belegt, dass die Prahlerei als Zugehörigkeitsmarkierung zur weißen Mehrheitsgesellschaft verwendet wurde. Durch diese Rahmung in der Interaktion konnte die Ladendiebin erstens den Status erlangen, Rechtsnormen zu klären und zweitens den schwarzen Angestellten gemäß dem rassistischen Stereotyp als Straftäter abwerten. Die situationsübergreifende Kontextualisierung erzeugte ethnische Hegemonie zugunsten der weißen Ladendiebin und zum Nachteil des Schwarzen. Der Vergleich der visuellen und hörbaren Elemente der Interaktion machte deutlich, dass die weiße Ladendiebin den Wendepunkt einleitete, indem sie mit Nachdruck »you« sagte. In diesem Moment wandte sie sich dem schwarzen Angestellten zu und das Gleichgewicht der Kräfte war plötzlich erreicht. Es kann festgehalten werden, dass an diesem Punkt die Gewalt begann, weil die Ladendiebin zum verbalen Angriff überleitete. Mit der Äußerung »when you insult someone you are going to fucking jail« verletzte die Ladendiebin nachhaltig die Würde des schwarzen Angestellten. Letzterer drehte daraufhin sofort seinen Kopf missbilligend zur Seite. Die Gedankenexperimente ergaben, dass in dieser Geste seine verletzte Würde zum Ausdruck kam. Im weiteren Verlauf der Sequenz wurde diese Interpretation dadurch validiert, dass der Angestellte die Gefängnisdrohung der Ladendiebin bei 1 Minute und 32 Sekunden mit »I really take you to jail for that« aufgriff. Der verbale Angriff bei Sekunde 17 bis 20 hatte seine Würde nachhaltig beschädigt, sodass er auch noch mehr als 60 Sekunden später darauf Bezug genommen hat. Der Ausbruch von Gewalt begann mit dem »you« der Ladendiebin. Mit diesem Wendepunkt in der Interaktion nahm der Konflikt einen sehr dynamischen Verlauf und eskalierte schließlich in körperliche Gewalt.

Literatur

Anderson, Elijah (1999): Code of the Street. Anstand, Gewalt und das moralische Leben der Innenstädte. New York, London: Norton & Company.

Braunsmann, Katharina/Wagner, Felicitas (2018): Spontan-authentische Liveaufnahmen als Forschungsmaterial. Empirische Perspektiven und methodische Überlegungen im Kontext der Gewaltforschung. In: ORDEX Working Paper 2, Bielefeld: Universität Bielefeld. https://doi.org/10.4119/unibi/2930500.

Coenen, Ekkehard/Tuma, René (2022): Contextural and Contextual – Introducing a Heuristic of Third Parties in Sequences of Violence. In: Historical Social Research. Sonderband 47, H. 1, S. 200–224.

Collins, Randall (2008): Violence. A Micro-sociological Theory. Princeton and Oxford: Princeton University Press.

Hoebel, Thomas (2022): Emplotments of Violence. On Narrative Explanations and their Audiovisual Data. In: Historical Social Research. Sonderband 47, H. 1, S. 225–246.

Kissmann, Ulrike T. (2014): Die Sozialität des Visuellen. Fundierung der hermeneutischen Videoanalyse und materiale Untersuchungen. Weilerswist: Velbrück Wissenschaft.
Knoblauch, Hubert/Wilke, René (2018): Forschungsdateninfrastrukturen für audio-visuelle Daten der Qualitativen Sozialforschung – Bedarf und Anforderungen. In: RatSWD Working Paper Series Nr. 267. https://www.konsortswd.de/wp-content/uploads/RatSWD_WP_267.pdf (02.07.2024).
Koloma Beck, Teresa (2011): The Eye of the Beholder: Violence as a Social Process. In: International Journal of Conflict and Violence 5, H. 2, S. 345–356.
Lindegaard, Marie Rosenkrantz/Liebst, Lasse Suonperä/Bernasco, Wim/Bruvik Heinskou, Marie/Philpot, Richard/Levine, Mark/Verbeek, Peter (2017): Consolation in the Aftermath of Robberies Resembles Post-Aggression Consolation in Chimpanzees. In: PLoS ONE 12, H. 5: e0177725. https://doi.org/10.1371/journal.pone.0177725.
Lindemann, Gesa (2017): Verfahrensordnungen der Gewalt. In: Zeitschrift für Rechtssoziologie 37, H. 1, S. 57–87.
Müller, Michael R./Soeffner, Hans-Georg (Hrsg.) (2018): Das Bild als soziologisches Problem. Herausforderungen einer Theorie visueller Kommunikation. Weinheim und Basel: Beltz Juventa.
Oevermann, Ulrich (1983): Zur Sache. Die Bedeutung von Adornos methodologischem Selbstverständnis für die Begründung einer materialen soziologischen Strukturanalyse. In: Von Friedeburg, Ludwig/Habermas, Jürgen (Hrsg.): Adorno-Konferenz. Frankfurt am Main: Suhrkamp, S. 234–288.
Raab, Jürgen (2008): Visuelle Wissenssoziologie. Theoretische Konzeption und materiale Analysen. Konstanz: UVK.
Reemtsma, Jan Philipp (2008): Vertrauen und Gewalt. Versuch über eine besondere Konstellation der Moderne. Hamburg: Hamburger Edition.
Reichertz, Jo/Englert, Carina Jasmin (2011): Einführung in die qualitative Videoanalyse. Eine hermeneutisch-wissenssoziologische Fallanalyse. Wiesbaden: Springer.
Soeffner, Hans-Georg (1984): Strukturanalytische Überlegungen zur gerichtlichen Interaktion. In: Reichertz, Jo (Hrsg.): Sozialwissenschaftliche Analysen jugendgerichtlicher Interaktion. Tübingen: Stauffenburg, S. 189–225.
Weenink, Don (2014): Frenzied Attacks: A Micro-sociological Analysis of the Emotional Dynamics of Extreme Youth Violence. In: The British Journal of Sociology 65, H. 3, S. 411–433.
Weenink, Don/Tuma, René/van Bruchem, Marly (2022): How to Start a Fight: A Qualitative Video Analysis of the Trajectories Toward Violence Based on Phone-Camera Recorded Fights. In: Human Studies 45, S. 577–605. https://doi.org/10.1007/s10746-022-09634-6.

Videoanalytische Rassismusforschung

Christian Meyer, Hanna Grauert und Frank Oberzaucher

Zusammenfassung: Das in der Rassismusforschung üblicherweise eingesetzte Methodenset beschränkt sich vorwiegend auf befragungsintensive oder textbasierte Forschungsmethoden, die soziale Sachverhalte rekonstruktiv erfassen und keine Aussagen über die Herstellung sozialer Ereignisse erlauben. Demgegenüber diskutieren wir in unserem Beitrag die Potenziale der analytischen Nutzung von Audio- und Videoaufzeichnungen, die rassifizierende oder als rassifizierend gedeutete Kommunikationen oder Handlungen dokumentieren. Dieser Typus von Forschungsdaten liegt inzwischen vermehrt vor, weil Alltagsakteur:innen *in situ* entsprechende Situationen mit verfügbaren Aufzeichnungstechnologien (meist Smartphones) selbst dokumentieren und in sozialen Medien veröffentlichen. In diesem Typus von Forschungsdaten liegt aus unserer Perspektive ein enormes Potenzial, das dringend sozialwissenschaftlich ausgeschöpft werden sollte. Denn Video-Daten dokumentieren den Prozesscharakter rassifizierender Kommunikationen und Handlungen, wie er *live* erfolgt, und erlauben auf dieser Basis eine detaillierte soziologische Analyse der inkrementellen Verläufe sozialer Situationen, die als rassifizierend gelesen und erlebt werden. Die Analyse von Videos, die in *Social Media* zirkuliert werden, kann also soziologisch zur Erforschung adverser oder gar antisozialer Facetten menschlicher Sozialität beitragen. Die Tatsache, dass diese Videos zudem auch in verschiedenen Kontexten vernakulär (etwa durch User, die diese Videos kommentieren und über sie diskutieren) sowie professionell (in polizeilichen oder juristischen Organisationen) aus anderen als soziologischen Perspektiven analysiert werden, erweist sich als zusätzlich gewinnbringend.

Stichwörter: Qualitative Rassismusforschung, Ethnomethodologie, Konversationsanalyse, Videoanalyse, Social-Media-Daten

1. Einleitung

In den letzten Jahren sind in den sozialen Medien vermehrt Videoaufzeichnungen rassistischer Ereignisse verbreitet und vehement diskutiert worden. Die Aufzeichnungen zeigen unterschiedliche Formen von manifest gewordenem Rassismus: rassistische Polizeigewalt wie im Fall von George Floyd, aber auch rassistische Gesänge wie das jüngste Festvideo aus Sylt. Die permanente Verfügbarkeit des videografischen Aufzeichnungsmediums »Smartphone« hat zu einer

omnipräsenten Möglichkeit der Dokumentation versteckt oder offen rassistischer sozialer Handlungen und Situationen geführt (vgl. Kissmann i. d. B.). Diese Allgegenwart des videografisch dokumentierten Rassismus bietet der soziologischen Forschung die neue Gelegenheit, Rassismus in interaktionalen Situationen zu beforschen und den Fokus auf den Verlauf und Vollzug sowie die Entstehung und Herstellung von Rassismus zu legen. Damit rücken Sozialdimensionen in den Blick, die bisher in der Forschung weitgehend ausgeblendet blieben und sowohl unabhängig von individuellen Einstellungen und Intentionen als auch unabhängig von gesellschaftlichen Strukturen und Vorab-Festlegungen sind. Mit Goffman (1983) gesprochen, kann es nun auch um die Interaktionsordnung anstelle der Gesellschaftsordnung oder des isolierten Individuums gehen (vgl. Rawls/Duck 2020).

Zudem bietet die Datensorte der Videoaufzeichnung eine lang ersehnte Gelegenheit, dem Problem zu begegnen, dass nicht offen rassistisches Handeln nicht immer eindeutig zu identifizieren und moralisierenden Deutungspraktiken unterworfen ist. Von der einen, anklagenden Seite, die spezifische Inzidenzen des Handelns und Sprechens als rassistisch interpretiert, wird es als typisches Vorkommnis eines unterstellten unterliegenden, wiederkehrenden Musters aufgefasst. Von der anderen Seite, den angeklagten Transgressor:innen, wird es demgegenüber schnell als auf exzessiven Empfindlichkeiten basierende Überempfindlichkeit abgestritten und anderen, eher kontingenten und unsystematischen Faktoren, Intentionen, Bedeutungen oder Umständen attribuiert. Somit liegt rassistisches Handeln und Sprechen zu einem bedeutenden Maß im Auge der Betrachterin und ist eine Frage der Interpretation der Beteiligten an der Situation. Daher ist Rassismus schwer zu operationalisieren.

Eine ethisch nicht-triviale und legitime Möglichkeit, dieser Schwierigkeit zu begegnen, ist Sartres Maxime »Die Opfer haben immer recht« (vgl. z. B. Sartre 2005). Sie weist bereits in die Richtung einiger heutiger Orientierungen, die »strukturellen Rassismus« unterstellen. Die sequenzielle Analyse videografisch dokumentierter sozialer Situationen, die mindestens von einem Teil der Beteiligten als »rassistisch« interpretiert werden, bietet eine weitere Option der Erforschung von Rassismus, die möglicherweise besser gegen Vorwürfe des Subjektivismus geschützt ist. Denn auf dieser Datenbasis können die interaktionalen Interpretationsketten in ihrer Inkrementalität nachvollzogen werden.

Im Text werden wir diese neuen Optionen an drei Beispielen illustrieren und diskutieren: Zuerst an einem Fall offenen Rassismus und dessen sowohl sequenzieller interaktionaler »Behandlung« durch die Anwesenden als auch seiner eigendynamischen Weiterentwicklung. Im individuellen Ausdruck umfasst das rassistische Handeln hier die Äußerung expliziter Invektive, aber auch die Kommunikation von Ekel-Gefühlen. Das zweite Beispiel betrifft ebenfalls offenen Rassismus, der gewaltförmig – zuerst sprachlich, dann gestisch und schließlich physisch – ausgedrückt wird. Das Besondere an diesem Beispiel liegt darin, dass

die unterschiedlichen Rassismus-Modalitäten sequenziell eingebettet sind und in einem Kontinuum verächtlich-erniedrigenden Handelns geschehen. Unser drittes Beispiel zeigt einen Fall nicht offenen Rassismus, der von Polizisten bei einem Einsatz in der Wohnung einer syrischen Familie zum Ausdruck kommt. Kennzeichnend für ihn ist interaktional gesehen die Verweigerung von normalerweise unterstellten Interaktionsrechten, während semantisch gesehen stark zwischen Gruppenzugehörigkeiten unterschieden wird, den Gegenübern existenzielle Daseins-Berechtigungen abgesprochen werden und Misstrauen, Verachtung und Abscheu ausgedrückt werden.

2. Social-Media-Daten in der Rassismusforschung

Mit Blick auf die Frage nach der Nutzbarkeit von videografischen Social-Media-Daten in der sozialwissenschaftlichen Forschung diskutieren wir in unserem Text gleich mehrere Fragen: Eine erste Frage bezieht sich darauf, ob die Videoaufzeichnungen aus den sozialen Medien ohne eigenes erworbenes ethnografisches Zusatz- oder Kontextwissen überhaupt sinnvoll verwendet werden können. Unsere Analysen zeigen, dass sich gerade im Fall öffentlicher Daten aus sozialen Medien die Gelegenheit ergibt, Zusatzwissen aus Forendiskussionen und juristische Prozesse einzubeziehen (vgl. unseren Fall 2). Hier handelt es sich mithin um bereits vorab vom Individuum der Ethnografin bzw. des Ethnografen und deren/dessen Subjektivität entkoppeltes Wissen, das von Anfang an multiperspektivische Analysen ermöglicht. Ferner können die Forschenden Anwält:innen oder gar die Opfer und Täter:innen kontaktieren (vgl. unseren Fall 3), um deren Perspektiven einzuholen. In vielen Fällen haben die Filmenden großes Interesse an der Bekanntmachung und Diskussion ihres Falls in der Öffentlichkeit, haben sie doch häufig selbst die Aufzeichnungen in den sozialen Medien geteilt (vgl. unseren Fall 1).

Eine zweite Frage bezieht sich darauf, ob die isolierte Analyse von Einzeldaten, die möglicherweise zudem vorselektiert und geschnitten im Netz geteilt wurden, überhaupt an sich wissenschaftlich verwendet werden können. Unsere Beispiele zeigen, dass sie in der Tat vielfältige Forschungsfragen exemplarisch beantworten können. Mit ihrer Hilfe können basale Sozialitätsformen rassistischer Interaktion identifiziert werden, selbst wenn sie zunächst relativ kontextfrei erscheinen. Gerade sequenzielle, praktische, emotionale und interaktionsnormative Dimensionen können beim Erfülltsein grundlegender Qualitätsbedingungen der Daten hervorragend untersucht werden.

Ein dritter Vorteil vorab öffentlich verfügbarer Daten in sozialen Medien liegt zudem darin, dass dies sowohl die kritische oder komplementäre Sekundärnutzung der wissenschaftlichen Gemeinschaft als auch die Diskussion oder sonstige Nutzung der Analysen durch die weitere mediale Öffentlichkeit (im Sinne einer *public sociology*) oder gar des Rechtssystems vereinfacht.

Fall 1: Straßenbahn Bremen

Das erste Datenstück ist ein Video, das vom Urheber selbst auf Social Media veröffentlicht wurde (Bremennext 2023). Es dokumentiert nicht nur die Herstellung und den Vollzug einer sozialen Situation offenen verbalen Rassismus, sondern auch die sequenzielle, interaktionale Abwendung durch das Opfer und weitere Beteiligte an der Situation.

Student F ist wie jeden Morgen mit der Straßenbahn auf dem Weg zur Universität. Dann steigt Herr E in einem blauen Trainingsanzug zu und setzt sich versetzt zu ihm gegenüber in einen Vierersitz. Nach der Schilderung von F beklagt sich E über Zuwanderung nach Deutschland und benutzt u. a. das Wort *afrikanisch*. F geht darauf ein und weist ihn darauf hin, dass dieses Wort problematisch ist, da es viele verschiedene afrikanische Sprachen gibt. An dieser Stelle äußert E zum ersten Mal das N-Wort, was F veranlasst, die Situation mit seinem Smartphone zu dokumentieren. Die Situation nimmt einen eskalatorischen Verlauf, da E den Studenten F weiter explizit mit verschiedenen Varianten des N-Wortes benennt. Daraufhin wird die Polizei gerufen und die Straßenbahn unterbricht die weitere Fahrt. F richtet sein Handy dabei nicht direkt auf E, sondern lässt es in seinem Schoß liegen. Dadurch ist die Audioaufnahme zwar vollständig zu hören, aber nur Ausschnitte der Mimik und Gestik zu erkennen. F reagiert aktiv auf die rassistischen Beschuldigungen und Beleidigungen und sucht das Gespräch mit E. Dabei äußern sich auch Mitreisende in der Straßenbahn und kritisieren ebenfalls E's Aussagen. Erst mit dem Eintreffen der Polizei wird das Video gestoppt. Im Folgenden werden einige aussagekräftige Ausschnitte aus der Aufzeichnung herausgegriffen und analysiert, um empirisch nachzuzeichnen, wie sowohl die Herstellung als auch die interaktive Abwendung eines rassistischen Ereignisses vollzogen werden.

1[1]	F	**aber sie dürfen zu mir *neger*[2] sagen.**
2	E	**ja:a natürlich darf ich das (.) jeden [tag.**
3	F	**[ja? das ist=das ist in ordnung [für sie?**

1 Die Transkription aller Datenstücke erfolgte nach GAT2 (Selting et al. 2009) unter Hinzufügung glossierender multimodaler Beschreibungen sowie von Standbildern. Wir danken Nino Karalashvili und Baha Ocak für deren Erstellung.

2 Der Begriff ebenso wie die anschließende, weiter verschärfte Form sind explizite rassistische Beschimpfungen und koloniale Fremdbezeichnungen, die Schwarze Menschen und People of Color herabwürdigen und als Abweichung von der »weißen Norm« darstellen. Die Abbildung der Begriffe hier soll ausdrücklich nicht der Reproduktion rassistischer Sprache dienen. Im Gegenteil: Die Begriffe werden im Transkript ausgeschrieben, um den Kontext der Äußerung unverändert zu dokumentieren, die genaue sprachliche Situation analysierbar zu machen und die unverblümte Sprachgewalt, die durch sie ausgeübt wird, zu dokumentieren und nachvollziehbar zu machen.

4 P1 [nein

5 E [ich DARF neger sagen. ich darf sogar nigger sagen. (.) alter nigger.

6 P2 halt die fresse=FRESSE du scheiss rassist ja?

7 F also ich glaube ich bin nicht der einzige der das gehört hat [oder?

8 E [ABER ich darfs doch sa:gen? ich muss ihn ja nicht anfassen.

9 F aber sie haben doch [direkt zu mir gesagt ((unverständlich))].

10 P3 [NEE? sie dürfen das auch nich SAGEN

11 P4 [sie sollen das auch doch nicht sagen sie sitzen [((unverständlich))

12 E [naTÜRlich darf ich das sagen (.) [sagen darf ich alles

13 P3 [wissen sie wissen sie was das bedeutet?

14 E ich lebe in einem frei:en land

Zu Beginn der Videoaufzeichnung stellt F die rassifizierende Benennung durch E infrage (Z. 1–3). E beruft sich auf die freie Meinungsäußerung (Z. 2–5) und wiederholt und eskaliert seine rassistische Äußerung. F bleibt ruhig, greift E nicht an und bietet ihm dadurch keine Angriffsfläche. Er stellt ihm eine rhetorische Rückfrage, ob E diese Bezeichnung als korrekt empfinde (Z. 3). Nun greift der erste Mitreisende in die Interaktion ein, indem er die rhetorische Frage lautstark mit »nein« (Z. 4) beantwortet. E's Eskalation durch Verschärfung der Wortwahl reagiert auch auf diesen Zwischenruf. Daraufhin mischt sich eine weitere Mitreisende ein und ruft, dass E seine »Fresse« halten solle und ein »Scheiß Rassist« sei (Z. 6). An dieser Stelle sind bereits wichtige Prozessdynamiken der Interaktion erkennbar. F, der direkt angesprochen wird und von der Beleidigung persönlich betroffen ist, reagiert weder aggressiv noch lässt er sich von Fragen einschüchtern. Er stellt rhetorische Rückfragen und sucht das Gespräch mit E. Gleichzeitig kritisieren einige Mitreisende E direkt und machen ihm öffentlich klar, dass seine Äußerungen nicht hinnehmbar sind. Damit durchbrechen sie die *civil inattention* (Goffman 1963), mit der normalerweise in öffentlichen Räumen die große Nähe zu anderen Personen neutralisiert wird (Blick aus dem Fenster, Zeitung lesen etc.). In der hier vorgestellten Situation wird demgegenüber eine Art *civil attention* praktiziert, indem einige Mitreisende sich in die Interaktion einschalten und intervenieren.

Dies gibt F die Möglichkeit, zu betonen, dass weitere Mitreisende ebenfalls die rassistischen Einlassungen E's gehört hätten und kritische Zeug:innen der Situation wurden (Z. 7). E reagiert auf die geäußerte Kritik der Mitreisenden, die signalisieren, dass seine Beleidigungen F's nicht auf Konsens stoßen. Dass er dies möglicherweise anders erwartet hatte, zeigt sich darin, dass er sich an die anderen Mitreisenden wendet und entgegnet, dass er sich sprachlich äußern dürfe, solange er F nicht körperlich angehe (Z. 8). Erneut widersprechen ihm zwei Mitreisende,

auch rassistische Aussagen dürfe er nicht machen (Z. 10, 11, 13). Interaktiv wird den Versuchen E's, F zu rassifizieren, immer wieder die Zustimmung durch die Mitreisenden verweigert. Vermutlich davon provoziert, beruft sich E auf seine vermeintlichen Rechte, über die er als Staatsbürger verfüge, darunter seiner Meinung nach auch, alles sagen zu dürfen. Damit versucht er, die Sanktionierung der Verwendung herabwürdigender Begriffe durch F und die anderen Mitreisenden als Sprechverbot zu interpretieren (Z. 12–14). Es kommt zu einem kurzen Austausch, bei dem E behauptet, F würde durch dessen Anwesenheit seine Rechte beschneiden. E versucht dadurch, F die Anerkennung von dessen Subjektrechten zu entziehen. F hakt daher nach und fragt, welche Rechte E's er in dieser Situation gebrochen habe (Z. 19).

19	F	**was=was habe ich denn für rechte gebrochen monsieur? was habe ich ihnen denn angetan?**
20	E	**gar nichts.**
21	F	**o:h? oh? (.) [sagen sie doch (--)**
22	P5	**[boah a:lter**
23	F	**sagen sie doch. [und sie sagen ich habe sie geschlagen]**
24	E	**[ja aber ich doch auch nicht ich habe doch auch [kein recht gebrochen**
25	P6	**[DO:CH**
26	F	**[eben nicht monsieur deswegen führen wir die diskussion. sie dürfen sich nicht diskriminierend [äußern.**
27	E	**[hör auf mich anzufassen. [ge:h weg.**
28	F	**[ich?=ich fasse sie doch überhaupt nicht an**
29	E	**ich mag keine schwarzen menschen.**
30	F	**aha. da haben wir es schon wieder. (.) [was unterscheidet mich denn?**
31	E	**[ich mag euch nicht.**
32	F	**was unterscheidet mich denn von ihnen?**
33	E	**mann lass mich in ruhe. ich komme aus einer kleinstadt. ich kenn sowas nicht in bremen**
		((steht auf und läuft nach hinten in die bahn))
34	F	**kleinstadt.**
35	P7	**bitte verlassen sie die bahn.**
36	F	**ey unfassbar. monsieur ich hätte gesagt wie gesagt ich steig mit ihnen aus**
		((steht ebenfalls auf und läuft E hinterher))
37	P8	**das ist nicht ihre bahn.**
38	F	**das ist nicht ihre bahn. monsieur steigen sie mit mir aus die leute haben es nicht verdient hier länger zu warten**

E erwidert schließlich, dass F nichts getan habe (Z. 20). Die Interaktion hat sich im fortlaufenden Prozess durch die intervenierenden Aussagen der Fahrgäste sowie F's Nachfragen so weit dynamisiert, dass E F dessen Subjektrechte wieder zuspricht. Interaktiv in Bedrängnis geraten, deeskaliert er mit dem Versuch, beide, E und F, zu entlasten, indem er betont, dass auch er keine Rechte gebrochen habe (Z. 24). Der Versuch scheitert, da ein Fahrgast laut ein »Doch« äußert und damit die Argumentation umkehrt – nicht die Meinungsfreiheit E's würde eingeschränkt, sondern die Diskriminierungsfreiheit F's (Z. 25). F ergreift diese Gelegenheit und expliziert, dass E seine Rechte gebrochen habe, indem er ihn diskriminiert habe (Z. 26). Daraufhin eskaliert E erneut, indem er von ihm verlangt, ihn nicht anzufassen und wegzugehen (Z. 27). Da keine reale Berührung erfolgt ist, kann dies nur ein Versuch der Dramatisierung der Situation und der Selbststilisierung als Opfer oder aber metaphorisch gemeint sein: nicht als räumliche oder körperliche, sondern als interaktive und argumentative Bedrängnis.

Öffentlich zu einer Erklärung bzw. Entschuldigung genötigt, bringt E erneut eine rassistische Begründung hervor, indem er äußert, dass er keine schwarzen Menschen mag. Sie ist insofern zwar weniger drastisch, als er nicht mehr das N-Wort benutzt (Z. 29), umfasst aber die rassifizierende Kommunikation von Gefühlen der Abwehr und des Ekels. F lässt diese essenzialisierende und rassifizierende Begründung erneut nicht gelten und fragt, worin denn der Unterschied zwischen beiden bestünde (Z. 30, 32). Offensichtlich verunsichert ergreift E die Flucht: Er bittet F, ihn in Ruhe zu lassen, da er aus einer Kleinstadt komme und eine solche Situation wie in Bremen nicht kenne (Z. 33). F lässt auch dies nicht zu, indem er E hinterherläuft, während dieser von weiteren Mitreisenden aufgefordert wird, die Bahn zu verlassen (Z. 35, 37). Somit werden E nicht nur verbal, sondern auch räumlich die Teilhaberechte an der Interaktion sowie an der weiteren Fahrt entzogen. Die Mitreisenden lassen damit gemeinschaftlich eine potenzielle weitere rassifizierende Praxis gegenüber F nicht zu. Auf dem Bahngleis bleibt F in der Nähe von E, da er warten möchte, bis die Polizei eintrifft. Auch gibt F der Straßenbahn zu verstehen, dass diese ihre Fahrt fortsetzen könne. Die Straßenbahn bleibt jedoch stehen und E betritt sie kurz darauf erneut.

46 E ((steigt wieder in die Straßenbahn und geht zurück zu seinem Platz))

47 F **sorry ich muss das machen, ne? ansonsten kommen leute immer glimpflich davon aber das wollen wir nicht. (.) das ganze muss konsequenzen mit sich bringen. (-) ansonsten wird sich dieses land nie ändern**
((läuft ihm hinter))

48 E ((setzt sich auf seinen vorherigen Platz))

49 F ((setzt sich auf seinen vorherigen Platz))

50 E **fass MICH NICHT AN (---) ich hab angst vor den niggern**

51 F **sie haben angst vor mir?**

52 P9 **wir haben angst vor i:hnen?**

53 F **richtig. (.) jeder hat angst vor ihnen.**

54 P9 **wir haben alle angst vor ihnen. [sie laufen hier aggressiv rum**

55 E **[kannst du mir den nigger mal vom hals nehmen?**
((schaut und zeigt auf einen Fahrgast vor ihm))

56 P10 **[KÖNNT IHR JETZT ENDLICH MAL RUHIG SEIN? DAS IST JA WOHL NICHT MEHR NORMA:L**

57 E [((schaut in Richtung P10))

58 P10 **ICH SCHMEISS DICH GLEICH PERSÖNLICH RAUS=DU VOLLIDIOT! ES REI:CHT**

Bei Wiedereintritt entschuldigt sich F bei den Mitreisenden dafür, dass sie erneut die Bahn betreten, aber er erklärt auch, dass das Verhalten von E nicht ohne Konsequenzen bleiben dürfe (Z. 47). Im Unterschied zur ersten Situation in der Straßenbahn, in der E von den Mitreisenden in vielfältiger Weise widersprochen wurde, reagiert nun niemand. Nachdem sie zu ihren Plätzen zurückgekehrt sind, fordert E F erneut auf, ihn nicht anzufassen und beschimpft ihn abermals rassistisch als N******, vor denen er Angst habe (Z. 50). Wiederum rhetorisch fragt F E, ob er auch Angst vor ihm habe und stellt ihn abermals zur Rede (Z. 51). Ein weiteres Mal schaltet sich eine Mitfahrerin in das Geschehen ein und widerspricht E, indem sie klarstellt, dass die Mitreisenden vielmehr Angst vor E wegen dessen Aggressivität hätten (Z. 52, 54). F greift diese Umkehrung der Argumentation E's auf und wiederholt, dass alle Angst vor diesem hätten (Z. 53). Erneut wird die rassistische Bezeichnung nicht nur durch F, sondern auch eine Mitreisende zurückgewiesen. Dem ungeachtet sucht E Blickkontakt zu einer/einem anderen Mitreisenden vor ihm und bittet diesen mit der erneuten rassistischen Typen-Bezeichnung (N*****) sowie einer dramatisierenden Körpermetapher, ihm F »vom Hals« zu schaffen (Z. 55). Auch hier zeigt sich erneut im rassistischen Handeln die Kommunikation von Gefühlen des Ekels und der Abwehr. Auch dieser Versuch scheitert, und lautstark schreit ein Fahrgast von links, dass die Dyade zur Ruhe kommen solle. Er greift also zunächst nicht nur E an, sondern formuliert eine kollektive Schuldzuweisung (Z. 56). Erst nachdem E ihn anblickt, droht er, ihn aus der Bahn zu werfen (Z. 57, 58).

Wie Medien berichten, habe E selbst im Beisein der Polizei seine rassistischen Beleidigungen gegen F wiederholt. Zudem habe er auch die 58-jährige Straßenbahnfahrerin sexistisch beleidigt. Die Polizei erhob Anzeige wegen Volksverhetzung und Beleidigung und erteilte einen Platzverweis. Offenbar ist es nicht das

erste Mal, dass E in dieser Form ausfallend wurde. Unter dem Video auf Instagram sammelten sich schnell zahlreiche, meist F Solidarität und Unterstützung kommunizierende Kommentare.

Auch die anderen in der Straßenbahn Anwesenden haben sich nach medialer Darstellung vorbildlich verhalten. Es sei in der geteilten Videoaufzeichnung zu sehen und zu hören gewesen, dass die anderen Mitreisenden Zivilcourage bewiesen hätten, sich eingemischt hätten und eingeschritten seien. Dabei wird explizit auf einige der Äußerungen der P1–P10 in den Zeilen 4, 6, 10, 11, 13, 22, 25, 35, 37, 52, 54, 56 und 58 verwiesen. In der Tat haben die Mitreisenden unterstützende Einwürfe gerufen, die F dann reformulieren konnte, tatsächlich körperlich eingeschritten sind sie jedoch nicht. Wie wir an Zeile 56 gesehen haben, betraf die Schuldzuweisung zudem nicht immer eindeutig nur E.

Das Beispiel zeigt dennoch, dass jeder Versuch seitens E, F im Konsens mit den anderen Anwesenden zu rassifizieren und ihm seinen Status als Subjekt abzusprechen, in der interaktiven Auseinandersetzung scheitert. F's Deutung (Z. 1), dass E sich ihm gegenüber rassistisch verhalte, wird von dritten Personen ratifiziert und expliziert (Z. 4; Z. 6), was F im Interaktionsverlauf immer wieder aufgreift. Die Abwehr der rassistischen Aggression erfolgt sowohl durch (rhetorische) Rückfragen von F als auch durch direkte, sprachliche Einwände und Kritik der Fahrgäste in der Bahn. Die Herstellung von rassistischen Situationen kann somit in deren Abwendung durch interaktiven Widerspruch und argumentative Umkehr durch eine »Interpretationsgemeinschaft« prozessiert werden, auch wenn der Aggressor dies nicht validiert. Anhand der videografischen Daten lässt sich rekonstruieren, wie die sequenzielle Abfolge von Redezügen unterschiedlicher Beteiligter in öffentlichen Räumen den Gesamtverlauf maßgeblich beeinflusst und Rassifizierung nicht allein zwischen den stabilen Rollen von Täter:in und Opfer subjektiv verhandelt wird. Durch die interaktive Beteiligung Dritter und der Reformulierung von deren Beiträgen durch das Opfer konnten rassifizierende Kommunikationen entgegen der Annahme der Genealogie und Vorab-Strukturiertheit einen eigendynamischen, interaktionalen Vollzug entwickeln und damit auch stets revidiert werden.

Fall 2: Straßenbahn in Erfurt

Das zweite Video, das wir analysieren, wurde in einer Tram in Erfurt von einer/einem unbekannten Mitreisenden aufgenommen. Es zeigt ebenfalls offenen Rassismus, der gewaltförmig – zuerst sprachlich, dann gestisch und schließlich physisch – ausgedrückt wird. Das Video der Attacke wurde auch in den sozialen Medien veröffentlicht und hat bundesweit eine Rassismus-Debatte ausgelöst (Nimz 2021). Auch wurde das Video als Hauptbeweismittel für einen Gerichtsprozess verwendet, in dem darüber verhandelt wurde, ob die Tat des Angreifers rassistisch einzuordnen ist oder nicht (MDR Thüringen 2023). Im Video sind

zwei Personen sichtbar: P, die mit einem Fahrrad in der Tram steht und deren Körper und Gesicht erkennbar sind (Abb. 1, links), und K, die im Video nur von hinten erkennbar ist und auf einem Sitzplatz in der Tram sitzt (Abb. 1, rechts). K trägt einen Kapuzenpulli und eine Jacke, sodass weder Gesicht noch Aussehen für den Betrachter erkennbar sind.

Abb. 1

Z. 01: willst du noch

Abb. 2

Z. 03: du kommst

Abb. 3

Z. 06: blicken sich

01	P	**willst du noch eine? (.) du drecksvieh. (.) HEH? (.) fang an zu brennen du?=du fotze junge (.)**
		((steht im Eingangsbereich der Tram, hält sich an seinem Fahrrad fest und hat Blickkontakt zur sitzenden Person K))
02	K	((sitzt und hält den Kopf gesenkt nach unten Richtung Boden oder Smartphone))
03	P	**[du kommst <<f>HIER IN MEIN LAND JUNGE?>**
04		[((schiebt sein Fahrrad ruckartig in Richtung K))
05	K	[((hebt den Kopf in Richtung P))
06	K/P	((K und P blicken sich gegenseitig an))

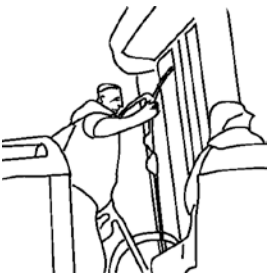
P beginnt im ersten Redezug, der aufgezeichnet wurde, K mit einer entmenschlichenden Bezeichnung als »Drecksvieh« und mit einer effeminierenden Beleidigung als »F****« zu beleidigen (Z. 01, Abb. 1 und 2). Des Weiteren schüchtert er K ein, indem er ihm wünscht, er solle brennen. In gewöhnlichen Alltagsinteraktionen wechseln sich Paarsequenzen ab, indem bspw. eine Begrüßung einen Gegengruß oder eine Frage eine Antwort auslöst: ein zweiter Redezug ist konditionell relevant und sozial verpflichtend (Meyer 2024, S. 12). Betrachtet man die Anfangssequenz im Video, ist bereits hier ein Ungleichgewicht zu erkennen, indem zwar P (Z. 02) Blickkontakt mit K aufnimmt und seine Beleidigungen äußert, K jedoch seinen Kopf gesenkt nach unten hält und keine Antwort ausspricht (Z. 03). Die Unterlassung des zweiten Zuges ist einerseits eine Folge der Bedrohungssituation, andererseits aber auch eine soziale Handlung: K lehnt die Konversation mit P durch die Verweigerung eines zweiten Zuges ab. Er widersetzt sich also nicht den

Beleidigungen P's und greift diesen auch nicht persönlich mit einer Gegenbeleidigung an, sondern verharrt in seiner Position und hält den Kopf nach unten geneigt (Z. 02, Abb. 1). Erst als P sein Fahrrad ruckartig in Richtung K schiebt (Z. 04) und ihm vorwirft, in sein Land zu kommen (Z. 03), hebt K seinen Kopf in Richtung P (Z. 05, Abb. 3). Verbal macht P eine vermeintliche Herkunft K's relevant und stellt seine Zugehörigkeits- und Präsenzrechte infrage (Z. 04). K's Kopfheben in diesem Moment ist für P zwar als Signal der Gesprächsbereitschaft interpretierbar, kann aber auch aus körperlicher Vorsicht aufgrund des ruckartigen Schiebens des Fahrrads seitens P erfolgt sein. P macht K *accountable* (interpretierbar), dass er die Kontrolle über den direkten Raum von K (dem Raum vor dem Sitzplatz) beansprucht und seine volle Aufmerksamkeit einfordert. Indem K nun den Kopf hebt, ermöglicht er P, in Interaktion mit ihm zu treten (Z. 6, Abb. 3).

Abb. 4

Z. 07: he du fotze

Abb. 5



Z. 07: fängt sich

Abb. 6

Z. 10: senkt den Kopf

07	P	**he [du fotze**
		[((verliert das Gleichgewicht und kippt mit dem Oberkörper nach hinten links, fängt sich mit dem rechten Arm, zieht sich wieder nach vorne rechts in Richtung K und richtet sich mit angespanntem rechtem Arm vor K auf))
08	K	[((blickt auf P))
09	P	**[HEH?**
		[((hält sich direkt vor K an der Stange fest, der rechte Oberarm ist sichtbar angespannt))
10	K	((senkt den Kopf))
11	P	**(--) du kleines drecksvieh (-) verpiss dich einfach dahin wo du herkommst du du=verfickte fotze.**

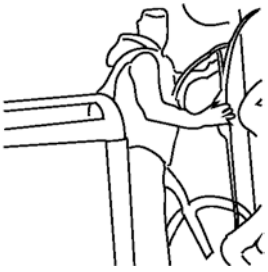
Im nächsten Redezug beschimpft P K erneut mit »F****« (Z. 07), verliert dabei aber sein Gleichgewicht, sodass er sich an der Haltestange der Tram fangen muss (Z. 07, Abb. 4, 5). Danach richtet er sich mit angespanntem Oberarm vor K wieder auf (Abb. 6). Während er die Kontrolle verliert, ist K ihm noch mit

seinem Blick zugewandt (Abb. 5, Z. 08). Es scheint für einen kurzen Augenblick, dass P durch den Gleichgewichtsverlust die Kontrolle über die Gesamtsituation verliert. Sein Gleichgewichtsverlust hat jedoch die gegenteilige Handlungskonsequenz für die darauffolgende Interaktion: P korrigiert seine Haltung, stellt sich nun noch dichter vor K und versperrt durch seinen angespannten Oberarm jegliche Auswegs- bzw. Fluchtoption für K (Abb. 6). Durch das Fahrrad, seinen angespannten Arm sowie seine direkte Körperstellung vor K besteht kaum eine Möglichkeit, dass K sich der Situation noch entziehen könnte, z. B. durch das Wechseln auf einen anderen Platz (Abb. 6). Die ausweglose Lage macht K interaktiv *accountable*, indem er wieder den Kopf nach unten senkt und damit anzeigt, dass er die weitere Interaktion mit P abbricht und sich zurückziehen möchte (Z. 10, Abb. 6). P lässt dies jedoch nicht zu, sondern wiederholt die Beleidigung K's als »Drecksvieh« und fordert ihn auf, sich dahin ›zu verpissen, wo er hergekommen‹ sei (Z. 11). Erneut macht P eine unterstellte Herkunft relevant und spricht K seine Existenzrechte ab. Diese Relevanzsetzung kann allenfalls anhand äußerlicher Merkmale K's erfolgt sein, die nur für P sichtbar sind und die er nun interaktiv relevant macht.

Abb. 7

Z.13: spuckt

Abb. 8



Z.20: dreht den Kopf

Abb. 9

Z. 26: verstehst DU

12	K	((hebt die linke Hand schützend über seinen Kopf))
13	P	((lehnt sich an der Haltestange nach vorne in Richtung K und spuckt von oben herab auf ihn))
14	P	((lehnt den Oberkörper wieder nach hinten, sodass er wieder aufrecht vor K steht))
15	P	**friss bist scheisse=du drecksvieh (-)**
16	P	**ruf nach deiner [mutter oder nach deinem papa? junge?**
17	Z	[((läuft in Richtung P))
18	P	**du FO:tze [DU**
19	Z	**[(unverständlich)**
20	P	((dreht den Kopf zu Z nach hinten links))
21	P	**[<<p>ich mach doch nichts**

22 K [((nimmt den Arm über seinem Kopf wieder herunter))

23 P ((dreht den Kopf wieder nach vorne rechts in Richtung K))

24 P **frag mich noch ei:nmal was ich von dir will junge?**

25 P **mach einfach deinen kopf runter wenn ich das nächste mal vor dir stehe=du drecksvieh du (-)**

26 P **verstehst DU DAS?**

27 P ((tritt mehrere Male mit seinem rechten Fuß in K's Gesicht))

28 P **JAH mach nächstes mal (unverständlich)?**

Nach entwürdigenden und entmenschlichenden Bezeichnungen wie »Drecksvieh« und der alterisierenden Aufforderung »verpiss dich dahin wo du herkommst«, spitzt sich P's menschenverachtendes Verhalten gegenüber K weiter zu, indem er K von oben herab bespuckt (Z. 13, Abb. 7). K reagiert darauf mit einer Schutzgeste, indem er seine linke Hand über seinen Kopf hält (Abb. 8). Ersichtlich wird, wie P's gewaltförmiges Handeln zuerst durch sprachliche, dann gestische Invektive und schließlich durch physische Gewalt ausgedrückt wird. Die sequenzielle Abfolge der Herabwürdigung, inklusive der Rassifizierung, seitens P wird weiter fortgesetzt, indem P K abermals entmenschlichend als »Drecksvieh« und effeminierend als »F****« (Z. 15, 18) bezeichnet. Hinzu kommt, dass er ihm verkindlichend Angst und Schwäche zuschreibt und ihm unterstellt, dass er nach seiner Mama und nach seinem Papa schreien würde (Z. 16).

Im Gang bewegt sich jetzt von der linken Seite die Mitreisende Z in Richtung P (Abb. 8). Im Video ist erkennbar, dass sie P anspricht, was auf der Aufzeichnung nicht zu verstehen ist (Z. 19). Wie Nimz (2021) berichtet, wollte Z nicht etwa zum Schutz des Opfers einschreiten, sondern den Schläger vielmehr mit den Worten »Denk an die Kameras« vor der Dokumentation seiner Taten warnen. P dreht sich kurz zu ihr um und erwidert, dass er doch gar nichts machen würde (Z. 20, 21, Abb. 8). Währenddessen nimmt K seine Hand herunter (Z. 22). Die Situation wird also nicht nur von der filmenden Person, sondern auch von anderen Personen in der Tram beobachtet. Jedoch führt P sein verächtlich-erniedrigendes Verhalten gegenüber K nicht nur fort, sondern eskaliert es weiter. Er diktiert ihm, dass er seinen Kopf zu verneigen habe, wenn er das nächste Mal zu ihm spreche (Z. 25) und beschimpft ihn erneut als »Drecksvieh«. Schließlich tritt er mehrmals mit seinem rechten Fuß in Kopfhöhe auf K ein (Z. 27, 28, Abb. 9). Danach bricht das Video ab.

Im an den Vorfall anschließenden Prozess um den rassistischen Angriff wurde P letztlich wegen gefährlicher Körperverletzung, Nötigung und Sachbeschädigung zunächst zu vier Jahren und neun Monaten Haftstrafe verurteilt. Das Urteil wurde von der Beratungsstelle für Betroffene rechter, rassistischer und antisemitischer Gewalt in Thüringen insgesamt sehr positiv beurteilt, da zwischen Tat und Urteilsspruch eine geringe Zeitspanne lag. Auch die schnelle Inhaftierung des mehrfach vorbestraften Täters und dessen Verbleib in der

U-Haft wurden positiv hervorgehoben. Dass es eine Videoaufnahme gab, die sich in den sozialen Medien massenhaft geteilt hatte, wodurch keine Zweifel an der Tatbegehung und Schuld des Täters bestanden, trug, so die Bewertung, dazu bei. Schon der Ermittlungsrichter sah den Angriff als »ganz offensichtlich menschenfeindlich«. Bei der Urteilsverkündung sagte der Richter dann, die Tat von P sei eindeutig als rassistisch einzuordnen, wie die Zeug:innenaussagen, aber maßgeblich auch das Video, das durch die sozialen Medien kursierte, belege. Der Bundesgerichtshof (BGH) hob das Urteil später in Teilen auf und reduzierte es auf vier Jahre Freiheitsstrafe. Das 17-jährige Opfer der Tat erlitt schwere Prellungen und ein Schädel-Hirn-Trauma und musste nach dem Angriff im Krankenhaus behandelt werden. Auch von psychischen Folgen wird in den Medien berichtet.

Das schockierende Beispiel ermöglicht es, den inkrementellen Verlauf von rassifizierendem Handeln zu rekonstruieren. Es zeigt, wie rassistische Praktiken durch audiovisuelle Daten analytisch darstellbar gemacht werden können. Dabei ist deutlich erkennbar, wie die unterschiedlichen invektiven und rassifizierenden Praktiken miteinander auch sequenziell verknüpft sind: Verbale Gewalt, gewaltförmige Körperpraktiken und Gestiken (Fahrradschieben, körperliches Aufbauen) sowie physische Angriffe (Spucken, Treten) sind in einem Kontinuum der Erniedrigung und Annihilierung *alters* sequenziell organisiert. Auch die Objektifizierung von K durch P und die Aberkennung eines etwa durch Dialogizität artikulierbaren Subjektstatus kann eindeutig rekonstruiert werden. Interaktional wird das Ereignis allerdings in der Situation selbst zunächst nicht explizit als rassistisch behandelt, da eine »third turn validation« (Schegloff 1992) zunächst ausbleibt, wenngleich Z die Gesetzwidrigkeit von P's Handeln interaktional einspeist. Eine Ratifizierung und Validierung als rassistisches Handeln erfolgt jedoch am Ende durch die Rechtsprechung und Verurteilung P's.

Fall 3: Polizeieinsatz in Berlin

Nach der Analyse verbaler Invektive und ihrer interaktionalen Umkehr im ersten Beispiel und der Kenntnisnahme der nur graduellen Kontinuität von verbalen Invektiven, körperlich-gestischer Transgression und physischer Gewalt in der interaktionalen Sequenzialität bzw. Inkrementalität im zweiten Beispiel soll nun im dritten Beispiel abschließend die Aberkenntnis von Interaktionsrechten und gängigen Anrechten auf die Zuerkenntnis von Intersubjektivitätsfähigkeiten analytisch im Fokus stehen. Im Kontrast zu den vorherigen Situationen handelt es sich nun um ein Ereignis, um das es in der Nachfolge Diskussionen um seine Deutung gab. Ein weiterer Unterschied besteht darin, dass es ein Vorfall im Kontext institutionellen Handelns in privaten Räumen war: es handelt sich um einen

Polizeieinsatz, bei dem am frühen Morgen des 9. Septembers 2022 zwei männliche Beamte der Berliner Polizei die Wohnung der Familie R aufsuchten. Anlass für den Polizeieinsatz war, dass gegen den 30-jährigen Herrn R (HR), dessen Familie R aus Syrien stammt, aber in Berlin lebt, ein vollstreckbarer Haftbefehl bestand, da er eine Geldstrafe in Höhe von 750 Euro für dreimaliges Fahren ohne Fahrschein nicht gezahlt hatte. Das Fahren ohne Fahrschein wird in Deutschland nicht als Ordnungswidrigkeit, sondern als Straftat geahndet.

Als die Beamten in Uniform, Handschuhen und mit schweren Stiefeln die Wohnung betreten, sind HR und dessen Frau FR ebenso wie ihre drei kleinen Kinder noch in leichter Nachtkleidung. Die Polizisten geben der Familie zu keinem Zeitpunkt die Möglichkeit, sich anzukleiden und dadurch als gleichwertiges Gegenüber in der Interaktion zu präsentieren. FR filmt den Einsatz der Polizisten mit ihrem Handy, zwischendurch gibt sie es auch an eines ihrer Kinder weiter. Das Video beginnt erst, als sich die beiden Beamten bereits im Flur der Wohnung befinden. Das Video zeigt, wie es zu verbalen Auseinandersetzungen zwischen den Eheleuten und den Beamten kommt, auch körperliche Handlungen der Polizisten gegen HR sind dokumentiert. Erst durch das Eintreffen von vier weiteren Polizeibeamt:innen kommt es zur Beruhigung der Situation. Der offene Geldbetrag in Höhe von 750 Euro wird von den Eheleuten R beglichen und die Beamt:innen verlassen die Wohnung der Familie.

Bruchstücke des Videos wurden durch einen Berliner Senatsabgeordneten in den sozialen Medien zirkuliert und lösten eine vehemente Diskussion über die Angemessenheit des Einsatzes und möglichen polizeilichen Rassismus aus (Fröhlich 2022; Christmann 2022). Nach dem Einsatz erhoben die Polizeibeamten Anklage gegen die Eheleute R wegen Widerstands, tätlichen Angriffs und versuchter Gefangenenbefreiung (Polizei Berlin 2022). Die Eheleute R wiederum erstatteten Anzeige gegen die Beamten wegen Körperverletzung im Amt und rassistischer Beleidigung. Dafür stellten sie dem Gericht als Beweismittel auch das komplette Video zur Verfügung, das uns vom Anwalt der Familie bereitgestellt wurde.

Im Video ist zu sehen, dass HR sich zunächst weigert, den Bußgeldbetrag zu begleichen. Als die Beamten ihn daraufhin festnehmen wollen, verweigert er sich dem Anlegen von Handschellen hinter seinem Rücken, indem er seine Hände vor seinem Körper überkreuzt. Dabei erklärt er sich bereit, das Bußgeld nun doch direkt zu begleichen, um die Festnahme abzuwenden, was ignoriert wird. Die Situation ist stark angespannt, da ein Teil der Kinder durch lautes Weinen deutlich macht, dass die Situation für sie angsteinflößend ist, denn sie beobachten, wie ihr Vater von der Polizei festgehalten und bedrängt wird. Als FR daran erinnert, dass ihr Mann das Geld hat und es den Beamten geben möchte, dreht sich Polizist B (PB) um und ermahnt sie lautstark, Abstand zu halten. Erst als HR wiederholt, dass er das Geld habe und bezahlen möchte, lassen die Beamten HR nach einigen moralisierenden Hinweisen los und fordern

ihn auf, das Geld zu holen. HR bittet die Beamten, zu warten, und begibt sich in das Schlafzimmer, in dem sich auch zwei seiner kleinen Kinder befinden. PB verweigert dies und folgt ihm. Polizist A (PA) wartet an der Türschwelle und blockiert somit den Zugang zum Schlafzimmer, während FR die Situation vom Flur aus filmt. Da die beiden Kinder im Schlafzimmer laut weinen, versucht FR PB davon abzuhalten, ebenfalls hinein zu gehen, indem sie sich mit einer Beschwichtigungsgeste ihrer Hand ihm nähert. PB herrscht sie an mit »verpiss dich«!

Während HR im Schrank nach dem Geld sucht, setzt PB ihn weiter unter Druck, indem er aufrückt und mit ausgestrecktem und emporgehobenem Zeigefinger in den Schrank zeigt. HR reagiert auf die Bedrohung mit einer Körpergeste (Abb. 1–2), indem er seine Arme und Hände nach hinten und seine Brust nach vorne streckt. Dies kann als Geduld und persönlichen Freiraum einfordernde Abwehrreaktion verstanden werden, wird von PB aber als Angriff und Widerstand gegen seine persönliche und staatliche Autorität gedeutet. Er stößt HR mit beiden Händen an der Brust nach hinten. Daraufhin betritt PA nun auch das Zimmer (Abb. 3), und HR wird von den beiden Polizisten zu Boden gebracht (Abb. 4–6). Das »zu Boden Bringen« stellt eine polizeiliche Zwangshandlung dar, bei der körperliche Gewalt angewendet wird. Dabei muss das Verhältnismäßigkeitsprinzip beachtet werden und der Gewalteinsatz situativ geeignet und angemessen sein. Durch die Handlung des »zu Boden Bringens« (als zweiter Teil einer Paarsequenz) wird HR's körperliche Geste (als deren erster Teil) von den Polizisten somit auf die Bedeutung einer akuten Gefahr und Bedrohung festgelegt, die es zu beantworten und zu unterdrücken gelte. Mit Wohlwollen hätte sie sicherlich auch anders interpretiert werden können.

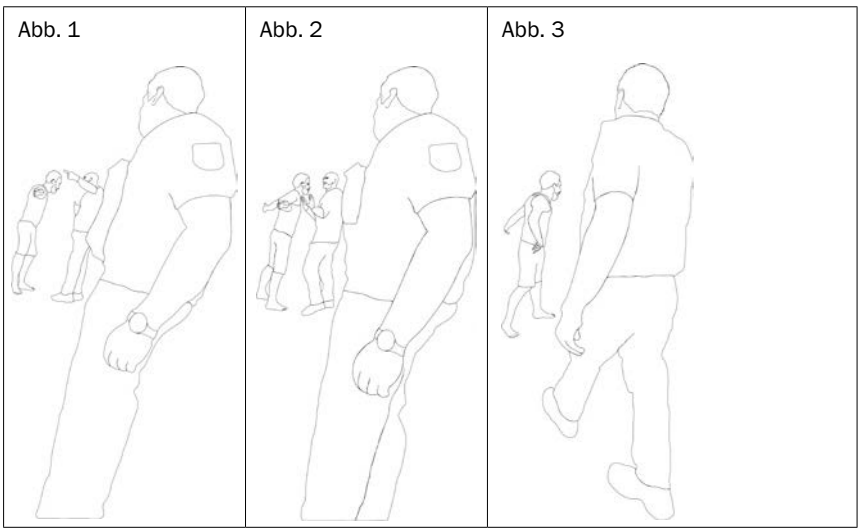


Abb. 4

Abb. 5

Abb. 6

Als FR das Schlafzimmer betritt, zeigt PB mit seinem im schwarzen Handschuh befindlichen Zeigefinger auf sie (Abb. 6) und befiehlt ihr dreimal in lautem Ton, zurückzubleiben. FR wird als permanente Bedrohung entworfen, die räumlich distanziert werden muss.

Nachdem HR Handschellen angelegt wurden, wird er erneut zu Boden gebracht. HR und FR sind empört und aufgebracht, dass die drei Kinder all dies mit ansehen müssen.

```
1   HR     warum machen das
2          (2.1)
3   PB     bin ich der verbrecher oder du (-)
4   Kind   (weint)
5   FR     <<f>habt drei kinder gesehen jetzt>
```

PB fordert sie nun auf, das Zimmer zu verlassen. Als sie erklärt, dass sie sich schließlich in ihrer Wohnung befänden, erwidert PB: »das ist mein Land und du bist hier Gast« (Z. 10).

```
6   PB     geh raus (.)
7   FR     <<ff>nicht raus> (.)
8          <<ff>das meine haus>
9   Kind   (weint)
10  PB     <<f>das ist mein land und du bist hier gast>
11  FR     <<f>wa::s>
```

FR beschuldigt daraufhin die beiden Polizisten, ihren Mann geschlagen zu haben. Dabei macht sie das »zu Boden Bringen« ihres Mannes körperlich-performativ sichtbar: Sie bewegt mehrere Male ihre rechte, geballte Faust nach oben und unten und symbolisiert so ein Schlagen (Abb. 7).

Abb. 7

Z. 01: schlagen

Abb. 8

Z. 05: fresse

Abb. 9

Z. 11: deine frau

Abb. 10

Z. 16: ich bringe

12 **FR** aber du- du- ih- (.) mu- mu- musstest schlagen (--)

13 **PA** hier hat keiner ge[schlagen=

14 **FR** [ja:

15 **PA** =wir haben ihn zu boden ge[bracht
16 **PB** [halt die fresse=
17 =fass mich nicht nochmal an (-)
18 [ich stecke dich in den knast
19 **HR** [<<acc>nicht äh sprechen mit [meiner frau> (.)
20 **PB** [ich stecke dich in den knast
21 **HR** <<ff>nicht so sprechen mit meiner frau> (.)
22 **PB** deine frau hat nicht so mit mir zu sprechen (.)
23 du bist hier in unserem land (.)
24 habt uns nach un- nach ä:h [unsern gesetzen habt ihr euch zu verhalten
25 **FR** [ja unseren anderen land aber wir- äh wir-
26 **PB** du schrei mich nicht so an und fass mich nie wieder an (.)
27 **PB** [ich bringe dich ins gefängnis=
28 **FR** [(nicht so verhalten)
29 **PB** =ich bringe DICH ins gefängnis

Ihre Bewegung kann als »Reinszenierung« und somit als verkörperte Praxis des Erinnerns verstanden werden (Bergmann 2000). Die beiden Polizisten reagieren unterschiedlich darauf. PA rechtfertigt die Handlung als amtliche Maßnahme (Z. 15). PB hingegen zeigt mit seinem Zeigefinger direkt auf die Stirn von FR (Abb. 8, 10), fordert sie in einer beleidigenden Art und Weise auf, nicht zu reden (»halt die Fresse«, Z. 16), und befiehlt ihr, ihn nicht zu berühren (Z. 17). Zudem droht er ihr rechtliche Konsequenzen an (»ich steck dich in den Knast«) (Z. 18). Der Videoaufzeichnung ist nicht eindeutig zu entnehmen, ob FR PB tatsächlich berührt hat. Angesichts des Abstands zwischen beiden erscheint es allerdings unwahrscheinlich. PB versucht hier eine Täter-Opfer Umkehr. FR's Kritik am Handeln der Polizei wird nicht nur als Angriff auf die polizeiliche Autorität interpretiert, sondern ihr Handeln wird als gesellschaftliche Gefahr umgedeutet und mit weiteren institutionell-behördlichen Maßnahmen bedroht.

Nachdem HR PB aufgefordert hat, nicht so mit seiner Frau zu sprechen, zeigt PB auf den Boden der Wohnung der Familie R und betont: »du bist hier in unserem Land« (Z. 23). Er macht HR und indirekt auch dessen Familie zu Anderen, die nicht zum Kollektiv gehören (Z. 23). Zudem knüpft er diese alterisierende Behauptung an die Forderung, sich nach »unseren Gesetzen« (Z. 24) zu richten. Das Zeigen auf den Boden stellt performativ eine Differenzmarkierung her, mit der Familie R aufgrund der ihr zugeschriebenen Herkunft die Zugehörigkeit zur deutschen Gesellschaft abgesprochen wird. Implizit wird ihnen damit auch das Grundrecht auf die Wohnung und deren Unverletzlichkeit aberkannt.

Das Beispiel verdeutlicht, dass auch in Fällen von nicht offenem Rassismus Daseinsberechtigungen abgesprochen sowie Misstrauen, Verachtung und Abscheu

ausgedrückt werden. Der inkrementelle Verlauf der Interaktion zeigt, dass das Verhalten der Polizisten durch eine unflexible, unempathische und misstrauische »Hermeneutik des Verdachts« (Ricoeur 1965, S. 40–49) charakterisiert ist. Eine solche »Hermeneutik des Verdachts« geht, wie wir an anderer Stelle ausgeführt haben (Meier zu Verl et al. 2024), stets von versteckten niederen Motiven des Gegenübers aus, kriminalisiert, alterisiert und diskriminiert diesen und ist insofern als missgünstig zu bezeichnen, als sie das Gegenüber grundsätzlich als Objekt, das sich zu unterwerfen hat, entwirft. In unserem Beispiel wird die Hermeneutik des Verdachts flankiert von einer gewaltförmigen Manipulation und Regulierung der alterisierten und objektifizierten Gegenüber, auch deren Körper im Raum betreffend, und einer fehlenden Zuerkennung von Legitimität und Moralität. Eine Hermeneutik des Verdachts, wie sie hier seitens der Polizisten angewandt wird, pocht auf die strikte Einhaltung von Regeln, die als eindeutig und unzweifelhaft entworfen werden, und darauf, dass selbst mögliches sprachbedingtes Missverstehen in Verantwortung der Gegenüber liegt und allein sie es sind, die in Vorleistung bezüglich der Herstellung von Verständigung gehen müssen.

Eine Hermeneutik des Verdachts ist damit das exakte Gegenteil einer Hermeneutik des Vertrauens (Ricoeur 1965, S. 36–40), die das Gegenüber auf eine wohlwollende, ambiguitätstolerante und großzügige Weise als kooperatives, kollaboratives und gleichberechtigtes Subjekt anerkennt. Aus sozialtheoretischer Perspektive ist der in einer solchen Hermeneutik des Vertrauens enthaltene Vertrauensvorschuss eine konstitutive Kooperationsbedingung zur Herstellung von Vergemeinschaftung. Eine Hermeneutik des Verdachts, wie sie in unseren empirischen Beispielen zum Vorschein kam, impliziert also auf sehr grundlegende Weise, das Gegenüber gar nicht erst als Subjekt anzuerkennen. Wie Heap und Roth (1973, S. 364) sagen: »intersubjectivity is contingently accomplished by members' situated practices«. Genau dies kann auch verweigert werden: Letztlich wird über eine Hermeneutik des Verdachts die Intersubjektivitätsfähigkeit *alters*, das zu *ego* prinzipiell gleichrangig ist und in Analogie steht, angezweifelt und verwehrt. Wie wir gezeigt haben, wird dies in Interaktionen auch mit institutionellen Repräsentant:innen noch immer angezweifelt und aberkannt. In Verbindung mit PB's Aussagen zur Daseinsberechtigung der Familie R in Deutschland scheint uns diese für die Interaktion durchgehend charakteristische Hermeneutik des Verdachts ein guter Grund zu sein, das Verhalten der Polizisten als rassistisch zu bezeichnen. Diskussionen in den Social Media haben sich kontrovers mit dieser Frage befasst: Während es zahlreiche Kritiken am Vorgehen der Polizei gab und zumindest PB auch suspendiert bzw. versetzt wurde, nahmen Vertreter:innen der AfD sowie Innenministerin Nancy Faeser die Polizisten in Schutz: ihrem Empfinden nach sei das Handeln der Polizisten nicht rassistisch gewesen. Urteile der Gerichte in den beiden angestrengten Verfahren stehen noch aus. Der einzige »third turn« ist somit bislang für Beispiel 3 derjenige in diesem Text. Die Videoaufzeichnung gestattet es uns, einen solchen auf fundierter Datengrundlage nachvollziehbar zu machen.

3. Fazit

In allen drei Fällen, die hier vorgestellt wurden, lässt sich anhand audiovisueller Daten erkennen, dass Rassismus in unterschiedlichen Weisen interaktiv hergestellt und prozessiert wird. Rassismus kann – wie in Fall 1 – offen artikuliert, aber in der sequenziellen Folge durch Bildung einer Gegen-Interpretationsgemeinschaft abgewendet werden. Er kann multimodal durch verbal-gestisch-körperliche Gewalt ausgedrückt und zum Gegenstand von Gerichtsverfahren werden. Ein zentrales Muster rassifizierender Interaktion kann sich auch – wie in Fall 2 und vor allem 3 – im Entzug von gängigen Intersubjektivitätsrechten ausdrücken. Wechselseitiges Vertrauen, das eine grundlegende Kooperationsbedingung für die Anerkennung des Gegenübers als gleichberechtigtes Subjekt darstellt, wird in rassifizierenden Interaktionen dem alterisierten Gegenüber aktiv entzogen. Unsere Fälle zeigen, dass bei rassistischen Vorfällen Täter:innen- und Opferzuschreibungen oftmals nicht entlang realer Täter:innen-Opfer-Achsen, sondern rassifizierender Zuschreibungen gemacht werden. Schuldigkeit wird dann aufgrund von vermeintlicher, d. h. über Sichtbarkeiten interpretierter Herkunft zugeordnet. Das wurde in Fall 1 erfolgreich abgewendet, in Fall 2 gewaltförmig ausgeübt und in Fall 3 sogar von der Polizei mit ihrem Gewaltmonopol situativ durchgesetzt. In Bezug auf die Fragestellung, ob veröffentlichte Social-Media-Daten als Datenquelle für wissenschaftliche Analysen und Erkenntnisse neue Perspektiven eröffnen, zeigen die drei vorgestellten Fälle, dass anhand dieser Datensorte insbesondere die Prozesshaftigkeit von rassistischen Ereignissen sowie deren inkrementeller Verlauf jenseits von Erhebungen zu individuellen Einstellungen und gesellschaftlichen Vorab-Strukturen sehr gut herausgearbeitet werden können. Auch bezüglich der Multimodalität von Sprache, Mimik und Gestik sowie physische Gewalt und deren Auswirkungen auf rassifizierendes Handeln können veröffentlichte Videos neue Forschungsperspektiven eröffnen. Da die Videos öffentliche Quellen sind, die von den Filmenden meist selbst ins Internet gestellt wurden, wäre auch eine unplausible Subjektivität der Ethnografin bzw. des Ethnografen am Material leicht zu demonstrieren und zu demontieren. Die/der Filmende selbst ist interessiert an einer öffentlichen Diskussion zu Rassismus, und so kann es zu einer wichtigen neuen Beziehung zwischen Wissenschaft und Öffentlichkeit kommen. Aus diesen Gründen sollte die Bedeutung dieser Datensorte für die Rassismusforschung keinesfalls unterschätzt werden. In Bezug auf mögliche Nutzungsszenarien erlauben Videodaten aus den Social Media somit die Erforschung auch antisozialer und adverser Formen von Sozialität und gestatten so eine Bearbeitung soziologischer Grundfragen nach der basalen Organisation von Sozialität. Zugleich ermöglichen sie gesellschaftsdiagnostische Forschung im Sinne einer *public sociology*. Was die Erhebung anbetrifft, so ist diese Form von Daten ungewöhnlich leicht erhältlich, insbesondere im Verhältnis zur Komplexität, die sie an Inhalten und Themen bieten kann. Entsprechende

Daten sind i. d. R. nur unter großen Schwierigkeiten selbst zu erheben und von zahlreichen Zufällen begleitet. Hinsichtlich der Analyse ist das Besondere, dass die soziologische Analyse nur eine von mehreren Analysekontexten darstellt. Sie gesellt sich zu professionellen juristischen und polizeilichen Analysen sowie zu situativen Live-Bewertungen der Anwesenden vor Ort und den öffentlichen Vernakuläranalysen in den sozialen Medien. All diese leicht zugänglichen Proto-Videoanalysen können soziologisch einbezogen werden und die eigene Analyse komplementieren.

Literatur

Bergmann, Jörg (2000): Reinszenierung in der Alltagsinteraktion. In: Streeck, Ulrich (Hrsg.): Erinnern, Agieren und Inszenieren. Enactments und szenische Darstellung im therapeutischen Prozeß. Göttingen: Vandenhoek & Ruprecht, S. 203–221.

Bremennext (2023): Video. https://www.tiktok.com/@bremennext/video/7253154928514927898 (18.09.2024).

Christmann, Karin (2022): »Das empfinde ich nicht als Rassismus«: Innenministerin Faeser verteidigt Berliner Polizisten. www.tagesspiegel.de/berlin/das-empfinde-ich-nicht-als-rassismus-innenministerin-faeser-verteidigt-berliner-polizist-8669461.html (23.09.2024).

Fröhlich, Alexander (2022): »Das ist mein Land, Du bist hier Gast«: Berliner Polizist beleidigt Familie in eigener Wohnung rassistisch. www.tagesspiegel.de/berlin/das-ist-mein-land-du-bist-hier-gast-berliner-polizist-beleidigt-familie-in-eigener-wohnung-rassistisch-8643700.html (23.09.2024).

Goffman, Erving (1963): Behavior in Public Places: Notes on the Social Organization of Gatherings. New York: Free Press.

Goffman, Erving (1983): The Interaction Order: American Sociological Association, 1982 Presidential Address. In: American Sociological Review 48, H. 1, S. 1–17.

Heap, James L./Roth, Phillip A. (1973): On Phenomenological Sociology. In: American Sociological Review 38, H. 3, S. 354–367.

MDR Thüringen (2023): Nach rassistischem Angriff: Haftstrafe für Straßenbahn-Schläger abgemildert. www.mdr.de/nachrichten/thueringen/mitte-thueringen/erfurt/angriff-strassenbahn-urteil-rassitisch-prozess-100.html (20.09.2024).

Meier zu Verl, Christian/Meyer, Christian/Grauert, Hanna (2024): Hermeneutiken des Verdachts: Institutionelle Diskriminierungspraktiken und Rassismus in öffentlichen und privaten Räumen der post-migrantischen Gesellschaft. Eine videoanalytische Untersuchung. In: Middell, Matthias (Hrsg.): Varianzen des Zusammenhalts Historisch und transregional vergleichende Perspektiven. Frankfurt am Main: Campus, S. 393–424.

Meyer, Christian (2024): Context-Sensitivity and Context-Productivity: Notions of »Practice« and »Practicality« in Ethnomethodology and Conversation Analysis. In: Frontiers in Sociology 9: 1221026, S. 1–14. https://doi.org/10.3389/fsoc.2024.1221026.

Nimz, Ulrike (2021): Rassistischer Angriff in Erfurt: »So ein feiger Mensch«. www.sueddeutsche.de/politik/rassismus-thueringen-erfurt-strassenbahn-1.5277929 (20.09.2024).

Polizei Berlin (2022): Fremdenfeindliche Beleidigung durch Polizist – EG Zentral ermittelt. www.berlin.de/polizei/polizeimeldungen/2022/pressemitteilung.1244807.php (18.09.2024).

Rawls, Anne/Duck, Waverly (2020): Tacit Racism. Chicago: The University of Chicago Press.

Ricoeur, Paul (1965): De l'interprétation. Essai sur Freud. Paris: Seuil.

Sartre, Jean-Paul (2005): Entwürfe für eine Moralphilosophie. Reinbek bei Hamburg: Rowohlt.

Schegloff, Emanuel A. (1992): Repair after Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. In: American Journal of Sociology 97, H. 5, S. 1295–1345.

Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung 10, S. 353–402.

Zwischen Interaktion und Produkt

Polizeiliche Bodycamaufnahmen als hybrider Gegenstand der Videoanalyse

Simon Egbert und Jasper Janssen

Zusammenfassung: Durch die zunehmende Ausstattung der deutschen Polizei mit Bodycams verändern sich nicht nur Interaktionen zwischen Bürger:innen und der Polizei, sondern die dabei entstehenden Videoprodukte wandeln auch die Bedingungen potenziell folgender Gerichtsverhandlungen. Als videoanalytischer Gegenstand wirft die Bodycam und ihre Aufnahmen dabei diverse Fragen und Herausforderungen auf, die wir in unserem Beitrag beleuchten. Neben Fragen zu technischen Begrenzungen (z. B. hinsichtlich der Kameraperspektive), stellen sich insbesondere Fragen zu den Bedingungen der Videoproduktion und ihre Einflüsse auf die gefilmte Interaktion. Polizeiliche Bodycams legen daher eine Kombination aus Videointeraktionsanalyse und Videoproduktanalyse nahe, die wir vorliegend als Verknüpfung aus Videointeraktionsanalyse und wissenssoziologisch-hermeneutische Fallanalyse umzusetzen vorschlagen. Die damit einhergehenden Analysepotenziale illustrieren wir exemplarisch anhand eines Bodycamvideos aus einem Münchener S-Bahnhof, in dem Bundespolizist:innen einen vermeintlich fahrscheinlosen Bürger überwältigen, der sich weigert sich auszuweisen.

Schlüsselwörter: Bodycams, Polizei, Camera Perspective Bias, Videointeraktionsanalyse, wissenssoziologisch-hermeneutische Fallanalyse

1. Einleitung[1]

Seit einigen Jahren haben sich körpernah getragene Videokameras, die sogenannten Bodycams, bei deutschen Polizeibehörden etabliert. Beginnend mit einer

1 Dieser Beitrag basiert auf dem Vortrag »Polizeiliche Bodycams als videoanalytischer Gegenstand« den der Erstautor zusammen mit Tom Berger auf dem Workshop »Qualitative Videoanalyse: Audiovisuelle Daten und ihre Kontexte« im November 2023 an der TU Berlin gehalten hat. Wir danken Tom Berger für seine Mitarbeit an dem Vortrag, der die Basis für diesen Beitrag bildet. Zudem danken wir Nele Dörl für ihre Mitarbeit in der Fertigstellung des Manuskripts sowie Mina Godarzani-Bakhtiari und René Tuma, René Wilke, Hubert Knoblauch und den Teilnehmer:innen des genannten Workshops für wertvolle Kommentare und Überarbeitungshinweise.

Pilotstudie in Frankfurt am Main 2013, haben sich die mobilen Kameras in mittlerweile allen deutschen Landespolizeien sowie bei der Bundespolizei verbreitet (Berger/Egbert/Janssen i. E.). Mit der Einführung der Kameras wird von der Polizei die Hoffnung verbunden, dass Gewalttaten gegen Polizist:innen durch Abschreckung verhindert werden können und gleichzeitig beweissicheres Material für etwaige Gerichtsprozesse, insbesondere mit Blick auf Widerstandshandlungen und Angriffe auf Vollstreckungsbeamt:innen, produziert werden kann (Mannich 2024). Aber auch die Herstellung einer »Waffengleichheit« (Teggatz zit. n. Brenner 2021, S. 21) gegenüber privaten Smartphoneaufnahmen wird (inoffiziell) angestrebt.

Die akademische Diskussion über Bodycams in Deutschland steht dabei noch am Anfang. Obgleich schon einige Texte zu dem Thema existieren, dominieren bislang noch Überblicksdarstellungen (Berger/Egbert/Janssen i. E.; Mannich 2024), rechtliche Studien (Kipker/Gärtner 2015; Martini et al. 2016) und Analysen des polizeilichen und politischen Diskurses (Berger/Egbert 2025; Lehmann 2017a). Soziologische Untersuchungen liegen bislang nur in Ansätzen vor, insbesondere bezüglich der Dimension der Visualität (Arzt/Schmidt 2021; Tuma et al. 2022). Die Fragen, welche Effekte mit den Bodycams als audiovisuelle Aufnahmegeräte für die polizeiliche Arbeit einhergehen, welche Rolle und Funktion sie in Polizei-Bürger:innen-Interaktionen einnehmen (können) und wie dies mit ihrer technischen Eigenart zusammenhängt, sind bislang empirisch unterbelichtet geblieben. Vor diesem Hintergrund diskutieren wir Bodycamaufnahmen in diesem Beitrag als videoanalytischen Gegenstand. Bodycamaufnahmen stellen, wie wir zeigen werden, einen besonderen Gegenstand für die Videoanalyse, und insbesondere die Videointeraktionsanalyse dar, da sie Aufnahmen von natürlichen Situationen von Polizei-Bürger:innen-Kommunikation liefern, die genutzt werden können, um diese im Nachhinein feingliedrig zu analysieren.

Im Gegensatz zu den normalerweise in der Videointeraktionsanalyse genutzten Aufnahmen sind Bodycamvideos nicht von den Forschenden selbst, sondern von den beteiligten Polizist:innen aufgenommen. Dies geht mit relevanten Einschränkungen und Besonderheiten für die Analyse einher. Diese werden im Beitrag ebenso diskutiert wie die ethischen Herausforderungen, die mit der Nutzung von polizeilichem Videomaterial, zumal wenn es sich um Leaks handelt, einhergehen. Am Beispiel des Stand Oktober 2024 einzigen öffentlich einsehbaren natürlichen Bodycamvideos von deutschen Polizeikräften möchten wir aber ebenso demonstrieren, welches Potenzial dieses Videoformat für die soziologische Videoanalyse und gleichermaßen für die Polizeiforschung bietet. Dabei fokussieren wir in unserer Analyse insbesondere auf die Art und Weise, wie Bodycams von Polizist:innen genutzt werden und was die Kameraaufnahmen (nicht) über die aufgenommenen Interaktionen verraten können.[2]

2 Die sich im gezeigten Fallbeispiel ebenfalls aufdrängende Frage nach der Verhältnismäßigkeit der gezeigten polizeilichen Zwangsmaßnahmen kann aus Platzgründen an dieser

Wir beginnen mit einer Einführung in die Nutzung und Legitimierung von Bodycams in der deutschen Polizei und verorten danach deren Implementierung in die polizeiliche Arbeit in den Kontext der neuen Sichtbarkeit von Polizeiarbeit (Goldsmith 2010). Daran anschließend führen wir in die Begleitumstände des von uns exemplarisch analysierten Bodycamvideos und in unsere videoanalytische Herangehensweise ein. Danach stellen wir unsere Ergebnisse vor, um anschließend Fragen hinsichtlich der Objektivität von Bodycamaufnahmen und der interaktionistischen Bedeutung der Kamera aufzuwerfen. Nach einer methodologischen Reflexion schließen wir den Beitrag mit einem Fazit, was insbesondere auf weiterführendes (video-)soziologisches Forschungspotenzial im Bereich der polizeilichen Bodycams verweist.

2. Bodycams in Deutschland

Bodycams werden von der deutschen Polizei seit dem Jahr 2013 eingesetzt. Seitdem haben sich die Kameras rasch bei den Polizeibehörden in Deutschland verbreitet, sodass derzeit alle 16 Landespolizeien sowie die Bundespolizei (Stand Oktober 2024) in Deutschland solche Kameras nutzen (Berger/Egbert/Janssen i. E.). Die schnelle Verbreitung ist bemerkenswert, da die den Kameras zugeschriebenen Wirkungsweisen nicht erwiesen sind. Die vorliegenden empirischen Ergebnisse dazu sind vielmehr widersprüchlich (vgl. z. B. Mannich 2024; Zander 2016).

Die erhoffte Wirkungsweise von Bodycams seitens der Polizei zielt vornehmlich auf die Reduzierung von Gewalttaten und Widerstandshandlungen gegenüber Polizist:innen. Durch die Ankündigung, dass eine Bodycam angeschaltet wird, sollen potenzielle Gewalttäter:innen von der Ausübung ihrer Handlungen gegenüber Polizist:innen abgehalten werden. Die Annahme dahinter ist, dass durch die (vermeintlich) beweissichere Aufnahme eine hohe Sanktionswahrscheinlichkeit antizipiert und gewalttätiges Handeln mithin abgeschreckt wird (Zander 2016). Diese Disziplinarlogik baut jedoch auf Rational-Choice-Annahmen auf, die hinsichtlich ihrer Übertragbarkeit auf unübersichtliche und stark emotionalisierte Polizei-Bürger:innen-Interaktionen, potenziell durch Alkohol- oder Drogeneinfluss begleitet, zweifelhaft sind und in der Forschung entsprechend kritisch gesehen werden (z. B. Mannich 2024, S. 406).

Ein weiterer Einsatzzweck der Bodycam ist aus polizeilicher Sicht die verbesserte Möglichkeit zur Erhebung von Beweismitteln. Durch die Möglichkeit, das

Stelle nicht näher beleuchtet werden, dies behalten wir uns für einen separaten Aufsatz vor, indem auch die Reaktionen z. B. von Polizeivertreter:innen auf das Video vorgestellt und diskutiert werden.

Einsatzgeschehen situationsnah aufnehmen zu können, versprechen sich Polizeibehörden, niedrigschwellig gerichtsverwertbare audiovisuelle Dokumentationen von Polizei-Bürger:innen-Interaktionen anzufertigen, die in einem Gerichtsverfahren einerseits als Belastung für angeklagte Bürger:innen, andererseits aber auch als Entlastung für Polizist:innen bei Vorwürfen von z. B. Körperverletzung im Amt genutzt werden können (Berger/Egbert 2025).

Diese bereits häufig als einseitig kritisierte Handhabung (z. B. Aden/Fährmann 2019) von Bodycams und ihren Aufnahmen in Deutschland, die der Polizei die alleinige Deutungshoheit und Handlungsmacht über die (Bilder der) Kameras zugesteht, hat ihre Quelle in dem deutschen Sonderweg der Legitimierung von Bodycams. Denn im Gegensatz z. B. zu den USA, wurden die Kameras in Deutschland allein mit der Zielsetzung angeschafft, die Polizei vor der Bevölkerung zu schützen, während anderswo hauptsächlich der Schutz der Bevölkerung vor der Polizei im Vordergrund steht (Lehmann 2017a).

Ein weiterer (inoffizieller) Grund für die Einführung von Bodycams wird darin gesehen, ein ›Gegenfilmen‹ zu ermöglichen – also auf das zunehmende gefilmt werden durch Betroffene und Beobachtende von Polizeihandeln zu reagieren. Daher sind die Bodycams auch vor dem Hintergrund der Diskussion um die »neue Sichtbarkeit« von Polizeiarbeit (Goldsmith 2010) zu verstehen, wie wir im folgenden Abschnitt zeigen.

3. Neue Sichtbarkeit polizeilicher Arbeit

Polizeibehörden werden zunehmend mit ungewollter Sichtbarkeit konfrontiert (vgl. Meyer/Grauert/Oberzaucher; Godarzani-Bakhtiari/Tuma i. d. B.), denn historisch etablierte Deutungshoheiten werden durch neue Formen und Kanäle der Kritik durch Sichtbarkeit deutlich herausgefordert, wie es Goldsmith (2010) unter dem Schlagwort der »neuen polizeilichen Sichtbarkeit« konstatiert. Denn auch wenn Medien- und Imagearbeit bereits seit den Anfängen moderner Polizeibehörden einen wesentlichen Bestandteil ihrer Arbeit ausmachen (Mawby 2002), spitzen die aktuellen Veränderungen die polizeilichen Herausforderungen und Aufwände im Feld polizeilicher (Un-)Sichtbarkeit auf neue Art zu. So ist die videobasierte Dokumentation von polizeilichem Handeln mit mobilen Endgeräten für Polizist:innen zu einer erwartbaren, wenngleich polizeilicherseits üblicherweise unerwünschten, Normalität geworden (Newell 2019).

Polizeivertreter:innen begründen diese Vorbehalte gegenüber Filmaufnahmen von Polizeihandeln mit der Sorge um Persönlichkeitsrechte individueller Beamt:innen sowie potenziell verfälschter und aus dem Zusammenhang gerissener Darstellungen (z. B. Probandt 2020). Zu erklären ist diese Abwehrhaltung aber auch damit, dass mit der neuen Sichtbarkeit der Polizeiarbeit für

die Institution Polizei vor allem das Risiko sinkender Legitimationsbestände und steigenden Misstrauens gegenüber der Polizei bei Teilen der Bevölkerung (Browning et al. 2021) einhergeht. Der (drohende) Verlust an definitorischer Macht über die eigene Sichtbarkeitspolitik konfrontiert Polizeien vermehrt mit einem Rechtfertigungsdruck, der die polizeiliche Arbeit zu einer verstärkten Gestaltung ihrer Sichtbarkeitspraktiken drängt (Ullrich 2018). Neben der Ausweitung von Videotechnologien führte dies u. a. dazu, dass die Öffentlichkeitsarbeit der Polizei offensiver und aktiver geworden ist, insbesondere mit Blick auf die Sozialen Medien (Graßl 2023, S. 50). Insofern begegnen Polizeibehörden der historischen Herausforderung neuer Sichtbarkeit zunehmend mit einer Aneignung von Sichtbarkeitspraktiken und -instrumenten, um diese im Sinne organisationaler Rationalitäten einzusetzen. Auch die Einführung von Bodycams bei deutschen Polizeien steht in dieser Kontinuität. So wird den Kameras, und den damit produzierten Aufnahmen, die Funktion zugesprochen dem (drohenden) Verlust diskursiver und definitorischer Macht durch eine filmende Zivilgesellschaft mithilfe eigener Visualitätsproduktion zu begegnen und so eine vermeintliche »Waffengleichheit« (Teggatz zit. n. Brenner 2021, S. 21) zurückzugewinnen.

4. Fallbeispiel: Bodycam-Video eines Bundespolizeieinsatzes in München

4.1 Kontextualisierung Fallbeispiel

Das Bodycamvideo, das wir im Folgenden analysieren, ist unseres Wissens das einzige Bodycam-Video eines realen Polizeieinsatzes in Deutschland, das öffentlich zugänglich ist. Dies liegt daran, dass Bodycamvideos in Deutschland von den Polizeien nicht veröffentlicht werden (müssen) und dass im hier analysierten Fall allem Anschein nach ein Leak für die öffentliche Verfügbarkeit des Videos verantwortlich ist. Es wurde dem Nachrichtenportal Focus Online im Sommer 2021, kurz vor Beginn des Prozesses gegen den betroffenen Bürger (B.), beinahe anderthalb Jahre nach dem Vorfall, zugespielt. Dem Medium nach wurde das knapp neun Minuten lange Video in »voller Länge« (Schattauer 2021c, o. S.) publiziert. Zu diesem Zweck wurden von den Redakteur:innen die Gesichter der unmittelbar beteiligten Personen und Artefakte, die Rückschlüsse über eine Identität zulassen (z. B. Brief), verpixelt und die Stimmen verzerrt. Gegenstand des Videos ist der Versuch zweier Bundespolizist:innen am Münchener S-Bahnhof Isartor, die Personalien des Fahrgasts B. aufzunehmen. B. wurde zuvor bei einer Ticketkontrolle von Sicherheitskräften der Deutschen Bahn verdächtigt, mit einem ungültigen Ticket zu fahren, da dieser seine

Zeitkarte beschriftete. Diese Notiz auf dem Ticket führte allerdings – entgegen der Annahme der Sicherheitskräfte – nicht dazu, dass es ungültig wurde, dem Polizeieinsatz liegt also eine Fehleinschätzung zugrunde (Schattauer 2021c). Den gerufenen Bundespolizist:innen wurde von den Sicherheitskräften der Deutschen Bahn eine aggressive Person angekündigt, weshalb, so kritisiert dessen Anwalt, »für die Beamten […] der Bahnsteig von vornherein ein Tatort und mein Mandant ein Täter (war)« (zit. n. Schattauer 2021b. o. S.). Das Bodycamvideo wurde im anschließenden Gerichtsverfahren als Beweismittel genutzt, in dessen Ausgang B. Anfang November 2021 zu einer Geldstrafe von 90 Tagessätzen à 30 Euro (insg. 2700 Euro) verurteilt wurde (Schattauer 2021a). Gegen den beteiligten Bundespolizisten wurde Anfang 2022 ein Ermittlungsverfahren, u. a. wegen möglicher Körperverletzung im Amt, eingeleitet – dieses Verfahren wurde mangels hinreichenden Tatverdachts allerdings eingestellt (Schattauer 2022).

4.2 Videoanalyse von Bodycamaufnahmen

Polizeiliche Bodycamaufnahmen stellen sowohl hinsichtlich ihrer Erhebungs- als auch der Analysekontexte eine videoanalytische Datensorte mit spezifischen Besonderheiten dar. Denn obgleich sie dafür prädestiniert sind, alltagsnahe Polizei-Bürger:innen-Interaktionen aufzunehmen, bilden sie kein klassisches Material für eine *Videointeraktionsanalyse*, da sie weder von den Forschenden selbst noch zu Forschungszwecken hergestellt werden (Tuma 2018b). Dies gilt gerade für jene Art von Videointeraktionsanalyse, die in ethnografische Feldaufenthalte eingebettet und mithin Teil von Videografien sind (Knoblauch/Vollmer 2018). Polizeiliche Bodycams sind demgegenüber Videoinstrumente der Polizei, deren Aufnahmen von Polizist:innen nach eigenem Ermessen angefertigt werden. Als Forscher:in ist man somit abhängig von den entsprechenden Kamera-Handlungen und -Entscheidungen der Polizist:innen, was mit spezifischen methodologischen Problemen zusammenhängt (siehe Abschnitt 7). Nichtsdestotrotz stellen Bodycamaufnahmen wichtige Ressourcen für soziologische Videoanalysen dar, nehmen sie doch natürliche Situationen auf, die Kommunikations- und Interaktionssequenzen zwischen Polizist:innen und Bürger:innen zeigen und einer detaillierten Analyse zugänglich machen (vgl. allg. dazu a. Meier zu Verl et al. 2024). In diesem Sinne wird die Aufnahme als »Forschungsmedium« benutzt, wie es für die Videografie bzw. Video-Interaktionsanalyse typisch ist (Tuma et al. 2013, S. 47).

Für die Analyse erscheint jedoch der externe Erhebungskontext von Bodycamaufnahmen eher eine Durchführung einer *Videoproduktanalyse* nahezulegen, bei denen Videos nicht primär als Forschungsmedium, sondern

als mediales Produkt im Fokus stehen und mit einem entsprechenden Erkenntnisinteresse untersucht werden (Tuma/Schnettler/Knoblauch 2013, S. 47; Tuma 2018b, S. 426 f.). Jedoch bringen die naheliegenden Verfahren der Videoproduktanalyse, etwa die wissenssoziologisch-hermeneutische Videoanalyse (Reichertz/Englert 2021) oder die dokumentarische Methode (Bohnsack 2011), ihre eigenen Problematiken hinsichtlich Bodycams mit sich. Denn diese klassischen Methoden für Videoproduktanalysen fokussieren in ihrer methodischen Analyse erstens grundsätzlich auf audiovisuelle Werke, die einen hohen inszenatorischen Charakter aufweisen und zahlreichen systematischen künstlerischen und dramaturgischen Gestaltungseingriffen – insbesondere auch vor und nach der Aufnahme – unterliegen, wie z. B. Werbeclips oder Kinofilme. Davon sind Bodycamvideos indes in nennenswertem Maße abzugrenzen. Obgleich sie ebenfalls fremderzeugt sind und hinsichtlich ihrer Laufzeit, Perspektive etc. dem Eingriff der Forschenden entzogen sind, sind ihre Aufnahmen hinsichtlich ihrer artifiziellen Qualität nicht mit typischen Videoprodukten vergleichbar. Insbesondere die perspektivische Besonderheit der Bodycam, welche in ihrem Bildausschnitt automatisch den verhältnismäßig grobmotorischen Bewegungen des (Ober-)Körpers folgt, lässt eine Analyse und Interpretation ikonologischer Sinndimensionen einer dokumentarischen Interpretation, wie sie bei der dokumentarischen Methode im Zentrum steht, ebenso ins Leere laufen, wie die strenge analytische Trennung von darstellenden und dargestellten Handlungen, wie sie die wissenssoziologisch-hermeneutische Videoanalyse vorsieht. Zweitens liegt in der situativen Involviertheit der Bodycam als Aktant (siehe Abschnitt 5) eine datenspezifische Besonderheit, die eine Analyse, die den Dokumentensinn allein aus dem Abgebildeten zieht, wie es die dokumentarische Methode vorsieht, nicht hinreichend zu erfassen mag. Denn Analysekontext und die der Aufnahme zugrundeliegende Erhebungssituation ergeben sich bei Bodycamaufnahmen stets erst aus dem, i. d. R. nur ausschnitthaft aufgenommen, Interaktionsgeschehen, weshalb das im Bodycamvideo sichtbare Handeln der gefilmten Akteur:innen stets eine besondere Bedeutung zukommen sollte. Dementsprechend liegt eine Analyse von Bodycamaufnahmen durch ihren *hybriden videoanalytischen Charakter* quer zu den etablierten Verfahren der Videoanalyse. Statt sich also auf eine dieser Verfahren zu beschränken, erscheint eine methodische Kombination notwendig, wobei die Schwerpunktsetzung freilich vom jeweiligen Erkenntnisinteresse ausgehen sollte. In unserem Fall liegt es nahe, die Videointeraktionsanalyse mit der videoproduktanalytisch interessierten wissenssoziologisch-hermeneutischen Videoanalyse zu ergänzen. Durch die Kombination dieser beiden sozialtheoretisch ähnlich fundierten Verfahren ist das methodologische Fundament hinreichend passgenau zu den Fragen, die wir an das Material stellen. Denn wir interessieren uns sowohl

für das Handeln vor der Kamera als auch für die Handlung der Kamera (Reichertz 2018, S. 103). Das Zeigen der Kamera – was stets bestimmte Deutungsangebote nahelegt (Englert 2013) – ist uns ebenso wichtig wie das auf den Aufnahmen Gezeigte.[3]

Uns interessiert im Folgenden daher nicht nur das Geschehen vor der Kamera, sondern auch die »(sich) im Material dokumentierenden Besonderheiten und Probleme der Erhebungssituation« (Tuma 2018b, S. 427), die sich auch auf die Charakteristika der genutzten Kamera und den damit zusammenhängenden Eigenarten der Aufnahme – was ist (nicht) sichtbar? – beziehen lassen, was eine Videoproduktanalyse nahelegt (hier als wissenssoziologisch-hermeneutische Videoanalyse angelegt). Gleichzeitig braucht es die Videointeraktionsanalyse, da unser Forschungsinteresse wesentlich in der Frage besteht, welche Bedeutung der Kamera in der Polizei-Bürger:in-Interaktion von den Beteiligten zugesprochen wird und wie die Anwesenheit der Kamera das Interaktionsgeschehen beeinflusst. Dafür eignet sich eine auf die Sequenzialität der aufgenommenen Interaktion fokussierte Herangehensweise, wie sie im Zentrum der Videointeraktionsanalyse steht (Tuma/Schnettler/Knoblauch 2013), besonders gut. Aus Darstellungsgründen stellen wir Analyse und Interpretation im folgenden Abschnitt in fließenden Übergängen dar, auch wenn diese sowohl bei der Videointeraktionsanalyse als auch der hermeneutischen Videoanalyse als getrennte zweistufige Verfahren konzipiert sind.

5. Videoanalyse

Wir konzentrieren uns in unserer Videoanalyse des exemplarischen Bodycamvideos auf zwei Schlüsselsequenzen – der Moment des Einschaltens der Kamera sowie das Handgemenge und die Fixierung von B. durch die Bundespolizist:innen. Im Zuge dessen fokussieren wir uns insbesondere auf die Bodycam als Aktant sowie die Rolle von Polizist:innen als Filmschaffende.

5.1 Schlüsselsequenz 1: Einschalten der Bodycam

Die erste (Schlüssel-)Sequenz des Videos beginnt mit der Frage der Polizistin 1 (die auch die Bodycam trägt), wie der Name von B. sei. Zu diesem Zeitpunkt ist Herr B. bereits deutlich aufgebracht und stellt rufend fest, z. B. – auf Französisch –,

3 Dass zwischen beiden Verfahren relevante Anschlusspotenziale bestehen, betonen aus Perspektive der Videografie/Videointeraktionsanalyse bereits Tuma, Schnettler und Knoblauch (2013, S. 53).

dass er die Verhaltensweisen der Polizist:innen als Provokation empfinde. Er bezieht sich dabei insbesondere auf eine Berührung der Polizistin 1, durch die er sich in seiner körperlichen Integrität beeinträchtigt sah, die aber nicht auf der Bodycamaufnahme zu sehen ist. Herr B. sitzt mittig von zwei S-Bahn-Gleisen auf einem der metallenen Sitzplätze einer Sitzbank. Frontal vor ihm steht der andere am Einsatz beteiligte Bundespolizist 2, in einem Abstand von rund einem Meter. Schräg links vor ihm – von B. aus gesehen – steht die filmende Polizistin 1 in ca. zwei Metern Entfernung. Ein paar Meter weiter weg steht eine DB-Sicherheitskraft (gelbe Weste). Gemeinsam bilden sie einen stehenden Halbkreis um den sitzenden B (siehe Abb. 1).

Abbildung 1: Schlüsselsequenz 1, Starten Aufnahme Bodycam

Polizist 2 weist B. darauf hin, dass sie ihn mit auf die Dienststelle nehmen würden, sofern er nicht seinen Namen nenne bzw. seinen Personalausweis übergebe. B. spricht in dieser ersten Sequenz nur auf Französisch, Polizist 2 konstatiert dennoch, dass B. Deutsch verstehe (Z. 14, Tab. 1).[4] Ab Sekunde 25 (Z. 19, Tab. 1) hört man die Polizistin 1 sagen, dass sie nun die Bodycam anschalte, nachdem sie ein weiteres Mal vergeblich um den Ausweis von Herrn B. gebeten hat, dessen erneute Weigerung sie mit dem Ausruf »Boah, Alder« (Z. 19, Tab. 1) quittiert.

4 Die Transkription der Videosequenzen erfolgte gemäß den Richtlinien des GAT2-Minimaltranskripts.

Tabelle 1: Auszug Transkript Schlüsselstelle 1, Starten Aufnahme Bodycam

1	P1:	[Ihren] Namen jetzt.	#00:00:01-6#
2	B:	[( )]	
3	B:	Vous vous foutez de moi ou quoi	#00:00:02-3#
4	P1:	Ihren Namen.	#00:00:03-3#
5	B:	Vous ne me touchez pas!	
6	P1:	Wenn Sie jetzt [nicht hören, dann gehen Sie]	
7		[Je ne vous ai pas autorisé]	#00:00:07-1#
8		[mit auf die]	
9		[à me toucher!]	
10	P1:	Dienststelle.	#00:00:07-6#
11		(3.1)	#00:00:10-7#
12	P1:	Ja. (2.5)	#00:00:13-6#
13	B:	Ah vous n'aimez pas les étrangers, eh!	#00:00:13-4#
14	P2:	Sie können deutsch. Sie können	#00:00:17-4#
15		[( )]	
16	B:	[Ne] me touchez pas! C'est de la provocation!	#00:00:19-1#
17	P1:	Geben Sie jetzt [Ihren Ausweis.]	#00:00:21-3#
18	B:	[C'est de la provocation.] ( )	#00:00:22-2#
19	P1:	Boah, Alder. Ähm ich mach jetzt die Bodycam an.	#00:00:27-8#
20		Alles, was Sie jetzt sagen wird gefilmt.	
19	B:	Je m'en fous! Comme ça on verra avec quelle ma-	#00:00:31-6#
20		nièr[e? ( )]	
21	P1:	[Sie werden jetzt gefilmt.] Haben Sie verstan-	#00:00:33-8#
22		den?	
23	B:	Mademoiselle! Elle m'a touché!	#00:00:36-1#
24	P1:	Sie /	#00:00:37-3#
25	B:	C'est une provocation. Elle n'a ni le droit, ni	#00:00:45-1#
26		l'autorisation de ma part! C'est une provocation.	
27			
28	P1:	Nein.	#00:00:45-9#
29	B:	PRO-VO-CATION	#00:00:47-2#
30	P1:	Nein. (…) Wir brauchen [Ihre Perso]nalien.	#00:00:50#

Die Videoaufnahme beginnt rund eine halbe Minute bevor die Polizistin 1 ankündigt, dass sie die Bodycam aktiviert. Dies zeigt an, dass von ihr eine Bodycam mit Pre-Recording-Funktion genutzt wird. Das bedeutet, dass die Bodycam durchgehend aufzeichnet, das Aufgezeichnete aber – sofern keine Aufnahme

aktiv gestartet wird –, nach einer gewissen Zeitspanne (hier: 30 Sekunden) automatisch wieder überschreibt.[5]

Die Polizistin 1 kündigt deutlich hörbar an, dass sie die Bodycam anschalten werde (Z. 19 ff., Tab. 1). Konkret sagt sie: »Ich mach jetzt die Bodycam an, alles, was Sie sagen, wird gefilmt«. Nachdem Herr B. zu einer erneuten Beschwerde ansetzt, wiederholt die Polizistin 1: »Sie werden jetzt gefilmt, haben Sie verstanden?« (Z. 21 f., Tab. 1). Auf die Ankündigung der Polizistin 1, die Bodycam anzuschalten, ist bei B. keine Veränderung im Auftreten erkennbar. Offenbar versteht er sie zwar – denn er sagt »ist mir egal« (»je m'en fous«, Z. 19, Tab. 1) – spricht aber anstatt auf das Filmen einzugehen, länger auf die Polizistin 1 in unverändert aufgebrachter Weise ein.

5.2 Interpretation und Kontextualisierung von Schlüsselsequenz 1

Die Entscheidung der Polizistin 1, die Aufnahme zu starten und die Bodycam aus dem Bereitschaftsdienst zu wecken, folgt unmittelbar auf ihren resignierenden Ausruf »Boah, Alder« (Z. 19, Tab. 1). Daraus lässt sich eine gewisse Resignation schließen, da sie die Situation als festgefahren empfindet. Das Einschalten der Bodycam scheint in dieser sich eskalativ zuspitzenden Interaktionsdynamik somit wie der Versuch, die Spannweite zwischen Androhung und Einsatz von Gewalt um eine Zwischenstufe zu ergänzen. Warum aber wird gerade in diesem Moment imaginierter situativer Verengung die Bodycam eingeschaltet?

Als Begründungen sind mehrere Erklärungsansätze denkbar, die als Reaktion und Ausdruck der beschriebenen Resignation auftreten. Einerseits ist die Bodycam als »Technofix« diskursiv als Instrument aufgeladen, das in der Lage sein soll, polizeiliche Handlungsprobleme zu »lösen«. Das Einschalten der Bodycam markiert hier einen polizeilichen Appell zum Handlungs-*Turn* des Gegenübers: Das antizipierte Risiko eines durch die Sicht- und Beweisbarkeit geschaffenen sozial-ökonomischen Schadens (durch Gerichtsprozess, Verurteilung etc.) seitens des sog. ›polizeilichen Gegenübers‹, der sich in der Bodycamaufnahme materialisiert, soll die Bodycam zum proaktiven Überwältigungsmittel machen. Die Macht dieses Mittels baut jedoch auf sozial-ökonomischer Verhaltensrationalität auf (siehe Abschnitt 3), die an der interaktiven Realität komplexer Situationen scheitern kann, wie die beschriebene Sequenz gleichsam par excellence verdeutlicht – B. registriert das Anschalten der Kamera, ändert sein Verhalten aber nicht. Das Einschalten der Kamera kann als eine Hoffnung nach einer von außen kommenden Deeskalation gelesen werden, in der das technische Gerät

5 Das Bundespolizeigesetz (BPolG) legt in § 27a Abs. 3 fest, dass Pre-Recording »bis zu einer Dauer von 30 Sekunden« erlaubt ist und mithin die Bodycams »im Bereitschaftsbetrieb in ihrem Zwischenspeicher kurzzeitig Daten erfassen« dürfen.

eine Entproblematisierung der Situation liefert, ohne dass die beteiligten Polizist:innen dafür tatsächlich etwas an ihrem Verhalten ändern müssten.

Jedoch liegt in der Handlung des Einschaltens der Kamera gleichzeitig ein definitorisches Moment, welcher B. als renitent, aggressiv und gefährlich markiert, denn die juristische Grundlage des Einschaltens liegt für Beamt:innen der Bundespolizei entweder in einer »erheblichen Bedeutung« der zugrundeliegenden Straftat bzw. Ordnungswidrigkeit, oder der »Gefährlichkeit« von Personen begründet (§ 27 Abs. 1 BPolG). Eine solche Markierung steht einer potenziellen Deeskalation diametral gegenüber, indem es das Gegenüber in einer binär-antagonistischen Selbst-Fremd-Figuration verortet (Schmidt 2021, S. 169–182), da dieses als »überwachenswert« (Herrmann 2022, S. 285) eingestuft wird (vgl. auch Lehmann 2017b, S. 186). Das Einschalten der Bodycam ist insofern gleichzeitig eine Vorbereitungshandlung für eine antizipierte körperliche Auseinandersetzung, die audiovisuell aufgezeichnet werden soll, um Beweisbarkeit für potenzielle Widerstände bzw. die Korrektheit des eigenen Handelns herzustellen. Dadurch schafft das Einschalten der Bodycam aber keine alternativen Situationsverläufe für die Zukunft, sondern markiert im Interaktionsgeschehen einen Kipppunkt in der Polizei-Bürger:innen-Interaktion, der zwar die Hoffnung nach alternativen Ausgängen sprachlich bezeugt, aber gleichzeitig in der Folge eher eskalative Handlungsrepertoires nahelegt.

5.3 Bodycam als Aktant: Polizist:innen als Filmschaffende

Wie die vorangehenden Überlegungen verdeutlichen, prägt die Anwesenheit der Kamera die antizipierten Handlungsmöglichkeiten, bedingt die tatsächlichen Handlungen der Beteiligten und prägt somit das Geschehen. Die Kamera wird somit selbst zum Aktanten (Latour 2022, S. 219). Dass die Bodycam die Art und Weise, wie Polizist:innen agieren, verändern kann, wurde in der Literatur bereits mit Blick auf die Kamera als Disziplinierungsinstanz für die Beamt:innen selbst thematisiert (vgl. auch Herrmann 2022, S. 283 f.). So wird beispielsweise der Kamera zugeschrieben, das Gewaltrisiko in Polizei-Bürger:innen-Interaktionen zu potenzieren, da deren Sichtbarkeit den Formalitätsgrad im polizeilichen Handeln steigert. Das Gewaltrisiko beeinflusst dies auf zwei Weisen: Einerseits durch formalisierungsbedingte Zurückhaltung (keine Fehler machen wollen, die auf Video dokumentiert sind), andererseits durch die durch formelle Sprache geschaffene Distanz (vermeintliche vorbildliche Kommunikation, mit dem Risiko die Distanz zum Gegenüber zu erhöhen) (Kersting et al. 2019, S. 19).

Zusätzlich manifestiert sich im Einschalten der Bodycam ein polizeiliches Rollenverständnis als aktiv Filmende, das Polizist:innen die Funktion zuweist, die Kamera nicht nur zum richtigen Zeitpunkt einzuschalten, sondern sich auch richtig zu positionieren, damit die Kamera das Geschehen hinreichend

dokumentieren kann (vgl. auch Lehmann 2017b, S. 188 f.). Polizist:innen werden durch den Einsatz von Bodycams somit zu Filmschaffenden (»filmmakers«) (Sandhu/Simmons 2022). Denn bereits der Akt des Einschaltens der Kamera benötigt aktives Handeln der Polizei und es bedarf auch weitergehender Filmhandlungen, um verwertbare Bilder aufnehmen zu können. Wir übernehmen im Folgenden den Begriff der *Film*schaffenden, obgleich wir die Aufnahmen von Bodycams gemäß ihres technischen Speicherhintergrunds, ihres Miniaturformats und mithin ihrer Mobilität und Alltagstauglichkeit als *Videos* bezeichnen (vgl. dazu allg. Tuma 2018a, S. 2). Denn mit Filmschaffenden kann die aktive Rolle der aufnehmenden Person in den Vordergrund gerückt werden, beispielsweise (s. u.) mit Verweis auf Narration, Perspektive und Schnitt. Wir folgen damit auch der Idee von Sandhu und Simmons (2022, S. 593), die den Begriff der Filmschaffenden vor allem deshalb wählen, weil sie Bodycams als »storytelling devices« verstanden wissen wollen, was ihren Aktantenstatus hervorhebt.

Sandhu und Simmons (2022, S. 598 f.) beziehen sich in ihrer Analyse von Polizist:innen als Filmschaffende auf zwei kinematografische Entscheidungen, die von den Polizist:innen zu treffen sind, mit Blick auf den Schnitt und die Kameraperspektive: Die Entscheidung hinsichtlich des *Schnitts* rekurriert insbesondere auf das Ermessen der Beamt:innen, wann die Kamera anzuschalten ist. In Deutschland ist es den Polizist:innen weitgehend freigestellt, die Kamera zu tragen und sie einzuschalten. Wie schon oben thematisiert, scheint sich die Polizistin 1 im dargestellten Fall mit dem Anschalten der Kamera zu erhoffen, dass die festgefahrene Situation aufgelöst wird. Womöglich war ihr auch klar, dass eine neue Eskalationsstufe in der Interaktion kurz bevorsteht, weshalb sie die Kamera angeschaltet hat. Demnach ist ihre Macht als Filmschaffende mit Blick auf den Schnitt stark ausgeprägt. Dies verdeutlicht sich auch im Vorwurf von B, dass die Polizistin 1 ihn bereits angefasst habe (»Elle m'a touché«; Z. 21, Tab. 1), ein Vorwurf, der sich auf einen Zeitpunkt vor der Aufnahme bezieht, sich gegen die Filmende richtet, aber durch ihr erst später stattfindendes Einschalten der Bodycam nicht visuell erfasst ist.

Mit Blick auf die *Kameraperspektive* sehen wir in der Aufnahme, dass die Polizistin 1, sobald sie nicht aktiv in die körperliche Auseinandersetzung eingebunden ist, in mehreren Situationen bemüht ist, ein paar Schritte Distanz zwischen sich und das Geschehen zu bringen und dabei ihren Oberkörper (an dem die Kamera befestigt ist) weiter auf das Geschehen richtet – womit eine intentionale Perspektivität naheliegt. Zudem macht die Polizistin 1 ihre Selbstreflexion als ›Kamerafrau‹ auch sprachlich explizit, in dem sie einem zu Hilfe kommenden (mutmaßlichen) zivilen Kollegen erklärt, dass sie auch alles filme (00:04:03). So markiert sie die eigene, aktive filmische Tätigkeit als interaktionsrelevant und informiert gleichzeitig den neu hinzugekommenen (mutmaßlichen) Zivilpolizisten, dass auch sein Handeln späterer Sicht- und Beweisbarkeit unterstehen kann.

Sowohl Schnitt als auch Kameraperspektive bedingten dabei die spezifischen Charakteristika eines »camera perspective bias« (Stoughton 2018, S. 1408), der

das Geschehen grundsätzlich aus der Perspektive des:der Polizist:in aufnimmt. Diese Verzerrungsform, die u. a. anhand von Experimenten mit Verhörsituationen untersucht wurde (vgl. Godarzani-Bakhtiari/Tuma i. d. B.), bedingt, dass der Blickwinkel, aus dem Menschen etwas beobachten, einen Effekt darauf hat, wie sie das Beobachtete wahrnehmen und bewerten – was als besonders prominent im Video wahrgenommen wird, wird auch als potenter angenommen (Stoughton 2018, S. 1409).

Während eine bewusste Positionierung der filmenden Polizistin 1 in einigen Momenten übersichtliche Aufnahmen produziert, zeichnet sich der Großteil des Videomaterials jedoch durch die Bodycam-typische, unübersichtliche Perspektivität ›mitten im Geschehen‹ aus, wie sich an einer zweiten Schlüsselsequenz verdeutlichen lässt.

5.4 Schlüsselsequenz 2: Handgemenge und Fixierung

Nach dem gescheiterten Versuch der beiden Polizist:innen, B. mit auf die Wache zu nehmen, gerät dieser unter den Händen der Polizist:innen ins Straucheln und stürzt mit nur wenig Distanz zur Bahnsteigkante zu Boden, wobei er einen der Mitarbeiter der DB-Sicherheit mit zu Boden zieht. Auch ein Teil des Inhalts der Gepäckstücke von B. werden über den Bahnsteigboden verteilt. Die zwei Polizist:innen bemühen sich derweil die Kontrolle über den Körper von B. zu erlangen. Durch die Involviertheit der Polizistin 1 und der daraus entstehenden Nähe der Kamera, deren niedrige Bildrate und unausgeglichenen Schaukelbewegungen, ist das weitere Handeln in der Videoaufnahme jedoch nur noch schwer zu erkennen.

Tabelle 2: Auszug Transkript Sequenz 2, »Handgemenge und Fixierung«[6]

1	P1:	Oh man!	#00:02:06-0#
2	B:	AIDEZ-MOI	#00:02:08-7#
3	P1:	I-diot.	#00:02:09-8#
4	B:	AIDEZ-MOI	#00:02:10-6#
5	P1:	Uurghh ((zieht am rechten Arm von B))	#00:02:11-2#
6	B:	AIDEZ-MOI! ((Eine S-Bahn fährt ein))	#00:02:12-8#
7	P1:	Lassen Sie locker!	#00:02:14-4#
8	B:	( ) AIDEZ-MOI! ((P1 greift nach dem Kopf von B.))	#00:02:19-9#
9		( ) AIDEZ-MOI.	

6 Wir bedanken uns bei Jacques de Maillard und Félicien Faury für ihre Hilfe bei der Übersetzung der französischen Inhalte.

10	P1:	((greift nach dem rechten Arm von B., dieser bewegt	
11		seinen Mund in Richtung des nach ihm greifenden	
12		Arms von P1))	
13	P1:	<Ugh. NEIN. <Schlägt B an die Schläfe>>	#00:02:22-3#
14	P1:	Der will nach uns beißen!	#00:02:24-1#
15	B:	((dreht sich auf die Seite und schlägt in Richtung	#00:02:24-3#
16		von P1 und der Kamera))	
17	P2:	((legt das Knie auf den Hals von B))	#00:02:25-8#
18	P1:	Aah! FUCK! [Der hat mir voll ins Gesicht ge-	#00:02:27-5#
19		schlagen!]	
20	B:	[AIDEZ-MOI. AIDEZ-MOI.]	#00:02:27-5#
21	P2:	[Alles gut. Alles gut.] <HEY!	#00:02:28-6#
22		( ) <HältB den Mund zu, greift nach dem rechten	
23		Arm von B, hält diesen fest>>	
24	P1	((unverständlich, ca. 4 Sek)) ((P1 hält den linken	#00:02:34-0#
25	und	Arm von B))	
26	P2:		
27	B:	AIDEZ-MOI, AIDEZ-MOI. AIDEZ-MOI! AIDEZ-MOI. [AIDEZ-	#00:02:37-1#
28		MOI, AIDEZ-MOI.]	
29	P1:	[Kriegst du's hin?] Bitte. Ugh. ((P2 hält beide	#00:02:39-6#
30		Arme von B fest, P1 entfernt sich einen Schritt vom	
31		Geschehen, zieht Handschuh aus))	
32	B:	AIDEZ-MOI. AIDEZ-MOI, AIDEZ-MOI. ((unverständlich,	#00:02:46-7#
33		ca. 4 Sek))	
34	P1:	( ) ((zieht Smartphone hervor, bedient das Dis-	#00:02:49#
35		play))	
36	B:	((unverständlich, ca. 2 Sek))	#00:02:51-4#
37	P2:	Komm runter. ((P2 ist nicht mehr zu sehen))	#00:02:52-3#
38	B:	((unverständlich, ca. 3 Sek))	#00:02:55-5#
39	P1:	Hey!	#00:02:56-6#
40	B:	((unverständlich, ca. 4 Sek))	#00:03:00-9#
41	P1:	Lukas, kannst du mal ne Streife ( ) Bahngleis 2.	#00:03:16-1#
42		Der ist grad komplett ausgerastet. (2.0) 'ne	
43		Streife. Wir sind zu zweit. ( ) Der rastet komplett	
44		aus. Danke, danke.	

Zuerst liegt B. auf dem Rücken, mit den zwei Bundespolizist:innen über ihm, dann dreht, der weiterhin um Hilfe rufende B. (Z. 2, Tab. 2), seinen Oberkörper in Richtung des Bodens und stützt sich mit den Armen ab. In dieser deutlich angespannten Situation beschimpft die Polizistin 1 den sich weiterhin versperrenden B. unter einiger Anstrengung als »Idiot« (Z. 3, Tab. 2). Der linke Arm von

B. wird derweil schon länger von Polizist 2 gehalten. Ohne freie Hände bewegt B. seine Mundpartie in Richtung des Armes der Polizistin 1, welche ein erschrockenes »Nein!« von sich gibt und in einer schnellen Bewegung den Arm von B. loslässt und gegen dessen rechte Schläfe zu schlagen scheint (Z. 13, Tab. 2). Ihr Handeln begründet sie verbal als Reaktion auf die vorherige Bewegung von B.: »Der will nach uns Beißen!« (Z. 14, Tab. 2). Ob B. tatsächlich in das von Polizeiuniform und Handschuh umschlossene Handgelenk biss oder die Polizistin 1 mit seinem Handeln nur zum Loslassen bewegen wollte, ist in dem Video nicht zweifelsfrei zu erkennen. In der entstandenen Dynamik packt die Polizistin 1 wieder den rechten Arm von B., während dieser seinen linken Arm losreißt. Auch wenn der Arm noch von seinem eigenen Oberkörper blockiert ist, führt er damit einen Schlag in Richtung der Polizistin 1 aus, wobei aufgrund der Blickrichtung der Kamera nicht zu erkennen ist, ob und gegebenenfalls wo er diese trifft (Abb. 2).

Abbildung 2: Schlüsselszene 2, Handgemenge

Laut Ausruf von Polizistin 1 hat er deren Gesicht getroffen (Z. 18 f., Tab. 2). Der Polizist 2 nutzt den Moment, um sein linkes Knie im Nacken-Hals-Bereich von B. zu platzieren. Er hält B. den Mund zu und platziert sein rechtes Knie auf dem Oberkörper von B., wobei ein Großteil seines Gewichtes auf dem Nacken-Hals-Bereich von B. zu liegen scheint. Der Polizist 2 übt durch diese Position so viel Druck auf B. aus, dass er diesen auch ohne Zuhilfenahme seiner Hände fixieren kann. Derweil sind die Hilferufe von B. wieder lauter und verzweifelter geworden. Seine kurzatmigen Atemzüge sind deutlich zu hören und erwecken den Anschein erschwerter Atembedingungen. Die Polizistin 1 bittet ihren Kollegen mit der Frage »Kriegst du's hin? Bitte« (Z. 29, Tab. 2) offenbar darum, B. kurz allein (mit den nicht im Bild sichtbaren zwei Mitarbeitern der DB-Sicherheit) zu fixieren, jedenfalls übergibt sie im direkten Anschluss den von ihr festgehaltenen Arm von B. an ihren Kollegen und distanziert sich einen Schritt vom Geschehen. Während sich die Situation für die Polizist:innen etwas entspannt zu haben scheint, wirken die Hilferufe von B. zunehmend panisch. Der auf seinem Hals kniende Polizist 2 quittiert dies derweil mit einem ruhig formulierten »Komm runter« (Z. 37, Tab. 2), während die Polizistin 1 mit dem Smartphone nach Unterstützung ruft:

»Lukas, kannst du mal ne Streife (unv.) Bahngleis 2. Der ist grad komplett ausgerastet« (Z. 42, Tab. 2). Während dieses etwa 15 Sekunden dauernden Telefongesprächs filmt die Kamera nur noch den Boden, da die Polizistin 1 mit einem Bein auf dem Boden kniet und den Oberkörper nach vorn gebeugt hat.

5.5 Interpretation und Kontextualisierung von Schlüsselsequenz 2

An dieser zweiten Sequenz zeigt sich zunächst, wie uneindeutig ein großer Anteil des Videomaterials der Bodycam ist und wie selektiv Sichtbarkeit über die spezifische Perspektivität hergestellt wird. Im Gegensatz zu den Formen einer sichtbarkeitsschaffenden Perspektivität, wie sie anhand der ersten Sequenz verdeutlicht wurden, verdeutlichen sich an dieser zweiten Sequenz jene Perspektivitätselemente, die Sichtbarkeit verhindern. Es zeigen sich verschiedene Momente, in denen die Kamera verdeckt, oder vom Interaktionsgeschehen abgewendet wird oder gesprochene Worte unverständlich bleiben. In der ganzen Sequenz sind beispielsweise die zwei Mitarbeiter der DB-Sicherheit nicht zu sehen und an einer entscheidenden Stelle ist das Handeln des Polizisten 2 ebenfalls nicht erkennbar. Außerdem kommt während des Telefonats der Polizistin 1 eine weitere (zivil gekleidete) Person hinzu, ohne dass dies im Video an dieser Stelle wahrnehmbar wäre. Auch viele der Handlungen, die für eine anschließende Wahrheitsfindung vor Gericht relevant sein können, stellen sich im Material nicht eindeutig dar: So kann aus dem Videomaterial weder mit Sicherheit geschlossen werden, ob der beschuldigte B. die Polizistin 1 tatsächlich gebissen hat oder ob sein (in Abb. 2 zu sehender) Schlag diese tatsächlich traf. Auch wenn die Klärung dieser Details für eine juristische Bewertung im Sinne von § 113 und § 114 StGB[7] nicht unbedingt relevant erscheinen mögen, so verdeutlichen sie doch die Uneindeutigkeiten von Bodycam-Aufnahmen. Diese beziehen sich einerseits auf technische Limitationen wie etwa das deutlich eingeschränkte Sichtfeld, übermäßige Dynamik und Verzerrung (Lehmann 2017b, S. 180; Stoughton 2018, S. 1399 ff.; Williams et al. 2016). Andererseits verweisen sie darüber hinaus auf zwei weitere relevante Dimensionen des bereits angesprochenen *camera perspective bias*: So ist festzuhalten, dass ein großer Teil der Handlungen der Polizist:innen, wohin sie schauen, welche Körperhaltung und Gesichtsausdruck sie haben, anhand des Videos gar nicht analysierbar sind, da diese zumeist ausschnitthaft oder gar nicht im Bild erfasst sind. Stattdessen wird die Interaktion aus einer »polizeilichen Perspektive« (Arzt/Schmidt 2021, S. 329) visualisiert, welche vornehmlich die Handlungen

7 Für die Erfüllung eines Tatbestandes »Widerstand« (§ 113 StGB) ist ein passiver Widerstand ausreichend und für den Tatbestand eines »tätlichen Angriffs« (§ 114 StGB) ist das tatsächliche Treffen eines Schlages nicht notwendig, er muss nur auf den Körper eines:einer Polizist:in gerichtet sein.

von B. fokussiert. Es handelt sich dementsprechend um eine filmische Reproduktion eines polizeilichen Blickes (Arzt/Schmidt 2021; Sandhu/Simmons 2022).

Der *camera perspective bias* wird durch die begleitende Narration der filmschaffenden Polizistin 1 verstärkt, indem sie das Interaktionsgeschehen und individuelle Handlungen kommentiert und bewertet. Ein Effekt, der auch in anderen Bodycamaufnahmen zu beobachten ist (vgl. auch Williams et al. 2016). Besonders deutlich wird die Bedeutung einer solchen narrativen Begleitung des polizeilichen Filmens bei den von der Polizistin 1 formulierten Vorwürfen des Schlagens und Beißens von B.: So kann das vorhandene Videomaterial diese Handlungen nicht zweifelsfrei belegen, auch wenn deren Ausführungen zumindest im Ansatz zu erkennen sind. Stattdessen geben die narrativen Elemente der filmenden Polizistin 1 hier den interpretativen Rahmen vor, was unter dem im Video zu Sehenden zu verstehen ist, indem sie das Nicht-Sichtbare ergänzen und das Visuelle so mit spezifischen Deutungsangeboten verknüpfen. Der *camera perspective bias* wird auf diese Weise um einen *narrative perspective bias* ergänzt.

Dementsprechend ist der eingeschränkten Kameraperspektive von Bodycams mithin die Dimension des eingeschränkten auditiven Ausschnitts hinzuzufügen. Denn ein Großteil der gesprochenen Worte sind im analysierten Fall gar nicht zu verstehen, mit Ausnahme der Rufe von B. und den meisten Worten der Polizistin 1, deren Nähe zum Aufnahmegerät ihre Worte wesentlich verständlicher machen als die der anderen Beteiligten, auch Zwiegespräche zwischen den Polizist:innen sind beispielsweise zumeist sehr schlecht zu verstehen. Gerade im *narrative perspective bias* liegt für retroperspektive Auseinandersetzungen vor Gericht das Potenzial interpretativer Einflussnahme durch Polizist:innen nahe: indem sich die, durch Richter:innen ohnehin schon häufig privilegiert behandelten Aussagen der Polizist:innen – als »Berufszeug:innen« (Theune 2020) – und deren selektiv angefertigten Bodycam-Aufnahmen miteinander verweben und so (vermeintlich) objektivieren.

Während sich die vorstehenden Befunde durchaus im Einklang mit der Diagnose von Arzt und Schmidt (2021, S. 329) lesen lassen, nach der Bodycams durch eine »polizeiliche Perspektive« geprägt sind und diese reproduzieren, verdeutlichen unsere Ergebnisse ebenfalls die Notwendigkeit, diese polizeiliche Perspektivität genauer zu bestimmen. Denn wie insbesondere in der Analyse der zweiten Sequenz deutlich wird, geben Bodycamaufnahmen trotz ihrer technisch induzierten Perspektivität (gefilmt aus Sicht der Polizist:innenoberkörper mit einem spezifischen Bildausschnitt) keineswegs eindeutige Deutungshorizonte vor, sondern es bedarf konkreter ergänzender Filmhandlungen (Schnitt, Perspektivität, Narration), um eine situative Perspektivität der Bodycam zu produzieren. Hierbei ist auffallend, dass gerade durch die besondere visuelle Unschärfe (Bildqualität, Perspektive, Dynamik etc.) der technischen Dimension eine besondere Erklärungsbedürftigkeit der Aufnahmen besteht, welche den filmischen Handlungen der Polizist:innen in der situativen Produktion von Perspektivität umso größere

Relevanz zukommen lässt. Diese situativen Dimensionen polizeilicher Perspektivität sind wiederum von einer übergeordneten Perspektive zu differenzieren, wie Polizei auf das Thema Bodycam blickt und wie sie diese einzusetzen gedenkt, was wiederum wesentlich bestimmt, was auf Bodycamaufnahmen am Ende zu sehen ist und zu erkennen geglaubt wird. Aus diesem Grund schlagen wir zur Analyse polizeilicher Perspektivität in Bodycamaufnahmen eine (vorläufige) Differenzierung zwischen einer *technischen Dimension*, einer *filmhandlungsbezogenen Dimension* und einer *diskursiven Dimension* der polizeilichen Perspektivität vor. Damit, so unsere These, können die spezifische Bodycam-Perspektive und mithin ihre nicht-neutrale Art des Sehens und Zeigens, ebenso wie empirisch beobachtbare Ambivalenzen innerhalb und zwischen den Dimensionen, analytisch hinreichend berücksichtigt werden.

6. Methodologische Reflexion

Welche Besonderheiten des Datenmaterials Bodycam sind im Sinne einer methodologischen Reflexion relevant? Als vorrangig relevante Dimension der Begrenzungen von Bodycamaufnahmen ist der Zugang zu entsprechendem Material zu nennen. Nicht nur sind Polizist:innen die Produzent:innen entsprechender Aufnahmen, sondern auch die Speicherung, Verwaltung und Löschung entsprechender Daten liegt in der organisationalen Verantwortung von Polizeibehörden. Videointeraktionsanalysen von Bodycamaufnahmen haben somit immer auch Elemente einer Videoproduktanalyse zu berücksichtigen (siehe Abschnitt 5.2), da zumeist das auf den Aufnahmen sichtbare Interaktions- und Kommunikationshandeln nicht allein relevant ist, sondern auch Fragen nach den technischen und institutionellen Vorbedingungen der Aufnahmen und deren Qualität von Wichtigkeit sind. Bodycamaufnahmen sind letztlich auch deshalb als besonderes Format für die soziologische Videoanalyse zu verstehen, da sie im Vergleich zu Hochzeitsvideos, Werbeclips, aber auch über soziale Medien geteilte Smartphonevideos, die filmende Person weniger davon abhalten, aktiv ins Interaktionsgeschehen einzugreifen, da ihre Funktion als Filmer:in vergleichsweise wenig aktives Tun benötigt – z. B., weil die Kamera nicht eigens festgehalten werden muss. Ob und inwieweit dies methodologische Konsequenzen haben sollte, gilt es noch zu prüfen (s. u.).

Neben Fragen zu potenziellem Missbrauch und Möglichkeiten der Prüfung wirft die alleinige Zugangskontrolle der Polizei zu den Aufnahmen auch Fragen hinsichtlich des Zugriffs für potenzielle Beschwerdeträger:innen und Forscher:innen auf. Denn auch wenn die Implementierung von Bodycams in deutschen Polizeibehörden weit fortgeschritten ist, gibt es keine institutionalisierte Zugangsmöglichkeit jenseits individueller Anfrageersuchen (Aden/Fährmann 2019; Arzt/Schmidt 2021). In jedem Fall stehen bei entsprechenden Videoaufnahmen

datenschutzrechtliche Bedenken den öffentlichen bzw. forschungsbezogenen Interessen gegenüber.

Eine methodologische und forschungsethische Besonderheit nimmt hier die Form des *Leaks* ein, welche die Verfügbarkeit des diesem Artikel zugrundeliegenden Materials und seiner Analyse ermöglicht. Aus methodologischer Sicht gehen mit Datenmaterial aus einem Leak weitere Probleme einher, da Forschenden nicht nur die Kontrolle über die Datenproduktion fehlt, sondern mindestens eine weitere Partei potenziell unentdeckbare Veränderungen am Material vornehmen kann, wodurch der Inhalt verzerrt wiedergegeben und falsche Schlussfolgerungen getroffen werden können. Hinsichtlich des diesem Artikel zugrundeliegenden (geleakten) Materials betrifft dies etwa die angesprochene Ungewissheit über das Ende des Videos (hat Polizistin 1 die Aufnahme gestoppt oder wurde das Video nachträglich gekürzt?) sowie die audiovisuelle Anonymisierung von allen beteiligten Personen. Insbesondere das vom Veröffentlichungsmedium veranlasste Verpixeln von Gesichtern schränkt die ohnehin eingeschränkte Erkennbarkeit des Interaktions- und Kommunikationsgeschehen im Video weiter ein. Etwaig nachträglich eingefügte Schnitte lassen sich immerhin durch wiederholte Betrachtung weitgehend ausschließen.

In Zusammenhang mit denkbaren Veränderungen am Material stellt sich aus forschungsethischer Sicht die Frage, ob die im Video sichtbaren Personen ihre Zustimmung zur Veröffentlichung gegeben haben – zumeist ist dies bei polizeilichen Bodycams als nicht gegeben vorauszusetzen. Zumindest im vorliegenden Video ist die Zustimmung von dem Betroffenen B. anzunehmen, da er selbst (wiederholt) mit dem Veröffentlichungsmedium gesprochen und augenscheinlich eine Löschung des Videos nicht verlangt hat.

7. Fazit

In diesem Beitrag haben wir polizeiliche Bodycams als videoanalytischen Gegenstand eingeführt. Dies bedeutet einerseits die Dimension der Visualität von Bodycams stärker in den analytischen Blick zu nehmen, als es die einschlägige deutschsprachige Forschung in der Polizeisoziologie bzw. Kriminologie bislang getan hat. Andererseits war es ebenso Ziel des Beitrags, polizeiliche Bodycamaufnahmen als ein für die soziologische Videoanalyse spezielles Datenformat zu präsentieren, das quer zu gängigen Kategorisierungen der Videoanalyse, insbesondere Videointeraktions- vs. Videoproduktanalyse, liegt.

Bei Bodycamaufnahmen haben wir es mit audiovisuellen Produkten zu tun, die zwar natürliche Situationen aufnehmen, dafür aber eine vergleichsweise geringe situative Präsenz umfassen, da die Kameras z. B. nicht extra festgehalten werden müssen. Sie ähneln somit den Aufnahmen, wie sie typischerweise in der Videointeraktionsanalyse gebraucht werden. Allerdings werden sie im Gegensatz

dazu nicht von den Forschenden selbst produziert, sondern von den jeweiligen Polizist:innen. Damit sind die Bodycamaufnahmen grundsätzlich auch als geeignete Gegenstände einer Medienproduktanalyse zu begreifen, die aber typischerweise nicht im Rahmen von Videointeraktionsanalyse getätigt werden. Daher scheint es erstrebenswert, die analytischen Potenziale einer solchen Kombination in zukünftigen Untersuchungen von Bodycamaufnahmen weiter auszuloten, auch und gerade, wenn die Videointeraktionsanalyse mit ethnografischen Feldaufenthalten ergänzt und somit zu einer Videografie gemacht wird. Die Integration der wissenssoziologisch-hermeneutischen Videoanalyse, so haben wir vorgeschlagen, bietet sich dabei insbesondere mit Blick auf Bodycams an, da sie in sozialtheoretischer Kontinuität mit der Videointeraktionsanalyse sowohl die Handlungen vor der Kamera als auch die Handlungen (mit) der Kamera in den Fokus der Analyse zu rücken vermag.

Die Forschung zu Bodycams in Deutschland hat sich dabei bislang noch nicht ausreichend mit den visualitätsbezogenen Charakteristika dieser von Polizist:innen körpernah getragenen Kameras auseinandergesetzt. Daher haben wir anhand eines konkreten Beispiels, des bis dato einzigen öffentlich zugänglichen Bodycamvideos eines realen Polizeieinsatzes aus Deutschland, aufgezeigt, welche Resultate eine empirisch-videoanalytische Untersuchung einer solchen Aufnahme generieren kann und wie dies mit dem existierenden Forschungsstand zu verbinden ist.

Unsere Analyse hat gezeigt, wie Polizist:innen als Filmschaffende in der Kombination aus Perspektive, Schnitt und Narration durch Bodycamaufnahmen produktiv auf die polizeiliche Deutungshoheit einwirken (können) und so proaktive Visualitäts- und Narrativbildung betreiben. Der videoanalytische Zugang ermöglicht es hierbei, diese filmhandlungsbezogene Dimension für die Produktion polizeilicher Perspektivität zu fokussieren und so bestehende Untersuchungen zur polizeilichen Deutungs- und Diskursmacht im Rahmen neuer polizeilicher Sichtbarkeit um Befunde zu den Produktionspraktiken zu ergänzen. Für weitere Untersuchungen erscheint eine weitergehende Beforschung der Wechselwirkungen zwischen der technischen, der filmhandlungsbezogenen und der diskursiven Dimension polizeilicher Perspektivität besonders angezeigt, um die diskursive Bedeutung von Bodycamaufnahmen und die dahinterliegenden Herstellungsprozesse empirisch weitergehend zu rekonstruieren.

Darüberhinausgehend konnten wir verdeutlichen, dass hinter der Einführung von Bodycams Annahmen der präventiven Abschreckung von Gewalt gegenüber Polizist:innen stehen, die sich in realen Polizeieinsätzen mitunter nicht wiederfinden lassen – insbesondere mit Blick auf die Hypothese der rationalen Risikokalkulation von potenziellen Angreifer:innen. Es wurde aber ebenso deutlich, dass die Kamera allein in dieser Hinsicht wenig ausrichten kann, wenn die Beamt:innen nicht auch anderweitig auf deeskalative Instrumentarien zurückgreifen. Stattdessen erhärtet unsere Analyse den Verdacht, dass Bodycams in

sich zuspitzenden Situationen ebenso sehr auf deren weitere Eskalation wie deren Entschärfung hinwirken können.

Angedeutet wurde in diesem Beitrag letztlich das enorme Potenzial von Videoanalysen im Bereich polizeilicher Bodycams. Einerseits steht methodisch im Raum, die videointeraktions- und videoproduktanalytischen Herangehensweisen stärker miteinander zu integrieren, um den hybriden Charakter von polizeilichen Bodycamaufnahmen gerecht werden zu können. Andererseits gilt es noch gezielter zu analysieren, wie die Anwesenheit der Kameras Polizei-Bürger:innen-Interaktionen in ihren Verläufen ko-produziert, welche Rolle solche Aufnahmen in womöglich daran anschließenden Gerichtsprozessen einnehmen und wie polizeiliche Vertreter:innen auf sichtbare Verfehlungen von Beamt:innen (öffentlich) reagieren.

Literatur

Aden, Hartmut/Fährmann, Jan (2019): Bodycams bei der Polizei – nicht nur zum Schutz von Polizistinnen und Polizisten! Verfassungsblog, 02.03.2019. https://verfassungsblog.de/bodycams-bei-der-polizei-nicht-nur-zum-schutz-von-polizistinnen-und-polizisten (15.10.2024).

Arzt, Clemens/Schmidt, Stefanie (2021): Bodycam als Objekt staatlichen Sehens und Zeigens – Wie das Recht den polizeilichen Bedürfnissen angepasst wird. In: Singelnstein, Tobias/Ruch, Andreas (Hrsg.): Auf neuen Wegen. Kriminologie, Kriminalpolitik und Polizeiwissenschaft aus interdisziplinärer Perspektive: Festschrift für Thomas Feltes zum 70. Geburtstag. Berlin: Duncker & Humblot, S. 319–331.

Berger, Tom/Egbert, Simon (2025): Bodycams und polizeiliche Fehlerkultur. In: Seidensticker, Kai (Hrsg.): Fehlerkultur in der Polizei. Wiesbaden: Springer VS.

Berger, Tom/Egbert, Simon/Janssen, Jasper (i. E.): Einsatz und Regulierung von Bodycams bei der Polizei in Deutschland: Eine Bestandsaufnahme. Working Paper: Universität Bielefeld.

Bohnsack, Ralf (2011): Qualitative Bild- und Videointerpretation: die dokumentarische Methode. 2., durchgesehene und aktualisierte Auflage. Opladen, Farmington Hills: Barbara Budrich.

Brenner, Jan (2021): Schutz für Einsatzkräfte oder Datenschutzproblem? In: dbb magazin 72, H. 5, S. 20–21.

Browning, Lana M./Merlino, Mara/Sharp, Johnathon (2021): Citizen Journalism and Public Cynicism toward Police in the USA. In: Journal of Police and Criminal Psychology 36, H. 3, S. 372–385.

Englert, Carina Jasmin (2013): Ein Schwenk sagt mehr als tausend Worte – Die sich aus der Praxis des Zeigens entfaltenden Deutungsangebote der Kamera. In: Lucht, Petra/Schmidt, Lisa-Marian/Tuma, René (Hrsg.): Visuelles Wissen und Bilder des Sozialen. Wiesbaden: Springer VS, S. 71–86.

Goldsmith, Andrew J. (2010): Policing's New Visibility. In: British Journal of Criminology 50, H. 5, S. 914–934.

Graßl, Michael (2023): Polizeikommunikation auf Social Media: Ziele, Strategien, Inhalte. Wiesbaden: Springer VS.

Herrmann, Luis-Miguel (2022): »Ich sehe dich!« – Polizieren im digitalen Blick von Bodycams. In: Thüne, Martin/Klaas, Kathrin/Feltes, Thomas (Hrsg.): Digitale Polizei. Einsatzfelder, Potenziale, Grenzen und Missstände. Frankfurt am Main: Verlag für Polizeiwissenschaft, S. 281–298.

Kersting, Stefan/Naplava, Thomas/Reutemann, Michael/Heil, Marie/Scheer-Vesper, Carola (2019): Die deeskalierende Wirkung von Bodycams im Wachdienst der Polizei Nordrhein-Westfalen: Abschlussbericht. Gelsenkirchen: Institut für Polizei- und Kriminalwissenschaft der Fachhochschule für öffentliche Verwaltung NRW. https://www.hspv.nrw.de/fileadmin/Newsletter/2018_05_Mai/Zwischenbericht_OnlineVersion.pdf (15.10.2024).

Kipker, Dennis-Kenji/Gärtner, Hauke (2015): Verfassungsrechtliche Anforderungen an den Einsatz polizeilicher »Body-Cams«. In: Neue Juristische Wochenschrift 68, H. 5, S. 296–301.

Knoblauch, Hubert/Vollmer, Theresa (2018): Soziologische Videographie: Fokussierte Ethnographie und Sampling. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 121–134.
Latour, Bruno (2022): Die Hoffnung der Pandora: Untersuchungen zur Wirklichkeit der Wissenschaft. 7. Auflage. Frankfurt am Main: Suhrkamp.
Lehmann, Lena (2017a): Die Erprobung von Bodycams bei der Polizei. Unterschiede in den Vereinigten Staaten, Österreich und Deutschland. In: SIAK-Journal 14, H. 2, S. 28–38.
Lehmann, Lena (2017b): Wen fokussiert das dritte Auge? Der Einsatz von Bodycams bei der Polizei. In: Liebl, Karlhans (Hrsg.): Empirische Polizeiforschung XX: Polizei und Minderheiten. Frankfurt am Main: Verlag für Polizeiwissenschaft, S. 175–195.
Mannich, Julia (2024): Polizeiliche Nutzung von Bodycams. Erwartungen und Erfahrungen im In- und Ausland. In: Neue Kriminalpolitik 36, H. 3, S. 393–414.
Martini, Mario/Nink, David/Wenzel, Michael (2016): Bodycams zwischen Bodyguard und Big Brother. In: Neue Zeitschrift für Verwaltungsrecht – Extra 35, H. 24, S. 1–18.
Mawby, Rob C. (2002): Policing Images: Policing Communication and Legitimacy. Collumpton: Willan.
Meier zu Verl, Christian/Meyer, Christian/Grauert, Hanna (2024): Hermeneutiken des Verdachts: Institutionelle Diskriminierungspraktiken und Rassismus in öffentlichen und privaten Räumen der post-migrantischen Gesellschaft. Eine videoanalytische Untersuchung. In: Middell, Matthias (Hrsg.): Varianzen des Zusammenhalts. Frankfurt am Main: Campus, S. 391–422.
Newell, Bryce C. (2019): Context, Visibility, and Control: Police Work and the Contested Objectivity of Bystander Video. In: New Media & Society 21, H. 1, S. 60–76.
Probandt, Melanie (2020): Prügelvideos und Hetze: Die neue Inszenierung der Gewalt gegen Polizisten auf Social Media. In: Focus Online, 24. Juni 2020. https://www.focus.de/politik/deutschland/social-media-als-mittel-gegen-die-polizei-pruegelvideos-und-hetze-die-neue-inszenierung-der-gewalt-gegen-polizisten-auf-social-media_id_12129060.html (27.09.2024).
Reichertz, Jo (2018): Die Kunstlehre der wissenssoziologisch-hermeneutischen Videointerpretation. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 101–119.
Reichertz, Jo/Englert, Carina J. (2021): Einführung in die qualitative Videoanalyse: Eine hermeneutisch-wissenssoziologische Fallanalyse. Wiesbaden: Springer VS.
Sandhu, Ajay/Simmons, Jonathan S. (2022): Police Officers as Filmmakers: The Cinematography of Body Worn Cameras. In: Policing and Society 33, H. 5, S. 593–603.
Schattauer, Göran (2021a): Krasses Urteil um Knie-Einsatz der Polizei: Strafe für Mann, der zu Boden gedrückt wird. In: Focus Online, 05.11.2021. https://www.focus.de/politik/prozess-am-amtsgericht-muenchen-krasses-urteil-um-knie-einsatz-der-polizei-strafe-fuer-mann-der-auf-foto-unten-liegt_id_24396747.html (27.09.2024).
Schattauer, Göran (2021b): Schockierendes Polizei-Video: Ermittler prüfen umstrittenen Einsatz in München. In: Focus Online https://www.focus.de/politik/deutschland/beamter-kniete-auf-hals-eines-mannes-nach-focus-online-bericht-staatsanwaltschaft-prueft-polizei-einsatz-in-muenchen_id_13467784.html (27.09.2024).
Schattauer, Göran (2021c): Schock-Video aus S-Bahnhof in München: Polizist kniet minutenlang auf Hals eines Mannes. In: Focus Online, 08.07.2021. https://www.focus.de/politik/deutschland/drama-in-muenchner-s-bahnhof-ich-hatte-todesangst-polizist-in-muenchen-kniet-minutenlang-auf-hals-eines-mannes_id_13458380.html (15.10.2024).
Schattauer, Göran (2022): Brutaler Knie-Einsatz an Bahnhof: Jetzt ermittelt die Staatsanwaltschaft doch gegen Bundespolizisten. In: Focus Online, 09.02.2022. https://www.focus.de/panorama/eskalierte-fahrkarten-kontrolle-nach-focus-online-bericht-jetzt-doch-ermittlungen-gegen-bundespolizisten-wegen-brutalen-knie-einsatzes_id_50914369.html (27.09.2024).
Stoughton, Seth W. (2018): Police Body-Worn Cameras. In: North Carolina Law Review 96, H. 5, S. 1363–1424.
Theune, Lukas (2020): Polizeibeamte als Berufszeugen in Strafverfahren. Baden-Baden: Nomos.
Tuma, René (2018a): Videoanalyse und Videographie. In: Geimer, Alexander/Heinze, Carsten/Winter, Rainer (Hrsg.): Handbuch Filmsoziologie. Wiesbaden: Springer VS, S. 1–18 (Living reference work entry). https://doi.org/10.1007/978-3-658-10947-9_93-1.

Tuma, René (2018b): Video-Interaktionsanalyse. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 423–444.
Tuma, René/Egbert, Simon/Reichertz, Jo (2022): Videobilder und die Mediatisierung des Polizierens: Praktiken, Aneignung und Folgen. In: Thüne, Martin/Klaas, Kathrin/Feltes, Thomas (Hrsg.): Digitale Polizei. Frankfurt am Main: Verlag für Polizeiwissenschaft, S. 263–280.
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie: Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.
Ullrich, Peter (2018): Videoüberwachung von Demonstrationen und die Definitionsmacht der Polizei: zwischen Objektivitätsfiktion und selektiver Sanktionierung. In: Österreichische Zeitschrift für Soziologie 43, H. 4, S. 323–346.
Williams, Timothy/Thomas, James/Jacoby, Samuel/Cave, Damien (2016): Police Body Cameras: What Do You See? In: The New York Times, 01.04.2016. https://www.nytimes.com/interactive/2016/04/01/us/police-bodycam-video.html (15.10.2024).
Zander, Jens (2016): Kein Schutz vor Angriffen durch Body-Cams im Polizeieinsatz? In: Monatsschrift für Kriminologie und Strafrechtsreform 99, H. 5, S. 379–384.

Polizeiliche Vernehmungen

Zum Wandel einer hochintensiven Kommunikationsform durch ihre audiovisuelle Aufzeichnung[1]

Mina Godarzani-Bakhtiari und René Tuma

Zusammenfassung: Der Beitrag untersucht die Einführung der Videodokumentation bei polizeilichen Vernehmungen in Deutschland. Diese erfolgt in einem spezifischen Wissensfeld, in dem sich verschiedene Expert:innen mit der Form der Vernehmung auseinandersetzen. Sie ist also reflexiv geprägt von Auseinandersetzungen um ihre Legitimität, Effektivität und Transparenz. Videos haben hier das Potenzial, die polizeiliche Arbeit systematisch zu verändern, weil sie eben eine besondere Objektivierung des Geschehens darstellen. Auf der Basis empirischer Feldforschung, die Beobachtungen und Interviews umfasst, analysieren wir die Bezugnahme auf rechtliche Rahmenbedingungen und die öffentliche Debatte. Dabei beziehen wir uns sowohl auf Transformationsprozesse der Vernehmungssituation selbst als auch auf deren trans-sequenzielle Einbettung und gesellschaftliche Deutungskämpfe sowie den akademischen Diskurs. Der Beitrag verdeutlicht damit die Herausforderungen der Untersuchung eines Feldes, in dem das Video nicht von außen kommt, sondern quasi emisch Teil einer vernakulären Videopraxis ist.

Schlüsselwörter: polizeiliche Vernehmung, Videodokumentation, mediatisierte Kommunikation, Trans-Sequentialität, polizeiliche Praxis

1. Einleitung

Polizeiliche Arbeit durchläuft derzeit einen Mediatisierungsprozess. Während die Digitalisierung der Arbeitsprozesse und (schriftliche) Dokumentationsformen vorangetrieben werden (vgl. Reichertz/Wilz 2016), findet über die Integration und Nutzbarmachung von Videotechnologien eine Umgestaltung situativer und kommunikativer Praxis statt. Entgegen der zeitlichen Vergänglichkeit des situativen Handelns, konservieren Videoaufzeichnungen (zeit-räumlich) begrenzt die temporale Ordnung und die Details von flüchtigen Abläufen.[2] Damit bieten

1 Für wertvolle Hinweise zum Text und Korrekturen bedanken wir uns bei den Herausgebern des Bandes sowie bei Nele Dörl, Simon Egbert, Jasper Janssen und Talia Tuana Yücel.

2 Für die entsprechende Diskussion in den Sozialwissenschaften siehe Bergmann (1985).

sie das Potenzial, verschiedenen Interessengruppen Einblick in den konkreten Vollzug polizeilicher Praktiken zu ermöglichen. Wir gehen davon aus, dass durch die zunehmende Nutzung von Videos (vgl. Egbert/Janssen i. d. B.) sowohl die Polizei selbst als auch Parteien außerhalb der Polizeiorganisation neue kommunikative Formen des Umgangs mit den Videos entwickeln, was wiederum grundsätzlich eine mögliche Transformation polizeilicher Arbeit nahelegt (vgl. auch Tuma et al. 2022).

2017 wurde in Deutschland der Einsatz der Videodokumentation bei Beschuldigtenvernehmungen durch § 136 Abs. 4 StPO eingeführt (faktisch in Kraft seit 2020). Diese vergleichsweise späte Einführung in Deutschland – z. B. Großbritannien führte sie bereits 1984 ein – ermöglicht es, die Adaptionen der polizeilichen Praxis sukzessive nachzuzeichnen. Während die objektivierten Ergebnisse – vor allem Protokolle und Transkripte – von Vernehmungen auch vor der Einführung der Videotechnologie darauf ausgelegt waren, von anderen (Richter:innen oder in seltenen Fällen der Öffentlichkeit) beurteilt zu werden, schafft die Kamera eine neue Sichtbarkeit (Goldsmith 2010) der Vernehmungssituation jenseits der textuellen Protokolle. Mit den in den Vernehmungsräumen aufgestellten Kameras materialisiert sich die Präsenz des imaginierten und nachträglich aktualisierten Dritten der Situation (vgl. auch Hee 2018). Vor dem Hintergrund dieser Überlegungen, aber auch aufgrund erster Berichte aus dem Feld und der dort geführten Aushandlungen, interessieren wir uns in diesem Beitrag für die Frage, ob eine neue Sensibilität der Beteiligten durch die Aufnahmen zu Veränderungen der kommunikativen Form von Vernehmungen führt. Ob es zu einer Transformation kommt oder ob sich das Beharrungsvermögen der etablierten Praxis durchsetzt, ist empirisch noch offen. Die Möglichkeiten der Analyse hängt stark von weiteren Fragen, insbesondere der Zugänglichkeit des Materials ab. Wir argumentieren, dass diese mediatisierte kommunikative Form der Vernehmung nicht losgelöst von den gesellschaftlichen Prozessen der außerpolizeilichen Verwertung audiovisueller Aufzeichnungen – etwa vor Gericht oder in den Medien, aber auch in der wissenschaftlichen Analyse – betrachtet werden kann, die international unter dem Stichwort Police-Oversight diskutiert werden. In diesem Beitrag konzentrieren wir uns auf die verschiedenen polizeilichen und nicht-polizeilichen Kontexte, die die Veränderungen der neuen kommunikativen Form der videodokumentierten Vernehmung prägen, und geben einen ersten Überblick über den Stand der Forschung. Dieser Beitrag will und kann keine detaillierte Analyse der Videos selbst bieten, sondern einen Überblick über die Auseinandersetzung mit den Daten im Feld, in der Wissenschaft und in der öffentlichen Debatte. Dies hat jedoch systematische Gründe, die wiederum reflexive Einblicke in eine vernakulare Videopraxis und das Feld der Polizei ermöglichen. Theoretisch interessiert hier also, wie sich eine sehr spezifische, von Expertenwissen geprägte, in einem geschlossenen Kontext institutionalisierte und hoch formalisierte Kommunikationsform – die Vernehmung – verändert, die immer auch reflexiv von

Auseinandersetzungen um ihre Legitimität, Effektivität und Transparenz geprägt ist. Dies geschieht in einem spezifischen Wissensfeld, in dem sich verschiedene Expert:innen mit der Form der Vernehmung auseinandersetzen. Die Vernehmung ist in Prozesse eingebunden, also Teil eines Arbeitsprozesses und der trans-sequenziellen Konstruktion von Ermittlungsverfahren. Videos verändern die Verkettung systematisch, weil sie eben eine besondere Objektivation des Geschehens darstellen.

Da sich bereits in der Gesetzgebung zur Einführung der Videodokumentation von Vernehmungen das Spannungsverhältnis zwischen dem Ziel der Verbesserung der Ermittlungspraxis und der neuen Sichtbarmachung polizeilicher Arbeit aufzeigen lässt, setzen wir hier an. Im darauffolgenden Abschnitt skizzieren wir die öffentliche Thematisierung am Beispiel der medialen Diskussion eines Vernehmungsvideofalls in Deutschland. Anschließend geben wir Einblicke in die emische Perspektive auf videodokumentierte Vernehmungen – Hoffnungen, Probleme und neue Bedarfe in der polizeilichen Praxis werden anhand von Beobachtungen bei der Polizei skizziert. Daran schließen wir einen Überblick über den internationalen Forschungsstand zu wissenschaftlichen Perspektiven auf videodokumentierte Vernehmungen an. Da forschungsethische und feldspezifische Herausforderungen in der wissenschaftlichen Auseinandersetzung mit (Video-)Vernehmungen immanent sind, widmen wir uns diesen im letzten Abschnitt des Beitrags.

2. Rechtlicher und politischer Diskurs um die Videodokumentation der Vernehmung in Deutschland

Hintergrund für die Einführung der Videodokumentation von Vernehmungen sind Diskussionen um (internationale) Justizskandale, bei denen gravierende Fehler im Justizsystem wie Fehlurteile, unfaire Prozesse oder andere Missstände zu ungerechtfertigten Verurteilungen geführt haben. Besonders »[f]alsche Geständnisse [die in Vernehmungen produziert werden] sind ein ernstes Problem in der Ermittlungspraxis und im Gerichtsverfahren« (Kroll 2014, S. 17). Dabei steht besonders zur Debatte, welchen Einfluss die Vernehmungstechnik auf ein Geständnis hat. Lassiter et al. (2006, S. 195) nehmen an, dass die Videodokumentation die nachträgliche Beurteilung der Freiwilligkeit und Zuverlässigkeit von Geständnissen verbessern kann. Kommt es zu Zwang oder Beeinflussung durch Polizeibeamt:innen, würden Betrachtende die Unrechtmäßigkeit oder Unzuverlässigkeit des Geständnisses erkennen. Bereits im Jahr 1995 wurde in Deutschland geregelt, dass *Zeug:innen*vernehmungen mit schutzbedürftigen Personen grundsätzlich aufzunehmen sind, wobei die Verwendung dieser Daten streng reguliert ist (§ 58a StPO). Betroffene sollen damit u. a. vor Mehrfachbefragungen bewahrt werden. Im Juni 2013 fand die Ausweitung des Gesetzes auf Opfer

sexuellen Missbrauchs statt. Die Videodokumentation bei *Beschuldigten*vernehmungen wurde hingegen Ende 2013 juristisch zulässig. Verpflichtend sind sie erst seit 2020, und zwar bei Ermittlungen, bei denen ein vorsätzlich begangenes Tötungsdelikt zugrunde gelegt wird oder bei Vernehmungen von Beschuldigten mit besonderer Schutzbedürftigkeit (Personen unter 18 Jahre oder Personen mit sichtbar eingeschränkten geistigen Fähigkeiten oder schwerwiegenden seelischen Störungen).

Die Videos dienen, so der Gesetzgeber, der »Verbesserung der Wahrheitsfindung« (Burhoff 2020, S. 5 ff.). Ihnen wird also zugesprochen, einen Beitrag zur Sachverhaltsaufklärung leisten zu können. Sie werden als »authentische« Aufzeichnungen des Geschehens eingeordnet (Deutscher Bundestag 2017). Ein wesentlicher Aspekt bei der Einführung der videodokumentierten Beschuldigtenvernehmung ist die Möglichkeit, dass die Videos als Teil der Ermittlungs- bzw. Gerichtsakte von verschiedenen Parteien im Nachhinein überprüft werden können (sofern ihnen Zugang gewährt wird). Sowohl die Schuldfrage bezüglich des Befragten als auch die Rechtmäßigkeit der Vernehmungspraktiken soll durch die neue Form der Dokumentation nachvollziehbar werden. Dabei wäre es über die Videos möglich, nicht nur das gesprochene und dokumentierte Wort, sondern auch das nonverbale kommunikative Handeln der Beteiligten zu betrachten (Tuma et al. 2022, S. 148 f.). Aus der Gesetzgebung wird somit bereits deutlich, dass Videotechnologie als audio-visuelles Aufzeichnungs-, Speicher- und Wiedergabemedium die Spannung zwischen Ermittlungsarbeit (Wahrheitsfindung) und Handlungsgeboten (rechtmäßige Vernehmung) in ein materialisiertes Objekt einschreibt. Die Einführung der Kamera in Vernehmungssituationen erlaubt sowohl eine Auseinandersetzung mit den Handlungen der vernommenen Person und mit den Handlungen der Polizei selbst. Die Vernommenen und die Polizeibeamt:innen, die Interaktion, in die sie eingebunden sind, und die Situation, die sie schaffen, sind dabei an die mögliche Beobachtung durch Dritte und damit an trans-sequentiell verknüpfte Situationen gebunden (vgl. Kolanowski et al. 2023 zum Forschungsprogramm der Trans-Sequentialität, vgl. Tuma 2023 für unseren spezifischen Theoriezugriff).

Wir stellen uns daher die Frage nach den Auswirkungen dieser rechtlichen und politischen und im Resultat technisch-mediatisierten Refiguration auf die Vernehmungspraxis. Da davon auszugehen ist, dass die Videodokumentation in Zukunft abseits der bisher noch begrenzten Anwendung, längst etablierte Verfahren wie Tonaufzeichnung und Echtzeit-Protokollierung – trotz allen Beharrungsvermögens etablierter organisationaler Strukturen und Praktiken – ergänzen oder möglicherweise ersetzen (Wickel 2020, S. 311) wird, wird diese Frage noch relevanter werden. Dafür steht auch der Koalitionsvertrag der Ampel-Regierung unter Bundeskanzler Olaf Scholz, in dem es hieß: »Wir machen Strafprozesse noch effektiver, schneller, moderner und praxistauglicher, ohne die Rechte der Beschuldigten und deren Verteidigung zu beschneiden. Vernehmungen und

Hauptverhandlung *müssen* in Bild und Ton aufgezeichnet werden« (Koalitionsvertrag 2021–2025, S. 85; eigene Hervorhebung).

3. Öffentliche Auseinandersetzungen mit videodokumentierten Vernehmungen

Das Thema der Videoaufzeichnung von Vernehmungen wird vor allem in Fachdiskussionen verhandelt, insbesondere im juristischen, psychologischen und polizeilich-fachlichen Bereich. Im Gegensatz zu audiovisuellen Technologien in der polizeilichen Praxis, die in der Öffentlichkeit sichtbar und wirksam werden, wie etwa Bodycams (Egbert/Janssen i. d. B.), bei Demonstrationen (Ullrich/Knopp 2018) oder dem Einsatz automatisierter Überwachungstechnologien (Hälterlein/Möllers 2016), finden die Aufzeichnungen von Vernehmungen meist in Räumlichkeiten der Polizei statt,[3] in dem i. d. R. nur die Beteiligten, eventuell ihre anwaltlichen Vertretenden und gegebenenfalls Dolmetscher:innen anwesend sind. Auch wenn es sich dabei nicht um eine geschlossene Hinterbühne handelt, denn die Interaktionen werden ja dokumentiert – schon lange mit Protokollen und Mitschriften und neuerdings eben mit Video –, so hat dennoch die Polizei die faktische Kontrolle über die Aufzeichnung. Der situativen Machtasymmetrie bei der Durchführung und vor allem auch bei der Interpretation der Aussagen durch eine Seite soll durch das Recht auf anwaltliche Begleitung bei der Vernehmung und auch durch die Formen der Dokumentationspflicht systematisch entgegengewirkt werden. Klassischerweise sind und bleiben schriftliche Protokolle zentraler Bezugspunkt und Objektivationen des Geschehens.[4] Bei den Textprotokollen bzw. Transkripten von Audioaufzeichnungen wird das Gesprochene in geschriebenen Text transformiert, wobei dieses Diktat oder die Transkription immer schon eine Transformation bedeutet (vgl. Komter 2019; Malsch et al. 2018).[5] Mit der Videoaufzeichnung ändert sich dies systematisch, denn die Videoaufzeichnung enthält mehr Details des Geschehens in seiner spezifischen Sequentialität (Bergmann 1985). Die Rolle der Videodokumentation von Vernehmungen wird vor allem in investigativen journalistischen Formaten thematisiert.[6]

3 Einen Sonderfall stellen hier Rekonstruktionen am Tatort dar.

4 Vor Gericht gibt es weiterhin eine mediale Hierarchie, nach der Schriftlichkeit und Mündlichkeit in den Verhandlungen ein Primat haben. Bilder bedürfen vor Gericht einer Expert:innendeutung (vgl. Schwartz 2009).

5 Wir gehen weiter unten im Beispiel auf das Diktat ein, welches diese Transformation exemplifiziert.

6 Der Unterschied zwischen der US-amerikanischen und deutschen Situation wird deutlich, wenn die Sendung *Last Week Tonight with John Oliver* (2022) das Problem falscher Geständnisse anhand breit verfügbarer, aber nicht in allen Staaten verpflichtender Videoaufnahmen aufgreift. Sie kritisiert die Reid-Methode (Kroll 2014), die auf fragwürdige

Das Rechercheformat STRG-F (2020) erhielt während des Gerichtsprozesses die Vernehmungsvideos von Stefan E., der im Juni 2019 zusammen mit einem Bekannten den Kasseler Regierungspräsidenten Walter Lübke erschossen hatte. Kurz nach der Tat gestand er die Tat in einer Vernehmung, die auf Video aufgezeichnet wurde. In weiteren Vernehmungen widerrief er sein Geständnis und revidierte seine Darstellung. Das Videomaterial wird von STRG-F exklusiv veröffentlicht (Abb. 1), die genaue Herkunft des Materials bleibt unklar. Die Journalist:innen, die die Videos in diesem Format präsentieren, gehen näher auf die Vernehmung und die dabei entstandenen, multiperspektivisch positionierten Videoaufnahmen ein.

Abbildung 1: Nachzeichnung der ausgestrahlten Videoaufzeichnung der ersten (oben) und zweiten (unten) Vernehmungen von Stefan E. (STRG-F 2020)

Ausschnitte aus den Videos werden gezeigt und zeitlich und inhaltlich kontextualisiert. Die unterschiedlichen Geständnisse von Stefan E. werden kontrastiert. Interessant ist, dass die Journalist:innen zentral das Problem thematisieren, ob das Video veröffentlicht werden darf. Sie führen sowohl rechtliche als auch moralische Einwände an. Rechtlich argumentieren sie, dass das Video einerseits bereits vor Gericht gezeigt wurde und damit öffentlich bekannt ist und andererseits ein Dokument der Zeitgeschichte (eines politischen Mordes) darstellt. Moralisch fragen sie sich, ob sie dem Täter als Rechtsextremisten eine Bühne bieten sollten. Sie

Deutungen nonverbaler Kommunikation setzt, und thematisiert stundenlange Verhöre, in denen Polizisten Lügen nutzen, um Geständnisse zu erzwingen – besonders bei weniger privilegierten Vernommenen. Die Sendung fordert eine vollständige Aufzeichnung aller Verhöre, um solche Methoden zu verhindern in allen US-Bundesstaaten. Dies verdeutlicht, wie kritische Journalist:innen als Kontrollinstanz wirken.

grenzen die Veröffentlichung jedoch von den Tatvideos terroristischer Anschläge ab, da sowohl die Tat im Geheimen geplant wurde als auch das Vernehmungsvideo kein Produkt des Täters, sondern Teil des Ermittlungsverfahrens ist. Die meta-reflexive Thematisierung der Videos, ihre Hervorhebungen als vertrauliche Daten, die ausnahmsweise »durchgesickert« sind, und die Kritik an der in den Videos sichtbaren polizeilichen Ermittlungspraxis verdeutlichen den entstehenden Diskurs um die Rolle von Vernehmungsvideos. Es wird deutlich, wie die Verfügbarkeit des Materials die Öffentlichkeit zur eigenen vernakularen Analyse einlädt.[7]

4. Entstehungskontexte: Hoffnungen, Probleme und Barrieren der Videodokumentation aus polizeilicher Sicht

Wenn man sich näher mit dem Feld der Vernehmungspraxis beschäftigt, wird deutlich, dass die Integration von Videodokumentation in Vernehmungen im polizeilichen Ermittlungsalltag noch in den Anfängen steckt und zahlreiche neue Handlungsprobleme für die Beamt:innen mit sich bringt. Die Hoffnungen und Herausforderungen durch videodokumentierte Vernehmungen zeigen sich besonders im Schulungskontext. Verschiedene Polizeihochschulen bieten für angehende oder bereits praktizierende Kriminalpolizist:innen Vernehmungsschulungen an (siehe Sticher/Schicht 2019 für einen Überblick). Wir haben eine Schulung für bereits praktizierende Beamt:innen, deren Plätze stark umkämpft sind,[8] in einem deutschen Bundesland teilnehmend beobachtet.[9]

In der dreitägigen Schulung standen Zeug:innenvernehmungen im Zentrum. Videodokumentation wurde dabei nur nebensächlich verhandelt. Gleichzeitig bildete die Auseinandersetzung mit der allgemeinen Dokumentationsanforderung ein zentrales Handlungsproblem, welches von verschiedenen Seiten immer wieder thematisiert und problematisiert wurde. Wie wir im Folgenden aufzeigen, wird die Integration technischer (audio oder audio-visueller) Dokumentation auf drei Ebenen verhandelt: erstens als relevant bezüglich der Anforderungen im Ermittlungsverfahren; zweitens als alltagspraktisches Problem der

7 Auf ein ähnliches Datum, das einer Tatortbegehung, greift etwa auch Forensic Architecture zurück, um den Mord von Halit Yozgat durch den NSU 2006 investigativ zu untersuchen. Forensic Architecture nutzte u. a. das Videomaterial als Grundlage für eine aufwändige Rekonstruktion und unternahm als öffentliche Dritte eigene Ermittlungen (vgl. hierzu Godarzani-Bakhtiari 2024).

8 Teilnehmende der Schulung berichteten, dass sie sich bereits seit (bis zu) acht Jahren auf die Teilnahme bewerben und nun das erste Mal einen Platz erhalten.

9 Polizeigesetze und Ausbildungsgestaltung liegen weitgehend in der Hand der Bundesländer, wodurch es zu entscheidenden Unterschieden kommt.

Polizeibeamt:innen und drittens im Hinblick auf die Organisationsstrukturen innerhalb der Institution.

Die Notwendigkeit der Beschäftigung mit audio- und audio-visueller Vernehmungsdokumentation wurde von den Teilnehmenden bereits in der Vorstellungsrunde thematisiert. Der Anforderung nachkommend, bei der Vorstellungsrunde darauf einzugehen, ob sie und wenn ja, wie sie Vernehmungen durchführen würden, antworteten viele der Teilnehmer:innen mit der Form der Dokumentation, die sie dabei anwenden. Es zeigte sich, dass von den Teilnehmenden keine einheitliche Vorgehensweise der Dokumentation durchgeführt wird. Die Anwendung unterschiedlicher Verfahren zur Dokumentation wurde der Forscherin in Gesprächen anschließend damit begründet, dass es für »kleine Delikte« weder die Notwendigkeit noch rechtliche Vorgaben zur detaillierten Dokumentation gebe. Die Tragweite und praktischen Spannungen, die sich daraus ergeben, werden aus einem Ausschnitt des Gedächtnisprotokolls der Forscherin deutlich:

> *Eine andere [Person] thematisiert, dass sie seit zwei Jahren Vernehmungen durchführt und immer noch nicht versteht, wie ein Aufnahmegerät funktioniert. In den weiteren Vorstellungen wird [i]mmer wieder […] davon gesprochen, dass ein ›Diktat‹ gemacht wird. Ich verstehe nicht, was das heißt. Später in der Pause erkundige ich mich und erfahre, dass ›Diktat‹ meint, dass man ohne Aufnahme vernimmt und am Ende, oder mittendrin das Diktiergerät zur Hand nimmt, um die Aussagen des Gegenübers zusammenfassend aufs Band zu sprechen. Man hört zu, merkt sich die Aussagen, notiert sich vielleicht dazu etwas und paraphrasiert dann das Gesagte in einigen wenigen Minuten aufs Band. Das, was der/die Polizist:in aufs Band spricht, wird später durch die Arbeit der Protokollanten zum Protokoll … Ich bin verwirrt und fassungslos. Also entstammt kein Wort des Protokolls dann von den Vernommenen?* (Gedächtnisprotokoll 52–59/60)[10]

Ausgehend von den Berichten der Schulungsteilnehmer:innen führt ein Dozierender die Integration von Videotechnologie in Vernehmungen normativ ein. Dabei thematisiert er die Anforderung einer wörtlichen Dokumentation und problematisiert die Praxis der Paraphrasierung der Zeug:innenaussagen durch Polizeibeamt:innen:

> *Wir diktieren nicht und wir schreiben nicht, weil [Video-]Aufzeichnung die beste Möglichkeit ist. Wie es derzeit ist [ohne Videoaufnahmen], ist es suboptimal. […] Ich werde niemals die Worte des Beschuldigten wiedergeben können* (Gedankenprotokoll 64–66).

10 Die Gedächtnisprotokolle sind von der Autorin des Aufsatzes verfasst (M. G.-B.). Diese durfte während der Teilnahme an der Schulung händische Notizen verfassen. Außerhalb des Schulungsortes wurden die Notizen digitalisiert und durch Erinnerungen ergänzt.

Eine praktizierende Polizistin, die seit einigen Jahren die in der Schulung gelehrte Vernehmungstechnik mit ihrer kriminalpolizeilichen Einheit einsetzt und aus praktischer Perspektive die Schulung pädagogisch begleitet, thematisiert die Einführung von Videotechnologie vor dem Hintergrund der nachträglichen Prüfung vor Gericht:

> *Man macht sich halt so unglaublich angreifbar [wenn man Vernehmungen nicht mit Video oder zumindest Audio aufnimmt]. Mir passiert das nicht mehr. Tut Euch selber den Gefallen [alles aufzunehmen], weil für das Verfahren ist das wirklich nicht förderlich [wenn nicht aufgenommen wurde]. [Pause] Es ist ja auch besser, weil dann müsst ihr auch nicht diktieren … weil da filtert man ja auch … man diktiert dann das, was für einen selbst wichtig ist. Da kommt es ja auch zu Missverständnissen, also unabsichtlich* (Gedächtnisprotokoll 724–729).

Die Lehrende konstruiert in diesem Zitat die Vernehmungssituation als eine, die von außen be- und verurteilt und im Zweifelsfall gegen die Polizeibeamt:innen gewendet werden kann. Damit ordnet die Polizeibeamtin die Vernehmungssituation als ein Legitimationsproblem der Polizei ein, das auf der trans-sequenziellen Organisation des Verfahrens beruht. Im Zentrum der Aussage steht nicht das rechtmäßige Handeln der Polizeibeamt:innen, dieses wird vorausgesetzt und jedem moralischen Zweifel enthoben. Mit dem Hinweis auf mögliche Missverständnisse räumt die Dozierende zwar die Möglichkeit von Fehlern ein, ordnet diese aber als Kommunikationsprobleme ein, die der komplexen Vernehmungssituation geschuldet sind. Die Kontrolle polizeilichen Handelns, die (u. a.) auch als Ziel in der gesetzlichen Grundlage der Videodokumentation verankert ist, wird umgedeutet in einen Prozess der Selbstermächtigung. »Schützenswert« sind die Polizeibeamt:innen, die sich gegen ungerechtfertigte Anschuldigungen zu verteidigen haben.

Über die *praktischen* Konsequenzen der videodokumentierten Vernehmung sprechen die Lehrenden widersprüchlich.[11] Einerseits wird deren Einfluss auf die Vernehmungssituation negiert, wie an der Aussage eines Dozierenden deutlich wird: »*Und außerdem die Technik, die wir einsetzen, verändert die Vernehmungstechnik [selbst] überhaupt nicht. Für uns ändert sich also überhaupt nichts durch Videos*« (Gedächtnisprotokoll 68–69). An anderer Stelle vergleicht derselbe Lehrende die Umgestaltung der Vernehmungspraxis durch die Videotechnologie mit dem Erlernen einer Sportart. So antwortet er auf die zuvor gemachte Aussage eines Teilnehmers, der erläuterte, dass für ihn der Polizeiberuf *lebenslanges »Lernen und Trainieren«* sei, folgendermaßen:

11 Hier unterscheiden sich sicherlich auch Perspektiven. In Gesprächen mit Polizist:innen, die vermehrt mit der Aufzeichnung arbeiten, wird deren Einführung als besonders relevant dargestellt.

L: Du hast was ganz Wichtiges angesprochen: Trainieren. […bezieht das auf den Einsatz von Audio/Videotechnologien in Vernehmungen]. Wenn man im Fernsehen jemanden sieht, der einen Salto macht, kann man das auch nicht von heute auf morgen … (Gedächtnisprotokoll 74–76).

Die Analogisierung mit dem Erlernen eines Saltos deutet auf die Notwendigkeit der Herstellung eines neuen, auch zu inkorporierenden Handlungswissens hin, was Übung voraussetzt. Die Vernehmungspraxis erscheint vor diesem Hintergrund als ein verkörperter, routinemäßiger Akt. Die Integration der Aufnahmegeräte wird als disruptives Objekt behandelt, welches den routinemäßigen Ablauf der Vernehmung stört. Die Notwendigkeit der Routineveränderungen ist an die Perspektive der Teilnehmer:innen der Schulung anschlussfähig. Immer wieder artikulieren die Teilnehmenden technische und kommunikative Unsicherheiten in der Gestaltung der audio(-visuellen) Dokumentation. Dafür spricht auch der folgende Moment: »*In der Pause erklärt eine Teilnehmerin einer anderen, wie das Diktiergerät funktioniert. Drei weitere Personen verfolgen das Gespräch aufmerksam und mit großen Augen*« (Gedächtnisprotokoll 276).

Dass sich das Handlungsproblem nicht nur auf die technische Bewerkstelligung der Ein- und Ausschaltung des Gerätes bezieht, wird immer wieder thematisiert. Teilnehmende diskutieren, dass die Positionierung des Geräts im Verhältnis zu den sprechenden Personen beachtet werden muss. Außerdem muss erreicht werden, dass Teilnehmer:innen der Vernehmungen nur nacheinander und nicht gleichzeitig sprechen, da sonst auf der Aufnahme nichts zu verstehen sei (Gedächtnisprotokoll 730–733). Besonders die Beschwerden der Schreibkräfte über die »*Qualität der Aufnahmen*« führen dazu, dass die Beamt:innen reflexiv das Verhältnis von Geräten, räumlicher Positionierung und Sprechkoordination überdenken.

Gleichzeitig knüpft die Auseinandersetzung mit den Schreibkräften auch an organisationsstrukturelle Verhältnisse an. Konsens besteht zwischen den Teilnehmer:innen darüber, dass es zu wenig Schreibkräfte gebe, um den Arbeitsaufwand, der durch die technischen Aufnahmen entstehe, zu bewältigen. Dies führe dazu, abwägen zu müssen, ob eine vollständige Aufnahme tatsächlich notwendig sei. So argumentiert eine Teilnehmerin: »*Wir müssen gucken, dass wir ökonomisch drangehen an die Sache [bezüglich der Länge der Vernehmungen], sonst kommen unsere Schreibkräfte nicht mit*« (Gedächtnisprotokoll 216–217). Diese Aussage knüpft an die Auseinandersetzung mit dem »soziologischen Ermessen« an, wie es Hunold (2015, S. 9) bezeichnet. Mit dem Begriff des soziologischen Ermessens, der von der ethnografischen Polizeiforschung geprägt wurde, wird auf die unzähligen Handlungsoptionen der Polizei verwiesen, die besonders dadurch gegeben sind, dass Rechtsnormen situativ unterbestimmt sind. Beamt:innen stehen deshalb häufig vor der Herausforderung, nach eigenem Ermessen zu beurteilen, welche Handlung in einer konkreten

Situation zu vollziehen ist. Der folgende Dialog macht deutlich, wie soziologisches Ermessen und strukturelle Bedingungen von den Teilnehmenden als zusammenhängend konstruiert wird:

> *TY: Also ich finde das (die komplette Vernehmung aufzunehmen) macht Sinn bei Sexual und Mord. Aber wenn ich jetzt noch bei allem alles mitlaufen lasse, dann komm ich auf 10 bis 12 Seiten. Da krieg ich was aufm Kopf.*
> *Forscherin: Von wem denn?*
> *T1: Von den Schreibkräften!* (Gedächtnisprotokoll 735–735)

Die Entscheidung über die Dokumentationsform der Vernehmung wird von der:dem Beamt:in zurückgeführt auf die Einordnung der Tatschwere des zu ermittelnden Falls. Gleichzeitig werden organisationale Verflechtungen in der Produktion der Dokumentation thematisiert und legitimierend herangezogen. Dass es dabei nicht nur um Zuständigkeiten, sondern auch um die Frage der Autorität in der Gestaltung der Dokumentation und damit letztlich auch inner-polizeiliche Ressourcenverteilung geht, zeigt eine Aussage der Dozierenden:

> *Und wenn die Schreibkräfte meckern […] ich führ die Diskussion seit 2019 […] aber ehrlicher Weise, die Schreibkräfte wissen nicht, was eine gute Vernehmung ist. Und wir müssen da Druck machen. Das kann jetzt nicht weitere 50 Jahre dauern. Da müssen wir Druck machen. Da muss das Land mehr Geld für die Schreibkräfte ausgeben, dass die hinterherkommen. Die sind ja auch billiger als wir. Wenn wir das alles diktieren und so, das dauert ja auch, nimmt Zeit und wir sind schließlich teurer als die Schreibkräfte* (Gedächtnisprotokoll 741–746).

Die Verfahrensschritte der Vernehmungsdokumentation sind also Teil organisationaler Aushandlungen, wobei sich zwischen den Akteur:innenperspektiven deutliche Unterschiede zeigen. Während die Dozierenden auf normative Vorstellungen der Vernehmungsgestaltung abzielen und die Barrieren auf die politische Gestaltung der Organisationsstruktur zurückführen, problematisieren die Teilnehmenden die praktischen Herausforderungen. Dies geht so weit, dass Teilnehmende, ausgehend von der Einordnung der Dokumentation innerhalb des Ermittlungsprozesses, die Notwendigkeit detaillierter Vernehmungsdokumentation als solche in Zweifel ziehen. Da letztlich die Staatsanwaltschaft entscheiden würde, wie in einem konkreten Fall vorzugehen sei, müsse man erst schauen, welche Relevanz die Form der Dokumentation für die Staatsanwaltschaften habe. Dazu ein Teilnehmer: »*[D]ie Staatsanwälte, die wollen sowieso nur den Endbericht lesen, denen sind die Protokolle oder Aufnahmen egal. Und die interessieren sich auch nur für Dinge, die vor Gericht verwendet werden können*« (Gedächtnisprotokoll 298–300). Ein anderer Teilnehmer erläutert seine Perspektive gegenüber der Forscherin in der Pause ausführlicher. Er bezeichnet die Forderung nach einer

detaillierten Dokumentation insgesamt als ›Wahnsinn‹, der sich ihm aus dem Ablauf der Ermittlungsarbeit nicht erschließt:

> *T: Der Dokumentationswahn nimmt immer mehr zu … Das ist Wahnsinn. Das macht auch keinen Sinn. Ist sowieso alles für die Tonne. Der Staatsanwalt liest sich das sowieso nicht durch. Der überfliegt höchstens einen Abschlussbericht. Was da alles produziert werden soll … liest eh keiner! Genauso mit Videos. Das ist so ein riesen Berg Arbeit für die Schreibkräfte, dafür dass das eh niemand liest. Die Videos werden sowieso nie wieder angeschaut. Das kommt in die Akte und fertig* (Gedächtnisprotokoll 149–153).

Der Beamte deutet die Relevanz der Dokumentation pragmatisch aus seiner alltäglichen Erfahrung mit der Verfahrensordnung. Fragen bezüglich der trans-organisationalen nachträglichen Herstellung von Transparenz spielen keine direkte Rolle. Hier zeigt sich eine Parallele zur normativen Einordnung der Dozierenden. Auch hier werden von dem Teilnehmer die Handlungen der Polizist:innen per se als normativen und moralischen Standards folgend angenommen.

Als weitere Beobachtungen liegen uns Einblicke in die Herausforderungen und wahrgenommenen Potenziale der Videodokumentation auf der Basis einiger erster explorativer Feldbesuche bei Polizeieinheiten und aus Gesprächen mit Polizist:innen vor, die seit einigen Jahren Pionierarbeit im Bereich der Videodokumentation leisten. Auch hier wird die Videodokumentation ambivalent gesehen, einerseits als Herausforderung für die etablierte Arbeitsorganisation und auch als Bedrohung für die Möglichkeit, Vernehmungen überhaupt durchführen zu können (Beamt:innen vermuten, Verdächtige verweigern zunehmend die Aussage), andererseits aber auch als Chance zur Verbesserung der eigenen Ermittlungsarbeit. Gerade bei Aufklärungsarbeit in komplizierten Fällen äußern Beamt:innen uns gegenüber die Hoffnung, durch eine genaue Analyse der Videos Ansatzpunkte für mögliche weitere Ermittlungen zu finden. Seien es Widersprüche in den Aussagen der Vernommenen, körperliche Anzeichen von Unsicherheit oder gar die Entlarvung von Lügen (was im Gespräch allerdings nicht explizit thematisiert wurde, vermutlich da auch der problematische Diskurs um die Unwissenschaftlichkeit der Lügenerkennung den Polizist:innen vertraut ist). Gerade aber die Möglichkeit, sich die Vernehmungen im Nachhinein gemeinsam mit den ermittelnden Kolleg:innen anzuschauen, wird als Chance gesehen, ebenso wie das Potenzial, solches Material für interne Fortbildung zu nutzen (was allerdings immer als rechtlich problematisch verstanden wird). Interessanterweise verfügen die Ermittler:innen an den Standorten, die über separate Räume für die Videodokumentation von Vernehmungen verfügen, zum Zeitpunkt unserer Interviews über eine etwas veraltete, aber für ihre Zwecke funktional vorbereitete Ausstattung,[12] d. h. entsprechend angeordnete Sitzgelegenheiten, ein Nebenraum

12 Es existieren Anbieter für hierfür optimierte Ausstattung, siehe z. B. Media-4-cast.de (o. J.)

für die Technik, eine Uhr im Bild, eine installierte Kamera mit direkter Aufzeichnung auf optische Datenträger (DVD), die als physische Objekte (nur bis zum Abschluss der Ermittlung) lokal in einem Schrank gelagert und der Akte beigefügt werden können. Der Umgang mit den DVDs und der Software, die es ermöglichen soll, die Aufnahmen mit Metadaten auf den DVDs zu versehen, ist allerdings noch einarbeitungsbedürftig, d. h. es bedarf Mitarbeiter:innen, die bereit sind, sich auf die Technik einzulassen, ein Problem, das sich in der Justiz vermutlich fortsetzen wird.[13] Hier werden die organisatorischen Herausforderungen einer systematischen Einführung der Technologie deutlich.

Betrachtet man nun konkrete Vernehmungen, so lässt sich auch an der Vernehmung selbst die Besonderheit der Situation verdeutlichen, die sich in die konkrete kommunikative Form einschreibt. Vernehmungen sind für die Betroffenen ohnehin eine enorme Drucksituation, insbesondere wenn es sich um Beschuldigtenvernehmungen bei Kapitaldelikten handelt. Die Befragten werden von den Beamten zu Beginn ausführlich über ihre Rechte belehrt, wozu auch die ausdrückliche Belehrung über die audiovisuelle Aufzeichnung gehört. Sie sind nicht verpflichtet, auszusagen, sondern können die Aussage verweigern bzw. sich anwaltlich vertreten lassen. Erklären sie sich dennoch zur Aussage bereit, so ist dies für die Polizei gerade zu Beginn der Vernehmung eine fragile Situation, da der Tatverdächtige bei der Belehrung jederzeit einen Rückzieher machen könnte. Eine Vernehmungssituation beinhaltet zu Beginn Abschnitte wie der im Folgenden dargestellte (die wir im Feld sichten konnten):

P.1: Okay. Bevor wir jetzt hier weitermachen, geht es hier auch noch einmal um diese videodokumentierte Vernehmung. Hier ist eine Kamera, nimmt auf. Da ist auch ein Diktiergerät, das zeichnet zusätzlich unser Gespräch, unsere Vernehmung auf, und das wird dann später transkribiert und auf DVD gebrannt und kommt dann mit ins Ermittlungsverfahren. So kann man immer nachvollziehen, was wir hier besprochen haben.
V.: Ja.
P.1: Ok. Sind Sie damit einverstanden, dass wir das in dieser Form hier durchführen?
V.: Ja.
P.1: Sind Sie auch damit einverstanden, wollen Sie grundsätzlich Angaben machen?
V.: Ja.
P.1: Möchten Sie, dass wir, ähm, für Sie noch einen Pflichtverteidiger bestellen?
V.: Nein, ich habe das bereits mit meinem Anwalt besprochen.
(…)

13 Auch wenn wir es hier nicht direkt bei Vernehmungen beobachtet haben, so ist doch eine Feldbeobachtung aus einem Fall bezeichnend, in dem eine Videoaufzeichnung von Überwachungskameras Gegenstand der Gerichtsverhandlung war. Diese wurde nach langer und umständlicher Vorbereitung der Technik durch die Richterin und einen Gerichtshelfer provisorisch zum Laufen gebracht, was sich in einem schief an die Wand des Gerichtssaals projizierten Bild und einer unbeholfenen Bedienung des Videos äußerte.

Formular mit schriftlicher Belehrung wird von P.1 an V zum Lesen übergeben. Ca. vier (!) Minuten nahezu bewegungsloses Verharren aller Beteiligten während der Beschuldigte das Formular liest.
(Gedächtnisprotokoll des Anfangs einer im Feld gesichteten videodokumentierten Vernehmung, P.1. Polizist 1, V: vernommene Person)

Gerade die Beobachtung, dass die beiden anwesenden Polizeibeamt:innen mehrere Minuten regungslos warten, bis die zu vernehmende Person das Formular gelesen hat, ist bezeichnend. Ob die Person subjektiv liest oder ob sie möglicherweise unter dem Druck (es handelt sich, wie gesagt, um einen schweren Vorwurf) dazu kaum in der Lage ist und nur auf das Blatt starrt, lässt sich anhand des Videos nicht sagen. Es wird jedoch deutlich, dass die Aktivität in der Situation, also der Handlungs-»Turn«, an sie übergeben wurde und trotz der ungewöhnlich langen Pause eben nicht unterbrochen wird (es wird nicht einmal ein Schluck aus der Kaffeetasse genommen). Auch in der beobachtenden Perspektive wird hier der Druck deutlich, der sich in der Situation kommunikativ – im Handeln der Beteiligten – aktualisiert. Wir beobachten hier, und das ließe sich aus der Videoaufzeichnung in einer Feinanalyse auch genauer rekonstruieren, dass die Präsenz der Videoaufzeichnung sich auch in subtilen körperlichen Bezugnahmen auf diese »mit« niederschlägt. Auch wenn der spezifische Einfluss der Videoaufzeichnung hier noch nicht eindeutig nachvollzogen werden kann, ist es doch plausibel anzunehmen, dass diese die Polizeibeamt:innen auch darin bestärkt, ihr »korrektes« Vorgehen sichtbar zu machen, also überdeutlich zu zeigen, dass sie die vernommene Person nicht ablenken oder die formale Situation verletzen. Und damit auch systematisch vor der Kamera zu artikulieren und zu zeigen, dass keine ihrer Handlungen als illegitime Ausübung von Druck interpretiert werden kann.

5. Analysekontexte: Wissenschaftliche Perspektiven

In den vorangegangenen Abschnitten wurde in den Diskurs um die Videoaufzeichnung und den Umgang mit und die Erwartungen an diese in Deutschland – von Pionierarbeiten abgesehen – relativ neue Anforderung an die audiovisuelle Dokumentation eingeführt. Gerade in Deutschland handelt es sich um ein unübersichtliches und für Außenstehende schwer zugängliches Wissensfeld, vor allem wenn es um interne Fortbildungen und konkretes Praxiswissen geht. Die Professionalisierung der Aus- und Fortbildung zur Vernehmung ist hier nach wie vor ein umkämpftes Desiderat, das durch die Pflicht zur Videodokumentation verschärft wird. Dem steht ein akademischer empirischer Fachdiskurs um Videodaten gegenüber, den wir im Folgenden thematisieren. Besonderes Augenmerk legen wir dabei auf Studien, die Anschlüsse an die sozialwissenschaftliche Videographie (Tuma et al. 2013) und die Konversationsanalyse bieten.

Es gibt, überwiegend im englischsprachigen Diskursraum, zunächst eine Vielzahl an Untersuchungen, die auf Grundlage der Videodaten den Zusammenhang von Vernehmungstechnik und Vernehmungserfolg untersuchen. Diese Untersuchungen, überwiegend anwendungsorientiert, bearbeiten feldspezifische Handlungsprobleme von Vernehmungen, wie etwa die Gefahr falscher Geständnisse oder die fehlende Bereitschaft der Kooperation der Vernommenen. Basierend auf einem Sampling von meist Dutzenden bis über hundert Vernehmungsvideos, wird in den überwiegend (sozial-)psychologischen Untersuchungen das Verhalten der Vernehmungsteilnehmer:innen kodiert – etwa um den Erfolg einer Vernehmungsmethode zu messen (King/Snook 2009; Kelly et al. 2021), den Zusammenhang von Vernehmungstechnik und Kooperation der Beschuldigten herauszuarbeiten (Kelly et al. 2015), die Effizienz von Fragetypen zu evaluieren (Kelly/Valencia 2021), die Rolle von Empathie herauszuarbeiten (Kepinska/Jakobson 2021; Baker-Eck/Bull 2022) oder Stressfaktoren in Vernehmungen nachzuvollziehen (Bélanger 2023). In diesen Studien werden die Vernehmungsvideos in Transkripte übersetzt und letztlich überwiegend basierend auf den Kodierungen Regressionsmodelle erstellt.

Andere Untersuchungen folgen eher dem interpretativen Paradigma, indem sie detailliert analysieren, wie polizeiliche Vernehmungen interaktiv hergestellt werden. In einigen der überwiegend ethnomethodologischen und konversationsanalytischen Arbeiten werden die Videos entweder selbst als Primärdaten verwendet, auf deren Grundlage erstellte Transkripte analysiert werden. May und Kolleg:innen (2020) untersuchen in verschiedenen polizeilichen und juristischen Handlungsfeldern, u. A. polizeiliche Vernehmungen, was durch das Stellen von Fragen interaktiv erreicht wird. Im Sinne der klassischen Konversationsanalyse analysieren sie Transkripte und arbeiten verschiedene Techniken heraus, mit denen ein »recording the facts of the legal story« (ebd., S. 28) vorgenommen wird. MacLeod (2020) untersucht Transkriptausschnitte von Zeug:innenvernehmungen in Vergewaltigungsermittlungen. Sie folgt dabei der Frage, wie polizeiliche Vernehmer:innen mit der Herausforderung umgehen, strafrechtlich relevante Aussagen zu produzieren, die im Nachhinein nicht unter Verdacht geraten, unfreiwillig zu sein. Svennevig et al. (2024) arbeiten in ihrer Untersuchung ebenfalls mit Videotranskripten. Sie analysieren die linguistischen und konversationellen Praktiken, mit denen norwegische Vernehmer:innen in Interviews mit Personen, deren Muttersprache nicht Norwegisch ist, die Rechte des Vernommenen verständlich machen. Diese Untersuchung ist an eine frühere Studie aus Deutschland anschlussfähig: Hee (2012) untersucht die Bewerkstelligung der Vernehmung mit nicht-deutschen Muttersprachler:innen. Dabei arbeitete sie nicht mit polizeilichen Videos als Grundlage ihrer Daten, sondern führt eine Ethnographie bei der Polizei durch. Da sie eine Videokamera bei der Polizei installiert, ähnelt ihr Datensammlungsprozess dem der Videographie. Andere konversationsanalytische Studien analysieren die Bedeutung

von Videos und Audios außerhalb ihrer Entstehungssituation im polizeilichen und justiziellen Kontext. Komter (2019) arbeitet zwar mit Audiodaten, jedoch ist ihre Arbeit einschlägig. Sie zeichnet die ›Karriere‹ von Beschuldigtenaussagen im niederländischen Justizsystem nach. Sie zeigt auf, inwiefern diese im Prozess transformiert werden: vom Gesprochenen (Audio), über das Geschriebene (Protokoll) zu re-produzieren Aussagen (Gericht). Die Untersuchung von Richardson et al. (2023) schließt direkt daran an. Die Autor:innen vergleichen ebenfalls konversationsanalytisch 29 digital (audio- oder video) dokumentierte Vernehmungen aus Großbritannien mit ihren polizeilichen Protokollen, mit dem Ziel, die sozialen Handlungen der Transformation von ›talk to text‹ nachzuvollziehen. Auch wenn die Autor:innen Fälle analysieren, die entweder mit Audio- oder Videotechnologie aufgenommen sind, gehen sie auf die Differenzen zwischen Dokumentationsformen nicht ein.

Andere Untersuchungen nutzen das Potenzial von Videodaten, indem sie die Analyse der Vernehmungssituation nicht auf die Analyse von Sprache begrenzen, sondern den körperlichen und räumlichen Vollzug von Interaktionen analysieren. Wegweisend ist die frühe Studie von Lebaron und Streek (1997), in der mit Rückgriff auf Videodaten von Vernehmungen die räumliche und interaktive Ordnung in Vernehmungen herausgearbeitet wird. Als Klassiker gilt auch die Analyse von Katz (2001), in der er anhand einer konversations- und emotionsanalytischen Analyse von zwei Videoaufnahmen eines Mordgeständnisses in den USA aufzeigt, wie in Emotionsarbeit eine verteidigende Aussage in ein Geständnis transformiert wird. Johnson (2020) argumentiert, anschließend an den »embodied turn« (Nevile 2015) (ebenfalls mit Bezug auf Lebaron/Streek), dass die linguistische Forschung entscheidende multimodale Aspekte der Bedeutungsherstellung übersieht, wenn Vernehmungen allein auf Grundlage von Transkripten analysiert werden. So schreibt sie: »A great deal is accomplished in pointing, beating, head nodding and shaking, and iconic gestures, with implications for all involved« (ebd., S. 292). Sie analysiert audio-visuelle Aufnahmen von Vernehmungen aus England und Wales, um zu verstehen, was in Interaktionen multi-modal hergestellt wird (wie es im Text heißt), denn »doing is integral to saying« (ebd., S. 293). Auch Monteoliva-García (2020) führt eine multimodale konversationsanalytische Untersuchung durch, die zwar nicht den Körper zentral setzt, dessen Relevanz in der Interaktion dennoch ernst nimmt. Sie fragt, wie Interpretationsregime in Vernehmungen verhandelt und über die Zeit hergestellt werden. Eine deutschsprachige Untersuchung (jedoch basierend auf einem Video aus dem US-Kontext), die ebenfalls das Visuelle des Datums fokussiert, ist die Arbeit von Meier zu Verl (2016). Er untersucht den praktisch-prozedural körperlichen Vollzug von Re-Enactments als Re-Aktualisierung eines vergangenen Gewaltereignisses. Indem er betrachtet, welche Rolle das Re-Enactment im weiteren Strafermittlungs- und Gerichtsprozess einnimmt, beschränkt sich seine Analyse nicht auf den polizeilichen Vernehmungskontext.

Der letzte Forschungsstrang, auf den hier Bezug genommen werden soll, und der direkt an die zuletzt vorgestellte Untersuchung anschließt, befasst sich konkret mit der Art und Weise, wie die Vernehmungsvideos als neues Datum im Feld nutzbar gemacht werden. Ein Teil dieser Arbeiten sind anwendungsorientiert ausgerichtet, ihre Erkenntnisse werden gleichsam evaluativ ins Feld zurückgespiegelt. In den (überwiegend) vergleichenden und teils experimentellen Untersuchungen wird sich z. B. mit der Frage beschäftigt, inwiefern sich die Medialität der Vernehmung (face-to-face; Videoaufnahme/Video-Interview; Telefoninterview) auf die Fähigkeit der Beamt:innen/Richter:innen/Geschworenen auswirkt, wahrheitsgetreue Urteile/Interpretationen, zu produzieren. Hartmann et al. (2004) untersuchen, inwiefern Polizist:innen Lügen besser identifizieren können, wenn sie nicht selbst vernehmen, sondern Videos von Vernehmungen betrachten. Hale und Kolleg:innen (2022) untersuchen ähnliches, nämlich die »accuracy of interpreting« (ebd., S. 222) indem sie die Bewältigung verschiedener Vernehmungssituationen (*face-to-face*-Interviews, Video-Call-Interviews und Telefoninterviews) experimentell vergleichen. Andere analysieren die Nutzbarmachung der Vernehmungsvideos im Gerichtssaal. Sie untersuchen, welche Rolle die Position der Kamera in der Vernehmungssituation für die Beurteilung des Beschuldigten im Gerichtssaal spielt, in dem das Vernehmungsvideo später abgespielt wird. Lassiter und Kollegen (u. a. 2006) waren die ersten, die den »camera perspective bias« nachgewiesen haben. Sie zeigen in ihrer experimentell-vergleichenden Studie, dass die Position der Kamera eine Auswirkung auf die Beurteilung des Beschuldigten hat. Blandón-Gitlin und Mindthoff (2018) führen Ergebnisse verschiedener Studien diesbezüglich zusammen und argumentieren, dass eine Fülle an Studien gezeigt hat, dass die Verurteilungsrate bei der Präsentation der Vernehmungsvideos höher ist als wenn keine Videos gezeigt werden, auch wenn die Vernehmung erzwungen wurde, oder die Gerichtsteilnehmer:innen aktiv dazu angehalten wurden, das Geständnis außer Acht zu lassen. Das zeigt, so die Autor:innen, dass Richter:innen und Geschworene nicht ausreichend sensibilisiert seien, polizeiliche Taktiken der Beeinflussung und des Zwangs in Vernehmungsvideos zu identifizieren.[14]

Die Forschung zu Vernehmungen und Vernehmungsvideos geht daher in verschiedene Richtungen: Zum einen geht es um die konkrete Evaluation von Vernehmungstechniken und Fragen des Einflusses der Videodokumentation auf die Wahrheitsfindung. Zum anderen besteht ein stärker sozial- und kommunikationswissenschaftliches Interesse an den Besonderheiten der kommunikativen Situation der Vernehmung. Die Körperlichkeit, aber auch die Materialität, die

14 Auch Lassiter et al. argumentieren, dass Techniken des offensichtlichen Unterdrucksetzens von Beschuldigten durch Beamt:innen durch nicht-angreifende psychologische Manipulation ersetzt worden ist, die nicht immer als Zwang erkannt wird, auch wenn sie zu falschen Schuldbekenntnissen führen kann (2006, S. 195).

Transformationen in andere Medien (z. B. die Protokolle), also die Ergebnisse der Situation, spielen im Anschluss an den embodied turn der Interaktionsforschung eine zunehmend zentrale Rolle. Sie sind für uns von besonderem Interesse, weil sie Einblicke in Verkettungen und Transsituativität geben. Allerdings ergeben sich hier eine Reihe von forschungspraktischen Komplikationen, wie wir im nächsten Abschnitt skizzieren.

6. Forschungspraktische Probleme und Herausforderung

Die Vernehmung, die stark formalisiert und institutionalisiert ist, hat klassischerweise einen starken Fokus auf sprachliche Kommunikation und weist eine starke Musterhaftigkeit im Sinne einer kommunikativen Gattung auf (vgl. Günthner/Knoblauch 1995). Vor allem internationale Studien, die sich, wie gezeigt, auf die raum-körperliche Anordnung, Emotionen, die Rolle nonverbaler Kommunikation und andere Aspekte der Gattung konzentrieren, in die die Subjekte oft essenziell und unter hohem Druck eingebunden sind, zeigen relevante Aspekte auf, die mit Videoanalysen auch in Deutschland weiter analysiert werden müssten. Die spezifische Institutionalisierung der formalisiert produzierten, stark rechtlich regulierten, vom Feld produzierten Videos, besteht auch in ihrer systematischen Einbettung in Verfahren, also in organisierte Arbeitsbögen. Durch das Video wird aus dem Flüchtigen etwas Aktenförmiges, damit hat diese Objektivation gerade das Potenzial, die trans-sequenzielle (Kolanowski et al. 2023) Konstitution von Untersuchungsverfahren systematisch zu verändern. Für die akademische Betrachtung des Phänomens bedeutet dies, die systematische Einbettung kommunikativer Formen und Gattungen in den Blick zu nehmen und die Situation videographisch – von außen wie von innen – und somit in ihrer reflexiven Thematisierung durch ein Wissensfeld zu erfassen.

Die Strukturen des Gegenstands und des umgebenden Wissensfeldes sind charakterisiert durch eine spezifische Nicht-Zugänglichkeit für Außenstehende. Da Vernehmungen auf der polizeilichen Hinterbühne stattfinden und der Zugang zu Aufzeichnungen für Akteur:innen außerhalb der Polizei oder der Justiz zumindest in Deutschland sehr selten ist, kann die Erforschung der Veränderungen durch die Integration von Videotechnologien bei der Polizei nur durch Feldkontakte und Feldaufenthalte gelingen. Gleichzeitig hat es die Forschung zur Polizei – besonders außerpolizeiliche, wie etwa universitäre Forschung – insgesamt schwer. Feldzugänge müssen mühsam erarbeitet werden und sind, selbst wenn etabliert, dauerhaft instabil. Reichertz sieht die Verantwortung auch bei der sozialwissenschaftlichen Forschung der 1960er und 1970er Jahre, welche der Polizei mit »empirieloser Verurteilung« begegnet sei (2003, S. 415). Ullrich hebt hervor, dass der Zweck und die grundlegende Handlungsrationalität der Polizei »Sicherstellung von Autorität, Kontrollhoheit und Dominanz« (2019, S. 165) ist.

Nicht steuerbare Einblicke in die Vollzugswirklichkeit ihrer Arbeit zu gewähren, ist nicht vereinbar mit der organisationalen Logik des Feldes (ebd., S. 159). Annäherungen durch Forschende bedrohen somit die organisationalen Prinzipien der Institution. Ihre strukturelle Außenposition, unabhängig von der inhaltlichen Perspektive, stellt die eigene Deutungshoheit infrage und wird deshalb innerhalb der Polizei als Sicherheitsfrage behandelt.[15] Gerade Vernehmungen zeichnen sich dadurch aus, dass ihre Durchführung auf einem reflexiven polizeilichen Sonderwissensbestand (Vernehmungstechnik und -strategie) basiert, welcher aufgrund des Ziels des strategischen Einsatzes als vertraulich behandelt wird. Dokumente dazu sind in der Institution teils als Verschlusssache klassifiziert. Soziolog:innen stehen außerdem vor der Herausforderung, dass die Position der wissenschaftlichen Expertise, was den Kontext Vernehmung angeht, weitestgehend bereits von der Psychologie besetzt ist. Zudem erfordert eine Vernehmungssituation von den Polizeibeamt:innen sensible emotionale Kontrolle, was für sie persönlich herausfordernd ist, wodurch sie kein Interesse an einer externen Betrachtung haben. Dass in Lehrkonzepten vorgeschlagen wird, darüber nachzudenken polizeiliche Supervisionen von Vernehmungen nicht im Team, sondern Einzeln vorzunehmen (Sticher/Schicht 2019, S. 64), zeigt auf, wie viel Unsicherheit mit der Begutachtung durch ein Außen bereits innerhalb der Institution besteht. Darüber hinaus – und dies ist von hoher Relevanz – haben auch die Vernommenen nicht unbedingt ein Interesse daran, die aus verschiedenen Gründen oft sehr belastende Vernehmungssituation durch zusätzliche Komplikationen wie beobachtende Wissenschaftler:innen zu erschweren. Die existierenden Schutzpflichten, die von der Polizei wahrzunehmen sind, führen außerdem dazu, dass gerade bei visuellen Daten mit dem Argument des Datenschutzes besonders vorsichtig umgegangen wird.[16]

Besteht ein Zugang – sei es klassisch ethnographisch oder insbesondere videographisch – eint alle Analysezugänge, dass sie mit vielfältigen analytischen Herausforderungen umzugehen haben, angefangen bei der zwingenden

15 Dies steht im engen Zusammenhang mit der Fehlerkultur des Feldes. Wie Seidensticker schreibt, zeichnet sich diese durch die Annahme »Fehler passieren nicht!« (2019) aus. Zur Einführung visueller Technologien (Bodycams) und polizeilicher Fehlerkultur, siehe Egbert/Berger (2025).

16 Besonders drastisch zeigt sich diese Debatte in der Diskussion um sexuelle und sexualisierte Gewalt. Hier wird diskutiert, ob Videoaufzeichnungen die Betroffenen von der wiederholten Schilderung traumatisierender Situationen entlasten können. Auch wenn es hier nicht um die Vernehmung, sondern um die Aufzeichnung der Tat geht, scheint die Diskussion um die öffentliche Vorführung der Vergewaltigung von Gisele P. durch eine Vielzahl von Männern, bei der die Klägerin dafür kämpfen musste, dass die Aufnahmen vor Gericht öffentlich gezeigt wurden, ein Beispiel zu sein. Ihr Argument für einen emanzipatorischen Paradigmenwechsel lautet: »Die Scham muss die Seite wechseln« (Tagesschau 2024). An diesem Beispiel lassen sich Verschiebungen im Diskurs um die Visualität von Gewalt auch in Prozessen beobachten.

Notwendigkeit von Anonymisierung im Forschungsprozess. Machtverhältnisse, die daraus in der Vernehmungssituation resultieren, gilt es analytisch konsequent zu beachten. Wie ein kontextsensibler analytischer Umgang auszusehen hat, ist gerade vor dem Hintergrund einer (bisher) weitestgehend unzureichenden Forschungslage offen. Hinzu kommt die Schwierigkeit, dass der Gegenstand selbst der Logik der Entschlüsselung eines Problems (Tat) folgt und damit implizit (auch wenn ganz anders gelagert) der wissenschaftlichen Logik ähnelt, jedoch bei weitem nicht gleicht. Anfragen seitens der Polizei, sozialwissenschaftliche Expertise für die Aufklärung von Ermittlungen einzubringen, sind mit äußerster Vorsicht zu begegnen. So kann und darf sich die soziologische Perspektive nicht für die Begründung polizeilicher und juristischer Schuldhaftigkeit instrumentalisieren lassen. Diese Differenzen zwischen Polizei- und Wissenschaftslogik gilt es unbedingt zu beachten und die Grenzen zwischen den Feldern sichtbar zu machen. Begeben sich akademische Forschende hier ins Feld, werden sie einerseits auch strukturell Teil der Situation (juristisch werden sie zu Zeug:innen des Verfahrens, mit allen Konsequenzen); gleichzeitig nehmen sie aber auch eine Positionalität gegenüber Parteien mit meist unterschiedlichen Interessen – in einem moralisch hochaufgeladenen Feld – ein. Gerade weil die Schuldfrage bzw. die Frage nach der Wahrhaftigkeit einer Aussage in Vernehmungen omnipräsent ist, wäre es ein Trugschluss, anzunehmen, dass deren Relevanz für die eigene Analyse (auch wenn dies nicht Teil der Fragestellung ist) durch eine distanzierte Betrachtung einfach entgangen werden kann. So hängt die Deutung der Interaktion maßgeblich von der Beurteilung des Hauptproblems ab (handelt es sich um glaubwürdige Aussagen durch die vernommene Person). Deshalb muss die Relevanz des Wahrheits- und Schuldverständnisses im Feld unbedingt analytisch reflektiert werden.

Da gerade in diesem Feld die direkte Analyse, d. h. die Videoanalyse der Daten selbst, nur in sehr ausgewählten und fokussierten Fällen möglich ist und insbesondere der Zugang und die Reflexion der internen Nutzungspraktiken mit Fragen des Zugangs zusammenfallen, lässt sich deutlich zeigen, dass in diesem Feld eine klassische Videographie zwar wünschenswert wäre, Erkenntnisse aber vor allem durch fokussierte Ethnographie, Expert:inneninterviews und andere angepasste Methoden sowie eine reflexive Erforschung der vernakularen Methoden, Einblicke in das Feld ermöglichen. Während Video vermeintlich als neue öffentliche Sichtbarkeit verstanden werden kann, zeigt sich hier, welche Unsichtbarkeiten systematisch mitproduziert werden.

Die geschilderten Probleme sind soziologisch und auch für die videoanalytische Forschung relevant, da Video eben nicht nur eine Forschungstechnologie, sondern eine Alltagstechnologie ist und die vernakulare Selbst- und Außenbeobachtung systematisch Teil des Feldes ist (vgl. Knoblauch/Wilke i. d. B.). Für die soziologische Forschung geht es dabei sowohl um die Integration und Weiterentwicklung von Methoden als auch um die Bewältigung von Herausforderungen

im Bereich der Archivierung, des (stark regulierten) Datenaustauschs und der Publikation. Dies sind Herausforderungen, denen wir uns in unserem Projekt »Visions of Policing« stellen, die aber sicherlich auch auf weitergehende Entwicklungen verweisen.

Literatur

Antaki, Charles/Richardson, Emma/Stokoe, Elizabeth/Willott, Sara (2015.): Dealing with the Distress of People with Intellectual Disabilities Reporting Sexual Assault and Rape. In: Discourse Studies 17(4), S. 415–32.

Baker-Eck, Bianca/Bull, Ray (2022): Effects of Empathy and Question Types on Suspects' Provision of Information in Investigative Interviews. In: International Journal of Police Science & Management 24(4), S. 406–16.

Bélanger, Philippe L./Fortin, Francis/Deslauriers-Varin, Nadine/Paquette, Sarah (2023): Exploring Interrogation-Related Stressors: Factors Influencing Apparent Stress in Investigative Interviews with Suspects. In: Canadian Journal of Criminology and Criminal Justice 65(3), S. 47–67.

Bergmann, Jörg (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. In: Bonß, Wolfgang/Hartmann, Heinz (Hrsg.): Entzauberte Wissenschaft. Soziale Welt, Sonderband 3. Schwartz, S. 299–320.

Blandón-Gitlin, Iris/Mindthoff, Amelia (2018): Do Video Recordings Help Jurors Recognize Coercive Influences in Interrogations? In: Najdowski, Cynthia/Stevenson, Margaret (Hrsg.): Criminal Juries in the 21 st Century. New York: Oxford University Press, S. 195–220.

Burhoff, Detlef (2020): Audiovisuelle Aufzeichnung von Beschuldigtenvernehmungen (§ 136 Abs. 4 StPO). StRR – StrafRechtsReport Ausgabe 1, S. 5–12.

Deutscher Bundestag (2017): Gesetzentwurf der Bundesregierung: Entwurf eines Gesetzes zur effektiveren und praxistauglichen Ausgestaltung des Strafverfahrens 18/11277. 18. Wahlperiode. Köln: Bundesanzeiger.

Egbert, Simon/Berger, Tom L. (2025): Polizeiliche Fehlerkultur und die Legitimierung des Einsatzes von Bodycams. In: Seidensticker, Kai (Hrsg.): Fehlerkultur in der Polizei. Wiesbaden: Springer.

Godarzani-Bakhtiari, Mina (i. E.): Die Analyse des Analysierens visueller Artefakte. Vergleichende-Videoartefakt-Analyse der Evidenzkonstruktionen von Forensic Architecture. In: FQS – Forum Qualitative Sozialforschung 26.

Goldsmith, Andrew J. (2010): Policing's New Visibility. In: British Journal of Criminology 50(5), S. 914–934.

Günthner, Susanne/Knoblauch, Hubert (1995): Culturally patterned speaking practices – the analysis of communicative genres. Pragmatics (5), S. 1–32.

Hale, Sandra/Goodman-Delahunty, Jane/Martschuk, Natalie/Lim, Julie (2022): Does Interpreter Location Make a Difference? A Study of Remote vs. Face-to-Face Interpreting in Simulated Police Interviews. Interpreting. In: International Journal of Research and Practice in Interpreting 24(2), S. 221–53.

Hälterlein, Jens/Möllers, Norma (2016): Deutungskonflikte um automatisierte Videoüberwachung. Zur sozialen Konstruktion einer Technologie als Instrument der Kriminalitätsbekämpfung. In: Zoche, Peter/Kaufmann, Stefan/Arnold, Harald (Hrsg.): Grenzenlose Sicherheit. Berlin: LIT, S. 163–180.

Hartmann, Tilo/Schramm, Holger/Klimmt, Christoph (2004): Personenorientierte Medienrezeption: Ein Zwei-Ebenen-Modell parasozialer Interaktionen. In: Publizistik 49(1), S. 25–47.

Hee, Katrin (2012): Polizeivernehmungen von Migranten: eine gesprächsanalytische Studie interkultureller Interaktionen in Institutionen. Heidelberg: Universitätsverlag Winter.

Hee, Katrin (2018): Das Aufzeichnungsmedium als Interaktant: Zur ›Invasivität‹ empirischer Forschung. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer, S. 365–83.

Howe, Christiane/Ostermeier, Lars (2019): Einleitung. In: Howe, Christiane/Ostermeier, Lars Ostermeier (Hrsg.): Polizei und Gesellschaft. Wiesbaden: Springer Fachmedien, S. 1–16.

Hunold, Daniela (2015): Polizei im Revier – Polizeiliche Handlungspraxis gegenüber Jugendlichen in der multiethnischen Stadt. Berlin: Duncker & Humblot.
Jakobsen, Kristina K. (2021): Empathy in Investigative Interviews of Victims: How to Understand It, How to Measure It, and How to Do It? In: Police Practice and Research 22(2), S. 1155–70
Johnson, Alison (2020): »Are You Saying You Were Stabbed … ?«: Multimodality, Embodied Action, and Dramatized Formulations in »Fixing« the Facts in Police Interviews with Suspects. In: Mason, Marianne/Rock, Frances (Hrsg.): The Discourse of Police Interviews. Chicago: University Press, S. 268–298
Katz, Jack (2001): How Emotions Work. Chicago, Ill.: University Press.
Kelly, Christopher/Meissner, Christian A. (2015): Interrogation and Investigative Interviewing in the United States: Research and Practice. In: Walsh, David/Oxburgh, Gavin/Redlich, Allison/Myklebust, Trond (Hrsg.): Contemporary developments and practices in investigative interviewing and interrogation: Volume II (suspects). London: Routledge.
Kelly, Christopher E./Meehan, Nathan/Mcclary, Michael/Jenaway, Elizabeth M. (2021): Just a Normal Conversation: Investigative Interviews in a County Jail. In: Criminal Justice and Behavior 48(8), S. 1166–1184.
Kelly, Christopher E./Valencia, Esteban J. (2021): You Ask and Do Not Receive, Because You Ask Wrongly. In: International Journal of Police Science & Management 23(1), S. 42–54.
King, Lesley/Brent Snook (2009): Peering Inside a Canadian Interrogation Room: An Examination of the Reid Model of Interrogation, Influence Tactics, and Coercive Strategies. In: Criminal Justice and Behavior 36(7), S. 674–94.
Kolanoski, Martina/Löffler, Marlen S./Küffner, Carla/Terjung, Clara (Hrsg.) (2023): Trans-Sequentiell Forschen: Neue Perspektiven und Anwendungsfelder. Wiesbaden: Springer VS.
Komter, Martha (2019): The Suspect's Statement: Talk and Text in the Criminal Process. UK: Cambridge University Press.
Kroll, Ottmar (2014): Wahre und falsche Geständnisse in Vernehmungen. In: SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis (2), S. 17–32.
Lassiter, G. Daniel/Ratcliff, Jennifer J./Ware, Lezlee J./Irvin, Clinton R. (2006): Videotaped Confessions: Panacea or Pandora's Box? In: Law & Policy 28(2), S. 192–210.
Last Week Tonight with John Oliver (2022): Police Interrogations. HBO www.youtube.com/watch?v=obCNQ0xksZ4 (10.11.2024).
Lebaron, Curtis D./Streeck, Jürgen (1997): Built Space and the Interactional Framing of Experience During a Murder Interrogation. In: Human Studies 20(1), S. 1–25.
MacLeod, Nicola (2020) »Tell Me in Your Own Words …«: Reconciling Institutional Salience and Witness-Compatible Language in Police Interviews with Women Reporting Rape. In: The Discourse of Police Interviews. Chicago University Press.
Malsch, Marijke/Kranendonk, P. Robin/Keijser, Jan de/Komter, Martha L./Boer, Meike de/Elffers, Henk (2018): Reporting on Police Interrogations: Selection effects and bias related to the use of text, video and audiotape. In: Investigative Interviewing: Research And Practice, 9(1), S. 61–76.
May, Alison/Holt, Elizabeth/Al Saeed, Neveen/Sani, Nurshafawati A. (2020): Legal talk. Socio-pragmatic aspects of legal questioning: police interviews, prosecutorial discourse and trial discourse. The Routledge handbook of forensic linguistics. 2. Auflage. London, New York: Routledge.
Media C 4 (o. J.): Vernehmungsraum, Video Vernehmungszimmer, Vernehmungsaufzeichnung. Multi Capture Version 4 www.media-4-cast.de/vernehmungsaufzeichnung/vernehmung (15.11.2024).
Meier zu Verl, Christian (2016): (Re)Konstruktionen des Tathergangs: Reenactments als epistemische Körperpraktiken der Strafverfolgung und -verhandlung. In: Dreschke, Anja/Huynh, Ilham/Knipp, Raphaela/Sittler, David (Hrsg.): Reenactments. Bielefeld: transcript, S. 297–326.
Monteoliva-Garcia, Eloisa (2020): Interpreting or other forms of language support? Experiences and decision-making among response and community police officers in Scotland. In: The International Journal of Translation and Interpreting Research 12(1), S. 37–54.
Nevile, Maurice (2015): The Embodied Turn in Research on Language and Social Interaction. In: Research on Language and Social Interaction 48(2), S. 121–51.
Reichertz, Jo/Wilz, Sylvia M. (2016): Polizeiliche Aufklärungsarbeit 2.0. In: SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis (1), S. 31–39.
Reichertz, Jo (2003): Hermeneutische Polizeiforschung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Richardson, Emma/Hamann, Magnus/Tompkinson, James/Haworth, Kate/Deamer, Felicity (2023): Understanding the Role of Transcription in Evidential Consistency of Police Interview Records in England and Wales. In: Language in Society, S. 1–32.
Schwartz, Louis-Georges (2009): Mechanical witness: A history of motion picture evidence in U.S. courts. Oxford: Oxford University Press.
Seidensticker, Kai (2019): Fehlerkultur der Polizei. Die Wirkung von Organisationsstruktur und Männlichkeitskonstruktionen auf den Umgang mit Fehlern. In: SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis (3), S. 78–91.
SPD/FDP/Die Grünen (2021): Koalitionsvertrag (2021–2025). Mehr Fortschritt Wagen: Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit. Zwischen der Sozialdemokratischen Partei Deutschlands (SPD), Bündnis 90/Die Grünen und den Freien Demokraten (FDP). Berlin: Schloemer & Partner.
Sticher, Birgitta/Schicht, Günter (2019):Gute Vernehmungen – Lernen kann jeder: Wie können die Kompetenzen der PolizeibeamtInnen während der gesamten Berufsbiographie gefördert werden, um gute Vernehmungen zu gewährleisten? Berlin: FÖPS Digital.
STRG-F (2020): Exklusiv: Die Vernehmungen von Stephan Ernst. www.youtube.com/watch?v=lF6_8sAHPZQ (10.11.2024).
Svennevig, Jan/Urbanik, Paweł/Diepeween, Aafke (2024): How Police Investigators Seek to Secure That Suspects Speaking a Second Language Understand Their Rights in Investigative Interviews. In: Police Practice and Research 25(3), S. 324–42.
Tagesschau (2024): Vergewaltigungsprozess in Avignon: Tatvideos erstmals öffentlich gezeigt. www.tagesschau.de/ausland/europa/avignon-prozess-tatvideos-vergewaltigung-100.html (05.10.2024 16:38 Uhr) (15.11.2024).
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie: Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer Fachmedien.
Tuma, René/Egbert, Simon/Reichertz, Jo (2022): Videobilder und die Mediatisierung des Polizierens. In: Thüne, Martin/Klaas, Kathrin/Feltes, Thomas (Hrsg.): Digitale Polizei: Einsatzfelder, Potenziale, Grenzen und Missstände. Frankfurt am Main: Verlag für Polizeiwissenschaft, S. 263–280.
Tuma, René (2023): Trans-Sequentielle Analyse und Kommunikativ-Konstruktivistische Videographie: Kommunikative Formen, interaktive Trajektorien und Arbeitsbögen. In: Sozialer Sinn, 24(2), S. 315–352.
Ullrich, Peter/Knopp, Philipp (2018): Protesters' Reactions to Video Surveillance of Demonstrations: Counter-Moves, Security Cultures, and the Spiral of Surveillance and Counter-Surveillance. In: Surveillance & Society 16(2), S. 183–202.
Ullrich, Peter (2019): Polizei im/unter Protest erforschen: Polizeiforschung als Entdeckungsreise mit Hindernissen. In: Howe, Christiane/Ostermeier, Lars (Hrsg.): Polizei und Gesellschaft. Wiesbaden: Springer Fachmedien, S. 155–189.
Wickel, Tobias (2020): Die Pflicht zur audiovisuellen Aufzeichnung von Beschuldigtenvernehmungen im Ermittlungsverfahren und ihre Bedeutung im Zusammenhang mit Beweisverwertungsverboten. In: ZIS, 6(20).

II. Videographische und medienanalytische Ansätze mit Vernakularen Videos natürlicher sozialer Situationen und inszenierter Inhalte

2. Medial-gattungsförmig gestaltete Vernakulare Videos

Videoanalyse und Videointeraktionsanalyse als Methoden der empirischen Sexualforschung

Sven Lewandowski

Zusammenfassung: Anhand von empirischen Beispielen aus dem DFG-Projekt *Die Praxen der Amateurpornographie* zeigt der Beitrag, dass sich mittels Video(interaktions)analysen amateurpornographischer Videos Zugänge zu *privaten* sexuellen Praxen gewinnen lassen, die klassischen Methoden der empirischen Sozialforschung – wie teilnehmender Beobachtung oder Interviews – verschlossen bleiben. Exemplarisch werden habituelle Interaktionsmuster, medial-sexuelle Praxen sowie Selbstbeobachtungsverhältnisse analysiert. Besonderes Augenmerk gilt zudem Triangulierungen von amateurpornographischen Videos und narrativen Interviews, die sich bereits während der Interviews entwickelten, da Interviewte Fotos und Videos zeigten *und* ›live‹ kommentierten. Thematisiert werden schließlich forschungsethische Herausforderungen, die sich einerseits aus der Arbeit mit amateurpornographischem Material und andererseits daraus ergeben, dass dieses Material (potenziell) auch Dritten zugänglich ist, sodass die Gefahr besteht, dass mittels KI und Bilderkennungssoftware Anonymisierungsstrategien ausgehebelt werden können.

Schlüsselwörter: empirische Sexualforschung, Pornographie, sexuelle Interaktion, Videointeraktionsanalyse, qualitative Forschungsmethoden

1. Einleitung

Während sich Sexualität vornehmlich, wenngleich nicht ausschließlich, als *körperliche Praxis* realisiert, bleibt diese wissenschaftlicher und insbesondere teilnehmender Beobachtung unzugänglich. Folglich dominieren in der empirischen Sexualforschung interviewbasierte Erhebungsmethoden, mittels derer sich jedoch, so der praxistheoretische Konsens (Hillebrandt 2014; Schäfer 2016; Schmidt 2012), habitualisierte körperliche Praxen, die zudem der Selbstbeobachtung der Akteur:innen entgehen, nicht adäquat erheben lassen.

Andererseits finden sich im Internet amateurpornographische Videos, die private sexuelle Interaktionen dokumentieren[1] und videointeraktionsanalytisch

1 Im Folgenden werden nur Fotos und Videos, die von privaten Akteur:innen selbst erstellt wurden, als Amateurpornographie bezeichnet.

sowie videoanalytisch untersucht werden können. Einen Zugang zu privaten sexuellen Praxen bieten sie insofern, als es sich bei ihnen um natürliche (Ernst 2015) bzw. vorgefundene, also nicht um durch reaktive Erhebungsverfahren gewonnene (Ego-)Dokumente (Salheiser 2014) und mithin um (quasi-)autoethnographisches Material handelt.

Das an der Universität Bielefeld durchgeführte DFG-Projekt *Die Praxen der Amateurpornographie*[2] triangulierte amateurpornographische Videos und narrative Interviews, um sexuelle Praxen wie auch die Entstehungskontexte der Videos sowie Motive und Selbstdeutungen der Akteur:innen zu erforschen. Die Videos dienten also sowohl als Mittel zur Erforschung der in ihnen dokumentierten sexuellen Interaktionen als auch als genuiner Untersuchungsgegenstand. Zudem wurden (Screenshots der) Videos als Stimuli im Rahmen der Interviews genutzt.

Im Verlauf des Forschungsprojekts wurden Triangulierungen (Flick 2011) sowohl ausgeweitet als auch verfeinert, sodass nicht nur Videos und Interviewaussagen trianguliert wurden, sondern beispielsweise auch unterschiedliche Videos derselben Akteur:innen, Selbstdarstellungen in Webforen mit Selbstdarstellungen im Interview sowie Interaktionsmuster, die sich in Videos zeigten, mit Interaktionsmustern, die während der Interviews beobachtet wurden. Ebenfalls wurden unterschiedliche Elemente von Videos miteinander trianguliert.

Da habitualisierte und verkörperte sexuelle Muster in ähnlicher Weise wie der Bourdieu'sche Habitus als »strukturierende Strukturen« (Bourdieu 1980, S. 98) fungieren und sich in jeder sexuellen Interaktion reproduzieren, erlauben Video*interaktions*analysen amateurpornographischer Videos Rückschlüsse auf die außerpornographische Sexualität der Akteur:innen und eröffnen so der Sexualforschung Möglichkeiten einer adäquaten Beobachtung und Analyse sexueller Praxis.

2. Forschungsdesign, Erhebungskontext und Feldzugang

Amateurpornographische Videos sind leicht zu finden. Schwieriger ist es hingegen sie als *echt* und *konsensuell* erstellt (und verbreitet) zu authentifizieren sowie ihre Akteur:innen ausfindig zu machen *und* als Studienteilnehmer:innen zu gewinnen. Hierbei erwies sich die anvisierte Triangulierung von Videos und Interviews als zusätzliche Hürde: Nicht wenige Personen wären bereit gewesen, Videos zu Verfügung zu stellen *oder* sich interviewen zu lassen – nicht aber beides.

Studienteilnehmer:innen lassen sich entweder ausgehend von Videos oder umgekehrt Videos ausgehend von Studienteilnehmer:innen finden. Der Vorteil des ersten Wegs liegt darin, dass er ein *theoretical sampling* anhand von Videos

2 DFG-Projekt-Nr. 397130176; GZ: LE 2224/2-1. Für ihr Engagement als wissenschaftliche Mitarbeiterinnen danke ich Tara Siemer und Nadine Giesbrecht herzlich.

und Kriterien wie Mindestdauer und dem Vorkommen sexueller Interaktion erlaubt. Allerdings besteht die Gefahr eines unbewussten ›pragmatischen‹ Bias, der dazu verleiten mag, Videos auszuwählen, die sich einfacher analysieren lassen, weil sie beispielsweise besser gefilmt oder ausgeleuchtet sind. Erfolgreicher als der erste Weg, der mitunter der Suche nach einer Nadel im Heuhaufen gleicht, erwies es sich, (potenzielle) Teilnehmer:innen über einschlägige Webforen, Sexkontaktseiten usw. zu kontaktieren. Ein Nachteil dieses Wegs mag sein, dass die Teilnehmer:innen die Videos auswählen, die sie zur Verfügung stellen. Allerdings lassen sich auf diese Weise auch unveröffentlichte Videos finden.

Insgesamt konnten sechszehn Fälle in die Studie einbezogen werden. Zehn narrative Interviews wurden mit Paaren, sechs weitere mit Einzelpersonen geführt. Die durchschnittliche Interviewdauer lag bei ca. 3,8 Stunden mit einer Bandbreite von knapp einer bis zu sechs Stunden.[3] Mehr als zehn Stunden Videomaterial wurden ausgewertet.[4] Bei den untersuchten Videos, die typischerweise aus sexuellen Motiven gedreht wurden, handelt es sich um vorgefundenes, also – im Gegensatz zu Videographien – nicht zu Forschungszwecken erstelltes Datenmaterial. Auch wurden Teilnehmer:innen nicht gebeten, extra für die Studie Videos anzufertigen. Da ihnen die Auswahl der zu untersuchenden Videos überlassen wurde, handelt es sich bei diesen freilich nicht um ›Rohmaterial‹, sondern um Produkte mehrfacher Selektionsprozesse (vgl. Lewandowski/Siemer 2021a, S. 87 f.). Manche Teilnehmer:innen teilten einfach einen Link zu Videoalben, die sie auch mit Freund:innen und Sexualpartner:innen teilen, oder zu ihren Profilen auf einschlägigen Websites, andere übermittelten ihre Lieblingsvideos, dritte wählten einige ihrer Videos mit Blick auf eine gewisse Repräsentativität aus oder stellten Videos zur Verfügung, von denen sie annahmen, dass sie für die Studie besonders interessant seien. Nicht zuletzt dürften primär Videos ausgewählt worden sein, in denen sich die Teilnehmer:innen (habituell) wiedererkennen und die ihrem Selbstbild entsprechen. Die Auswahl dürfte mithin im Großen und Ganzen ähnlichen Kriterien folgen wie die Auswahl der Videos, die sie hochladen oder mit anderen Personen teilen. Zur Verfügung gestellt und/oder gezeigt wurden aber auch unveröffentlichte Videos. In zwei Fällen hatten die Teilnehmer:innen noch nie eines ihrer Videos anderen gezeigt, geschweige denn hochgeladen.

Unterschiedliche Praxen amateurpornographischen Filmens eröffnen, strukturieren und limitieren zudem die Möglichkeiten der Video(interaktions)analyse. Gleiches gilt für nachbearbeitete Videos. Nachbearbeitungen waren in den untersuchten Videos jedoch selten und beschränkten sich meist darauf, dass

3 Alle Interviews wurden vom Autor und einer wissenschaftlichen Mitarbeiterin gemeinsam und – abgesehen von zwei pandemiebedingten Interviews per Zoom – *face-to-face* geführt.

4 Weitere etwa sechs Stunden ausgewertetes Videomaterial konnten nicht in die Studie einfließen, da keine Interviews zustande kamen.

Anfänge und/oder Enden abgeschnitten wurden. Schnitte innerhalb von Videos waren sehr selten.[5]

Während händisches Filmen sexuelle Interaktionen oftmals unvollständig in relativ kurzen und nicht selten technisch eher schlecht gefilmten und mitunter verwackelten Videos dokumentiert, sind Videos, die mit einer statisch fixierten Kamera aufgenommen wurden, meist länger, zeigen sexuelle Interaktionen aber häufig aus der Totalen oder Halbtotalen, also quasi ›von außen‹, jedoch oftmals in besserer Bildqualität. Allerdings erlauben händisch aus *point-of-view*-Perspektiven gefilmte Videos den Blick der kameraführenden Person auf die sexuelle Interaktion und das Kamerahandeln zu analysieren. Zudem zeigen sie mehr Details sowie *close-ups*, die Mikrointeraktionen leichter erkennbar machen. In greifbarer Nähe befindliche Smartphones erlauben es schließlich *während* sexueller Interaktionen und ohne diese nennenswert zu unterbrechen, spontan mit dem Filmen zu beginnen, es zu unterbrechen, zu einem späteren Zeitpunkt fortzusetzen oder zu beenden. Für die videointeraktionsanalytisch verfahrende Sexualforschung eröffnen solche ad hoc entstandenen Videos Möglichkeiten, gleichsam noch näher an (alltags-)sexuelle Interaktionen ›heranzurücken‹.

Im Gegensatz zu den untersuchten Videos, die von Fall zu Fall in Anzahl, (Mach-)Art, Länge und technischer Qualität erheblich variieren, handelt es sich bei den durchgeführten Interviews nicht um natürliche, sondern zu Forschungszwecken generierte Daten. Ihre ›Nicht-Natürlichkeit‹ wurde freilich durch die gewählte Methode des narrativen Interviews ›abgemildert‹, die auf die Erzeugung einer möglichst natürlichen Gesprächsatmosphäre zielt (Küsters 2009), was insbesondere bei Paarinterviews gelang, in denen sich Gespräche *zwischen* den Partner:innen entwickelten.

Der skizzierte Erhebungskontext beeinflusst die Analyse des Datenmaterials zunächst dadurch, dass sie sich auf Videos beschränken muss, die die Teilnehmer:innen nach eigenen Kriterien, teilweise aber mit Blick auf ihre Erwartungen hinsichtlich der Studie auswählten. Zwar wäre es denkbar, anhand von Forschungsfragen ein Sample z. B. aus den beliebtesten oder aus zufällig ausgewählten amateurpornographischen Videos, die auf einschlägigen Websites hochgeladen wurden, zu bilden. Abgesehen von datenschutzrechtlichen Fragen wird es bei derartigen Samples jedoch schwierig, die Akteur:innen für die Studienteilnahme *und* für Interviews zu gewinnen.

Aufgrund der gewählten Erhebungsmethode sind andererseits Fülle, Art usw. des zu analysierenden Video- und sonstigen Materials nicht steuerbar. Während manche Teilnehmer:innen nur wenige und/oder kurze Videos verfügbar machten,

5 Eine Ausnahme bilden jedoch zwei Studienteilnehmer, die als ambitionierte Hobbyfotografen sowohl über das technische Knowhow als auch über einen künstlerischen Ehrgeiz verfügen, der es ihnen verbietet, Videos unbearbeitet zu lassen. Eine Analyse eines semiprofessionellen (pseudo-)amateurpornographischen Videos, die herausarbeitet, wie die Kamera mittels Schnitten ›lügt‹ und Pannen kaschiert, findet sich in Lewandowski (2021).

überschütteten andere die Studie geradezu mit (teils überaus langen) Videos, hunderten von Fotos usw., was die Forschenden zur Auswahl einzelner Videos bzw. Videosequenzen zwingt, wobei die angesprochene Gefahr besteht, nicht nur im Sinne eines *theoretical samplings*, sondern (unbewusst) auch nach forschungspragmatischen und/oder individuellen Präferenzen auszuwählen. Derartigen ›Versuchungen‹ lässt sich dadurch begegnen, dass Projektmitarbeiter:innen zunächst unabhängig voneinander Videos, Szenen und Sequenzen auswählen. Dass nur Videos analysiert werden können, wenn Personen, die sie gefilmt haben, zur Teilnahme an der Studie gewonnen werden können, führt schließlich unter Umständen zu einer Über- oder Unterrepräsentation bestimmter Personengruppen, bestimmter Videos und/oder sexueller Praxen wie Praktiken.[6]

2.1 Datenaufbereitung

Um Interaktionen und Details besser erkennen zu können, wurden Videos vor ihrer Analyse gegebenenfalls aufbereitet, indem sie aufgehellt, wichtige Sequenzen in Einzelbildfolgen zerlegt sowie Screenshots angefertigt und vergrößert wurden. Um verbale Äußerungen besser zu verstehen, wurden mitunter Teile der Audiospur gefiltert.

Nach einer ersten Sichtung und kurzen Beschreibung wurden die Videos mittels atlas. ti zunächst offen kodiert und in Sequenzen gegliedert. Die Sequenzierung folgte dabei anhand der Geschehensabläufe (vgl. Tuma 2018, S. 434 ff.), zu den einerseits Interaktionen und andererseits Kamerahandlungen zählen. Ausgewählte Sequenzen sowie alle verbalen Äußerungen wurden umfangreich transkribiert und analysiert. Zusätzlich wurden Screenshots erstellt, in denen wiederum relevante Objekte (etwa Sexspielzeug oder auch Einrichtungsgegenstände), Personen, Interaktionen usw. kodiert wurden. Während der jeweiligen Arbeitsgänge wurden die kodierten ›Zitate‹ umfangreich kommentiert, Beschreibungen wichtiger Videosequenzen verfasst und erste Deutungsansätze entwickelt. In analoger Weise wurde mit den vollständig transkribierten Interviews verfahren, die sowohl als Transkript als auch als Audiodatei in atlas.ti eingespeist wurden. Schließlich wurden Interviewsequenzen, Videos, Screenshots, Fotos und weitere Egodokumente der Akteur:innen miteinander verknüpft.

6 Beispielsweise erwies es sich als besonders schwierig, Studienteilnehmer:innen zu gewinnen, die Videos ausschließlich für sich selbst drehen. Zudem zeigen praktisch alle untersuchten Videos heterosexuelle Praktiken.

2.2 Exkurs: Forschungspraxis

Empirische Pornographieforschung konfrontiert Forschende mit Darstellungen, die auf sexuelle Erregung abzielen, und stellt sie vor die Aufgabe, entsprechende Affizierungen zu managen. Ein solches Management erfordert genrespezifische Erfahrungen und Kenntnisse, sexuelle Toleranz sowie die Bereitschaft *und* Fähigkeit, auch sexuelle Praktiken und Praxen nüchtern zu betrachten, die dem eigenen Begehren und/oder eigenen moralischen Vorstellungen *nicht* entsprechen. Insbesondere die *gemeinsame* Videoanalyse benötigt formale wie informelle Desexualisierungs- und Entschärfungsstrategien sowie die Implementierung von *op-out*-Möglichkeiten (vgl. Lewandowski/Siemer 2021b, insbesondere S. 93 ff.).

Desexualisierungs*effekte* stellen sich in der Forschungspraxis allerdings bereits dadurch ein, dass Videos in Sequenzen, Einzelbildfolgen und Screenshots zerlegt, in Zeitlupe angesehen, die Aufmerksamkeit auf Details, genaue Bewegungsabläufe, Kamerahandlungen, habituelle Gesten oder auch auf Einrichtungsgegenstände im Hintergrund gerichtet wird und – last but not least – die Videos wieder und wieder angesehen werden, sodass sich das explizit Sexuelle zunehmend verflüchtigt und – zumal mit zunehmender Forschungserfahrung – die sexuellen Interaktionen primär als ›ganz normale‹ Interaktionen wahrgenommen werden. Nicht von ungefähr lautet ein zentrales Mantra der soziologischen Analyse sexueller Interaktionen, dass man sich nicht vom Sexuellen blenden und von der Interaktion ablenken lassen darf.

3. Verwendung von Videos im Rahmen narrativer Interviews

Zur Vorbereitung auf Interviews gilt es, die Videos der Studienteilnehmer:innen nicht lediglich oberflächlich zu sichten, sondern sie möglichst vollständig zu transkribieren und zu kodieren sowie erste Deutungen zu erarbeiten. Vertrautheit der Forschenden mit den Videos wird von den Interviewten nicht nur als Ausdruck von Kompetenz und Anerkennung geschätzt, sondern ist unabdingbar, um zu wissen, worüber gesprochen wird und um gegebenenfalls Nachfragen zu einzelnen Videos und Interaktionen stellen zu können. Zum Zwecke der *photo elicitation* (Harper 2002; Dimbath 2013) wurden Screenshots ausgewählt, von denen angenommen wurde, dass sie Erinnerungen stimulieren und Gesprächsimpulse setzen können, sowie solche, die Szenen, Handlungen oder Gegenstände zeigen, auf die sich die Forschenden keinen Reim machen konnten und/oder die in den Videos nicht recht zu erkennen waren. Die Methode der photo elicitation erwies sich in den meisten Fällen als überaus erfolgreich und wurde zudem dadurch ›validiert‹, dass sich Interviewte unaufgefordert und teilweise bereits *vor* der Vorlage von Screenshots analoger Methoden bedienten, indem sie Fotos und

Videos zeigten, um so das Problem zu lösen, dass sich sowohl sexuelle Interaktionen als auch Videos nur schwer beschreiben lassen. Zugleich testete das Vorführen von Videos und Fotos, die (noch) nicht zur Verfügung gestellt wurden respektive unveröffentlicht waren, die Interviewer:innen und validierte das während des Interviews aufgebaute Vertrauen. Videovorführungen hatten vor allem den Effekt, dass über Videos und sexuelle Interaktionen nicht nur ›abstrakt‹ gesprochen wurde, sondern diese live und ad hoc kommentiert wurden, sodass eine Triangulation von Interviews und Videos nicht nur ›vertieft‹, sondern *direkt* im Interview hergestellt werden konnte.

Szenen und Szenenkommentare wurden später in atlas. ti präzise verknüpft, indem Geräusche oder Dialoge aus den Videos, die im Mitschnitt des Interviews eindeutig auszumachen waren, als Referenzpunkte der Synchronisierung bzw. Zuordnung von Video- und Interviewsequenzen gewählt wurden. Eine solche Form der Verknüpfung geht über die üblichen, vergleichsweise unspezifischen Verknüpfungen von Interviewaussagen über Videos bzw. sexuelle Interaktionen mit den entsprechenden Videosequenzen hinaus und eröffnet so neue Analysemöglichkeiten, da sie u. a. erlaubt, nicht nur ad hoc auftauchende Erinnerungen und spontane Reaktionen der Interviewten auf ihre Videos einzufangen, sondern auch nachvollziehbar macht, auf was bzw. welche Videosequenzen sie wie reagieren und was sie als erklärungsbedürftig, zeigens- und/oder kommentierenswert wahrnehmen. Die Vorführung von Videos erinnerte mitunter an Datensitzungen, in der Personen mit unterschiedlichem Kenntnisstand, gemeinsam Videos deuten, zumal die Interviewten auch herauszufinden versuchten, was die Forschenden in diesen sehen. Jedoch müssen sich letztere mit Deutungen zurückhalten, da das Ansehen von Videos während der Interviews nicht ihrer gemeinsamen Interpretation und der in ihnen dokumentierten Interaktionen, sondern der Generierung von Datenmaterial dient. Selbstdeutungen der Interviewten lassen sich gleichwohl mittels Formulierungen vorsichtig hinterfragen, die zu weiteren Reflexionen anregen (»wir haben zuerst vermutet …«).

Der Nutzen einer Triangulierung von Interviews und Videos erweist sich u. a. darin, dass die Interviewten nicht nur ›Expert:innen in eigener Sache‹ sind, sondern vor allem als Teilnehmende und Zeug:innen der Interaktionen gegenüber den Forschenden, die diese nur aus den Videos kennen, einen Wissensvorsprung derart haben, dass sie Interaktionen aus Perspektiven erlebt und gesehen haben, die den Forschenden unzugänglich bleiben. Die Kommentierungen und Deutungen ihrer Videos durch die Interviewten geben jedoch nicht nur Auskunft über ihre Selbstdeutungen, sondern auch über blinde Flecke. Sie sind also nicht nur zu analysierendes Datenmaterial, sondern insofern auch Beiträge zur Forschung, als sie, wie folgendes Beispiel zeigt, die Forschenden veranlassen (sollten), eigene Deutungen kritisch zu überprüfen, zu präzisieren und gegebenenfalls zu revidieren.

Ein Video zeigt, wie Manuela auf dem Bauch liegt und von Murat,[7] der hinter ihr steht, penetriert wird. Aufgenommen wurde es mit einem Tablet, das so vor Manuelas Kopf platziert ist, dass ihr Gesicht prominent im Bild ist und von *beiden* beobachtet werden kann. Während Murat zustößt, verzieht Manuela immer wieder ihr Gesicht, was zunächst als Ausdruck von Schmerz gedeutet wurde. Zugleich erweckt die Sequenz den Eindruck, Murat würde Manuela – möglicherweise gegen ihren Willen – dominieren. Im Interview gibt Manuela an, gerne Rollen zu spielen. Murat verleiht hingegen der Befürchtung Ausdruck, dass das Video den Eindruck erweckt haben könnte, er unterdrücke Manuela. Während des Interviews wird zudem deutlich, dass in ihrer Beziehung Manuela den Ton angibt.

Vor diesem Hintergrund rückte bei einer Re-Analyse des Videos die Ausrichtung des Tablets noch stärker in den Fokus der Aufmerksamkeit. Dass das Tablet Murat, obwohl er hinter Manuela steht, ermöglicht, ihr Gesicht zu sehen, macht es plausibel, dass es nicht primär auf Schmerzen reagiert, sondern Manuela mit Blick auf die für Murat entworfene Inszenierung bewusst grimassiert.

Wenngleich das Beispiel illustriert, wie sich aus Interviews gewonnene Informationen für die (Re-)Analyse von Videos fruchtbar machen lassen, können die Deutungen der Videos durch ihre Akteur:innen nicht als die zentrale Referenz bzw. letzte Instanz der Analyse gelten. Vielmehr müssen sich *alle* Deutungen am Videomaterial messen und belegen lassen. Die Interviews steuern aber Kontextwissen bei, das sich weder mittels Videoanalysen noch mittels Videointeraktionsanalysen gewinnen lässt, aber manche der in den Videos dokumentierten Interaktionen zu deuten bzw. zu erkennen erlaubt (z. B. Rollenspiele). Nicht zuletzt wissen die Interviewten, was passierte, bevor die Kamera ein- und nachdem sie wieder ausgeschaltet wurde. Die Selbstdeutungen ihrer Videos durch die Interviewten haben für die Forschung also eine doppelte Funktion, indem sie einerseits als Datenmaterial fungieren, aus dem sich insbesondere Erkenntnisse zu Selbstbildern, Selbstverständnissen und Selbstdarstellungen/-inszenierungen sowie Selbstdeutungen, aber auch möglichen Selbsttäuschungen gewinnen lassen. Andererseits leisten sie aber auch, indem sie kritische Überprüfungen der Videoanalysen durch die Forschenden anregen, einen Beitrag zur Forschung.

4. Analysekontexte: Video- und Videointeraktionsanalysen amateurpornographischer Praxen

Im Gegensatz zu Videoanalysen, die auf die Gestaltung und Machart von Videos (Kameraführung, Schnitt, Montage usw.) fokussieren, widmen sich Videoi*nteraktions*analysen *nicht* Videos, sondern den in ihnen dokumentierten Interaktionen

7 Alle Namen von Studienteilnehmer:innen, denen herzlich gedankt sei, sind selbstverständlich Pseudonyme.

(vgl. Knoblauch 2011, S. 141 ff.; Knoblauch/Schnettler/Tuma 2010, S. 22 ff.; Tuma 2018; Tuma/Schnettler/Knoblauch 2013, S. 85 ff.).[8] Amateurpornographische Videos zeigen sexuelle Interaktionen freilich nicht selten unvollständig und aus für *Interaktionsanalysen* ungünstigen Perspektiven, beispielsweise in Detail- und *close-up*-Aufnahmen, und/oder wurden mit handgeführten Kameras sowie aus *point-of-view*-Perspektive gefilmt. Obendrein sind nicht wenige Videos recht kurz.

In den folgenden Abschnitten wird jeweils anhand von Fallbeispielen illustriert, wie sich mit video(interaktions)analytischen Methoden (paar-)sexuelle Habitus, sexuell-mediale Praxen und Selbstbeobachtungsverhältnisse in, anhand und mittels amateurpornographischer Videos untersuchen lassen.

4.1 Rekonstruktion habitueller Interaktionsmuster

Gerade anhand von Szenen, die *keine* allzu persönlichen Interaktionen zeigen *sollen*, lässt sich gut veranschaulichen, dass sich habituelle Prägungen trotz gegenteiliger Intention der Akteur:innen in Videos niederschlagen. So zeigen sich in den Videos, die Sabrina mit ihrem Mann dreht, um sie zu verkaufen,[9] immer wieder (paar-)habituelle Prägungen, die sich daran erkennen lassen, dass sie nicht nur in mehreren ihrer Videos zu beobachten sind, sondern nicht selten ›quer‹ zum intendierten Sinngehalt der jeweiligen Videos stehen. Ein Video inszeniert beispielsweise, wie Sabrinas als Geschäftsmann verkleideter Mann ein Zimmermädchen (Sabrina) überwältigt und mit seiner rechten Hand auf ein Bett stößt. Allerdings fängt er ihren Fall zugleich mit seinem linken Arm ab. Als sie ihn schließlich fellationiert, sucht und findet ihre Hand die seine und hält sie zärtlich. Der Widerspruch beider Handlungen – das Abfangen ihres Falls wie das Händchenhalten – zur intendierten Inszenierung ist so deutlich, dass sie wohl unbewusst geschehen und als Elemente des privaten paarsexuellen Habitus zu betrachten sind, die in die pornographische Inszenierung hineinragen. Dass es sich insbesondere beim Händchenhalten nicht um eine Pose, sondern um eine habituelle Geste[10] handelt, zeigt sich nicht zuletzt daran, dass sie im Gegensatz zu anderen Elementen der Inszenierung nicht betont wird, sondern eher beiläufig geschieht bzw. sich gewohnheitsmäßig einschleicht. Während sich in Sabrinas Videos habituelle Gesten aufgrund des Kontrastes zu inszenierten Rollenspielen relativ leicht erkennen lassen, muss sich die Analyse habitueller Prägungen in *privaten* amateurpornographischen Videos stärker an Homologien halten, die sich

8 Videos sind für Videointeraktionsanalysen mithin Mittel zum Zweck der Analyse von Interaktionen, nicht aber genuiner Forschungsgegenstand (vgl. Knoblauch/Wilke i. d. B.).

9 Sabrina stellt insofern einen Kontrastfall innerhalb des Forschungsprojekts dar.

10 Zur Unterscheidung von Pose und habitueller Geste, vgl. Bohnsack (2017).

in verschiedenen Videos ein und desselben Paars, aber auch in seinen Interaktionen während des Interviews beobachten lassen.[11] Sabrinas Beispiel verdeutlicht aber, dass sich paarsexuelle Habitus auch gegen die Intentionen der Akteur:innen durchsetzen.

4.2 Medial-sexuelle Praxen

Amateurpornographische Videos dokumentieren nicht lediglich sexuelle Interaktionen, sondern amateurpornographisches *Filmen* ist selbst eine spezifische mediale Praxis, die sich indirekt in den Videos dokumentiert. Obgleich sie eine *sine qua non* amateurpornographischen Filmens ist, haben viele Studienteilnehmer:innen ein ambivalentes Verhältnis zur Kamera. Insbesondere das Filmen mit einer Handkamera, die aktiv geführt werden muss und nur eine Hand freilässt, lenkt vom Sex bzw. umgekehrt der Sex vom Filmen ab. Folglich wird die Kamera häufig beiseitegelegt oder abgeschaltet, um sich auf die sexuelle Interaktion zu konzentrieren, sodass diese nur unvollständig bzw. ausschnitthaft dokumentiert wird. Hingegen erinnert eine auf einem Stativ befestigte Kamera manche Interviewte zu sehr an kommerzielle Pornographie, während anderen die Suche nach einer optimalen Kameraposition zu zeitintensiv ist. Manuela schätzt jedoch das *aktive* Filmen mit einer handgeführten Kamera als sexuelle Praxis, die Sex qualitativ verändere. Würde mit einer statischen Kamera gefilmt, »wär's einfach ganz normaler Sex wie sonst auch«. Murat argumentiert hingegen, dass Sex, der mit einer statischen Kamera gefilmt werde, »geschmeidiger« sei.

Die folgende Analyse einer Videosequenz,[12] die zeigt, wie Michael seine Partnerin Andrea leicht würgt, arbeitet mittels videointeraktions- und videoanalytischer Verfahren ein Zusammenspiel von sexuellem Handeln und Kamerahandeln heraus.[13] Im Interview betont Michael, dass sie diese Praktik »immer mal probieren« wollten und er es »in diesem Video echt (,) sehr spontan […] gemacht« habe. Die aus Michaels Perspektive gefilmte Szene zeigt Sex in der Reiterstellung, wobei Andrea auf Michael sitzt und sich die Kamera zwischen Andreas Brüsten und ihrem Gesicht, die jeweils fast bildfüllend gezeigt werden,

11 So lässt sich sowohl in Katjas und Stefans Videos als auch im Interview mit ihnen ein spielerischer Umgangsstil miteinander beobachten.

12 Videointeraktionsanalysen zum Umgang mit Pannen während sexueller Interaktionen sowie von Livestreams finden sich in Lewandowski/Siemer (2021a) respektive in Lewandowski (2023).

13 Die Analyse des Kamerahandelns erfordert einen videoanalytischen Zugriff, da es selbst nicht gefilmt wird und folglich nicht im Bild ist, sondern anhand des Videos rekonstruiert werden muss. Allerdings ist es, wie das folgende Beispiel zeigt, durchaus Teil der sexuellen Interaktion. Im Folgenden wird es primär mit Blick auf die sexuelle Interaktion analysiert, wobei freilich auch andere Analysemöglichkeiten offen stünden.

hin und herbewegt. Sodann zoomt die Kamera zurück, sodass Andreas Kopf *und* Oberkörper zu sehen sind. Kurz wackelt das Bild unkoordiniert und ein Schatten huscht durch es hindurch. Wie die Zeitlupe zeigt, rühren das Schwanken des Bildes und der Schatten daher, dass Michael seine linke Hand zur Kamera führt, die sie von seiner rechten übernimmt. Nachdem sich das Bild wieder stabilisiert hat, löst sich Michaels rechte Hand, die sich nun auf Andreas linker Brust befindet, von dieser, fährt über ihren Brustkorb an ihren Hals und umfasst ihn mit einem Würgegriff. Liest man – anstatt sich auf den Übergang zum Würgen zu konzentrieren – das Kamerahandeln retrospektiv, wird deutlich, dass nicht erst der Wechsel der kameraführenden Hand, sondern bereits die Veränderung der Kameraeinstellung der Vorbereitung der Würgeszene diente. Die Kamera wusste also, wo die ›Action‹ stattfinden würde, *und* antizipierte, dass das beschriebene Kamerahandeln erforderlich sein würde, um diese gut ins Bild zu bekommen. Das Würgen ergab sich folglich nicht spontan aus der Interaktion, sondern ihm ging ein Entschluss Michaels *voraus*, der sich in gleich zwei vorbereitenden Kamerahandlungen manifestiert, die sicherstellen, dass die Action adäquat dokumentiert wird. Was zunächst wie eine zufällig verwackelte Kamerabewegung aussieht, erweist sich als komplexe (Kamera-)Handlung, die ein kommendes Geschehen antizipiert *und* vorbereitet.[14]

Dass die sexuelle Handlung, die gefilmt werden soll, erst ausgeführt wird, *nachdem* ihre audiovisuelle Dokumentation sichergestellt wurde, zeigt zudem wie sexuelles Handeln und audiovisuelles Handeln (mitunter) interagieren. Nicht lediglich die Kamera reagiert auf sexuelle Handlungen, sondern sexuelle Handlungen (mitunter) auch auf das Handeln der Kamera. Dass nicht die Kamera der Action, sondern umgekehrt die Action der Kamera folgt und sich an dieser orientiert, zeigt, dass eine sexuelle Handlung (auch) medial motiviert sein kann bzw. der Wille zu ihrer Dokumentation handlungsleitend werden kann. Jedenfalls wird sie nicht ausgeführt *bevor* nicht die Kamera entsprechend ausgerichtet ist. Dass das Kamerahandeln den sexuellen Handlungsfluss unterbricht, zeugt schließlich davon, dass ein Handlungsvorsatz getroffen wurde, Michael also nicht im Affekt handelt. Indem sie zeigt, dass Michael im Gegensatz zu seiner Aussage nicht spontan gehandelt hat, verdeutlich die Analyse zugleich, dass Interviews keine adäquate Methode sind, um körperliche Mikrointeraktionen zu erforschen.

14 Das Wissen der Kamera um künftiges Geschehen rührt freilich daher, dass Michael sie führt und seine Handlungsabsicht kennt.
Die hier beschriebene Kamerabewegung bzw. -handlung ließe sich im Sinne von Bohnsacks dokumentarischer Methode der Videoanalyse analysieren (Bohnsack 2011). *An dieser Stelle* zielt ihre Analyse aber nicht auf den Habitus des abbildenden Akteurs, sondern fungiert als Mittel, um zu zeigen, dass eine bestimmte sexuelle Handlung nicht spontan aus dem Interaktionsfluss heraus entsteht. Die Analyse der Kamerabewegung dient hier also der Interaktions- bzw. Handlungsanalyse, d. h. dazu etwas über Michaels Handeln, nicht aber über seinen Habitus als Bildproduzenten zu erfahren.

4.3 Selbstbeobachtungsverhältnisse

Videos lassen sich zur Selbstbeobachtung *ex post factum*, als Erinnerungen an frühere sexuelle Interaktionen und nicht zuletzt zu pornographischen Zwecken im engeren Sinne, d. h. zur Generierung sexueller Erregung nutzen. Wie Murats und Manuelas Nutzung des Tablets zeigt, eröffnet Filmen beim Sex aber auch Möglichkeiten, sich *während* sexueller Interaktionen selbst zu beobachten.[15]

Murat und Manuela integrierten das Tablet so in ihre sexuelle Interaktion, dass es ihnen (Selbst-)Beobachtungen aus Perspektiven erlaubte, die dem bloßen Auge unzugänglich sind.[16] Eine Platzierung von Kamera und Monitor in einiger Entfernung zu einer sexuellen Interaktion birgt hingegen, wie ein Video von Herbert und Marion zeigt, die Gefahr, dass Selbstbeobachtungsmöglichkeiten vom Sex ablenken. Im Gegensatz zu einer handgeführten Kamera lenken sie zwar nicht dadurch ab, dass sie aktiv bedient werden müssen, aber (große) Bildschirme ziehen Blicke an, woraus die Aufgabe erwächst, zwischen Selbstbeobachtung und sexueller Interaktion zu navigieren, sich also nicht zu sehr vom Geschehen auf dem Bildschirm faszinieren zu lassen und mehr mit Blick auf ihn als mit Blick auf den:die Partner:in zu agieren.

Im besagten Video schaut Herbert immer wieder an Marion vorbei auf den Monitor und obwohl sie ihn ermahnt, sich auf sie zu konzentrieren, fällt es ihm sichtlich schwer, seinen Blick vom Bildschirm zu lösen. Dass er im Interview bestreitet, andauernd auf den Bildschirm zu schauen, lässt vermuten, dass ihm nicht bewusst ist, wie stark dieser seine Aufmerksamkeit absorbiert. Dass er diese aus der Interaktion ›heraussaugt‹ liegt freilich *auch* an seiner Platzierung außerhalb der Interaktion. Murats und Manuelas im Raum der Interaktion platziertes Tablet zieht hingegen ihre Aufmerksamkeit nicht voneinander ab. Indem es Murat ermöglicht, Manuela aus zwei Perspektiven *zugleich* zu sehen,[17] und beiden – *obwohl* er hinter ihr steht – medial vermittelten Augenkontakt erlaubt, bietet ihnen das Tablet nicht nur eine Art ›augmented reality‹, sondern verbindet sie miteinander.

Eine weitere Erfahrung mit medialer Selbstbeobachtung schildert Stefan. Während einer sexuellen Interaktion habe er sich dabei »ertappt«, diese auf dem Bildschirm zu beobachten, obwohl

> *ich hätte nur nach unten gucken müssen dann hätte ich das gesehen was ich sehen wollte und ich hab mir das die ganze Zeit aber auf dem Monitor angeschaut.*

15 Livestreams und die Veröffentlichung von Videos ermöglichen zudem synchrone wie diachrone (Selbst-)Beobachtungen zweiter Ordnung, also zu beobachten, wie Zuschauer:innen die Akteur:innen beobachten.

16 Interessant ist zudem, dass sie eine (weitgehend) statische Kamera *aktiv* in die sexuelle Interaktion einbanden.

17 Ihren Rücken, Nacken und Hinterkopf mit bloßem Auge und ihr Gesicht auf dem Tablet.

Im Interview reflektiert bzw. rätselt Stefan, was ihn dennoch so am Bildschirm faszinierte. Im Gegensatz zu Herbert, dem selbst ex post seine Faszination für die Selbstbeobachtung auf dem Bildschirm nicht bewusst ist, »ertappte« sich Stefan *in actu*.[18] Synchrone Selbstbeobachtungsmöglichkeiten erweisen sich mithin als ähnlich paradox wie das Verhältnis zur Kamera – sie können von sexuellen Interaktionen ablenken, diese selbst und/oder das sexuelle Erleben aber auch qualitativ verändern sowie mit zusätzlichen Reizen aufladen. Freilich kann, wie andere Fälle zeigen, kein medientechnologischer Determinismus angenommen werden. Vielmehr müssen Selbstbeobachtungsmöglichkeiten und Ablenkungsrisiken in der Situation gemanagt werden.

Die hier behandelten Videos unterstreichen erneut die Vorzüge der Triangulation von Interviews und Videos. Die Analyse letzterer erlaubt es, Verhaltensweisen, wie etwa Herberts Faszination durch den Bildschirm, Sabrinas habitualisierte Zärtlichkeitsgeste oder Michaels Kamerahandeln, zu erfassen, die sich, da sie den Akteur:innen nicht oder nur halb bewusst sind, nicht mittels Interviews erheben lassen. Mittels photo elicitation und/oder Nachfragen lassen sich aus Videos gewonnene Erkenntnisse wiederum in Interviews einspeisen. Auch erlauben sie es, die Angaben der Interviewten kritisch zu hinterfragen, da sich in Videos Aspekte, Muster und Praxen beobachten lassen, die die Akteur:innen selbst nicht sehen (können oder wollen), weil sie ihrer Aufmerksamkeit entgehen, blinde Flecke berühren, zu tabuisiert sind oder aus ›(paar-)ideologischen‹ Gründen nicht thematisiert werden können. So hadert etwa Murat damit, dass sich sein Selbstbild modern(isiert)er Männlichkeit nicht recht mit seiner Präferenz für sexuelle Inszenierungen, in denen er (leicht) dominiert, verträgt. Zugleich wird am Beispiel von Murat und Manuela aber auch deutlich, dass amateurpornographisches Filmen zu einer eigenständigen sexuellen Praxis werden kann, die sich freilich nicht allein aus Interviews rekonstruieren lässt.

5. Nutzungsszenarien und forschungsethische Herausforderungen

Wissenschaftliche Untersuchungen an und mit vorgefundenem amateurpornographischem Material sehen sich mit forschungsethischen sowie datenschutzrechtlichen Herausforderungen konfrontiert, die Publikationen, Nachnutzungen, Sekundäranalysen, aber auch der Verwendung in Forschung und Lehre enge Grenzen ziehen.

Jegliche Forschung mit vorgefundenem Datenmaterial, das den Forschenden nicht exklusiv zur Verfügung steht, birgt prinzipiell die Gefahr, dass sich

18 Da dieses Video nicht vorliegt, konnte die geschilderte Situation freilich nicht beobachtet werden.

Studienteilnehmer:innen mittels eines Abgleichs des wissenschaftlich publizierten mit in nicht anonymisierter Form online zirkulierenden Material identifizieren lassen.[19] Diese Möglichkeit stellt namentlich Studien, die öffentlich zugängliches und nicht-veröffentlichtes Datenmaterial triangulieren, vor besondere Herausforderungen, die sich hinsichtlich des Datenschutzes von der Analyse von Interviews, von Forschenden erstellten Videographien und von kommerziellem Video- und Filmmaterial deutlich unterscheiden. Gelingt es nämlich im Fall von Triangulationsstudien einen Teil der Rohdaten – etwa *ein* hochgeladenes Video – in nicht anonymisierter Form aufzufinden, wird es nicht nur möglich, Anonymisierungen auszuhebeln, sondern Triangulierungen zu nutzen, um mittels narrativer Interviews erhobene Daten, Wissensbestände und Informationen, die sich *nicht* aus Videos erschließen lassen, aber in Publikationen einfließen und in anonymisierter Form zitiert werden, ebenso wie Forschungsergebnisse mit identifizierbaren Personen zu verknüpfen.

Wenngleich es angesichts der Menge an online zirkulierender Pornographie als unwahrscheinlich erscheinen mag, dass sich ein bestimmtes Video auffinden lässt, so erhöhen der zunehmende Zwang zu maschinell auswertbaren open-access-Publikationen einerseits und die Weiterentwicklungen von KI und Bilderkennungssoftware andererseits die beschriebenen Risiken erheblich. Folglich wachsen die Schwierigkeiten, Videos und die in ihnen dokumentierten Interaktionen so präzise zu beschreiben, dass ihre Analyse nachvollziehbar ist, zugleich aber so vage zu bleiben, dass die Anonymität der Akteur:innen gewahrt wird. Selbst bei der Publikation verpixelter Screenshots, verfremdeter Zeichnungen sowie allzu präziser Beschreibungen von Videoszenen und erst recht bei wörtlichen Zitaten verbaler Äußerungen, die Videos identifizierbar machen, ist Vorsicht geboten. Beschreibungen sollten nicht nur anonymisiert, sondern auch dadurch verfälscht werden, dass nebensächliche Aspekte und Details wie Settings, Artefakte und körperliche Merkmale, die für die Analyse keine Rolle spielen, verändert, weggelassen oder aber hinzugedichtet werden, um De-Anonymisierungen zu erschweren (vgl. Kühl 2020). Ebenso können Kamerabewegungen oder Interaktionsverläufe, auf die es nicht ankommt, in verfälschender Weise beschrieben werden. Ist ein Teil des untersuchten Datenmaterials (potenziell) öffentlich zugänglich, muss man bei der Verfälschung von Daten weiter als allgemein bzw. bei nicht-öffentlichem Material üblich gehen. Kurz gesagt: Während man aus Interviews, die im Rahmen von Forschungsprojekten geführt wurden und auf die allein die Forschenden Zugriff haben, weitgehend problemlos wörtlich zitieren kann, gilt dies für öffentlich zugängliches Videomaterial gerade nicht. Lässt sich audiovisuelles Untersuchungsmaterial jedoch nicht hinreichend anonymisieren

19 Diese Gefahr besteht aufgrund der ubiquitären Selbstdokumentation privaten Lebens im virtuellen Raum selbst dann, wenn sich nicht dieselben, sondern unterschiedliche Bilder derselben Person(en) online und in wissenschaftliche Publikationen finden lassen.

bzw. verfremden, ohne dass für die Analyse zentrale Aspekte verloren gehen, so ist gegebenenfalls auf seine Publikation und/oder die Offenlegung mancher Aspekte des Forschungsprozesses zu verzichten.[20]

Für die akademische Lehre (vgl. Wilke i. d. B.) ergeben sich ebenfalls besondere Restriktionen. Studierende sollten sorgfältig ausgewählt werden sowie Einverständnis- *und* Verschwiegenheitserklärungen hinsichtlich des (amateur-) pornographischen Materials, das sie zu sehen bekommen, unterschreiben. Um sicherzustellen, dass nur befugte Personen Zugang erhalten, und um zu dokumentieren, wer welches Material sieht, sind Anwesenheitskontrollen bei *jeder* Seminarsitzung unerlässlich. Ebenso unerlässlich sind Triggerwarnungen vor jeder Vorführung von Videos. Visuelles Datenmaterial darf Studierenden zudem ausschließlich im Seminarraum zugänglich gemacht werden, wobei zu verhindern ist, dass es mitgeschnitten oder abfotografiert wird.

Aufgrund der geschilderten Problematiken und der Tatsache, dass es sich um Material handelt, das einem hochsensiblen Bereich privater Lebensführung entstammt, der durch das Recht auf sexuelle Selbstbestimmung zudem besonders geschützt ist, ist eine Zugänglichmachung amateurpornographischer Videos zur Nachnutzung oder Sekundäranalyse ohne die explizite einzelfallspezifische Zustimmung der Studienteilnehmer:innen ausgeschlossen (vgl. Kramer/Hitzler, Reichert/Nebowsky i. d. B.). Ohnehin haben die Teilnehmer:innen des Forschungsprojekts *Die Praxen der Amateurpornographie* der Verwendung ihrer Videos, Fotos und Interviews zu ausschließlich wissenschaftlichen Zwecken nur unter der Bedingung zugestimmt, dass das Material nur Projektmitarbeiter:innen direkt zugänglich ist, nur in geschlossenen Kreisen realweltlich anwesender Wissenschaftler:innen sowie im Rahmen der akademischen Lehre vorgestellt und nur in anonymisierter Form aus ihm publiziert werden darf.

6. Ausblick

Video- und Videointeraktionsanalysen auf Basis nicht zu Forschungszwecken erstellter, sondern vorgefundener ›natürlicher‹ audiovisueller Daten entfalten vor allem für die Erforschung sozialer Felder und Praxen – wie etwa Sexualität – großes Potenzial, die Methoden der teilnehmenden Beobachtung unzugänglich sind, zu denen Akteur:innen nicht befragt werden können, weil ihre Selbstbeobachtungen zu unpräzise, sie selbst nicht vertrauenswürdig oder die Praxen, die untersucht werden sollen, nicht verbalisierbar sind. Allerdings stößt die Erforschung sexueller Praxis mittels amateurpornographischer Videos trotz vielversprechender Potenziale sowohl an forschungsethische als auch an forschungspraktische

20 Allerdings muss nicht jede Studie die videoanalytischen Methoden, derer sie sich bedient, an *jedem* Video lehrbuchartig exemplifizieren.

Grenzen. Zu Letzteren zählt insbesondere die Verfügbarkeit von Videos, die sexuelle Interaktionen in einer Weise zeigen, die ihre Erforschung erlauben, und deren Verwendung zu wissenschaftlichen Zwecken die Akteur:innen zustimmen.

Die Analyse amateurpornographischer Videos mag zwar nicht die *via regia* der Erforschung privater sexueller Interaktionen und Praxen sein. Sie ist aber ein gangbarer Weg und hat anderen Wegen voraus, dass sie es erlaubt, private sexuelle Interaktionen nicht nur präzise zu beschreiben, sondern *überhaupt* zu beobachten. Die Triangulation amateurpornographischer Videos und narrativer Interviews mit ihren Akteur:innen ermöglicht es schließlich, sexuelle Interaktionen aus mehreren Perspektiven zu untersuchen, sodass sich die Schwächen beider Methoden durch die Stärken der jeweils anderen ausbalancieren lassen.

Literatur

Bohnsack, Ralf (2011): Qualitative Bild- und Videointerpretation. Die dokumentarische Methode. 2., durchgesehene und aktualisierte Auflage. Opladen/Farmington Hills: Barbara Budrich.

Bohnsack, Ralf (2017): Fotointerpretation. In: Gugutzer, Robert/Klein, Gabriele/Meuser, Michael (Hrsg.): Handbuch Körpersoziologie. Band 2: Forschungsfelder und methodische Zugänge. Wiesbaden: Springer, S. 423–441.

Bourdieu, Pierre (1980/1993): Strukturen, Habitusformen, Praktiken. In: Bourdieu, Pierre: Sozialer Sinn. Kritik der praktischen Vernunft. Frankfurt am Main: Suhrkamp, S. 97–122.

Corsten, Michael/Krug, Melanie/Moritz, Christine (Hrsg.) (2010): Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen. Wiesbaden.

Dimbath, Oliver (2013): Visuelle Stimuli in der qualitativen Forschung. Potenziale und Grenzen des fotogestützten Interviews. In: Soziale Welt (Themenheft Visuelle Soziologie) 64, S. 137–152.

Ernst, Stefanie (2014): Literarische Quellen und persönliche Dokumente. In: Baur, Nina/Blasius, Jörg (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 829–840.

Flick, Uwe (2011): Triangulation. Eine Einführung. 3., aktualisierte Auflage. Wiesbaden: VS.

Harper, Douglas (2002): Talking about pictures. A case for photo elicitation. In: Visual Studies 17, S. 13–26.

Hillebrandt, Frank (2014): Soziologische Praxistheorien. Eine Einführung. Wiesbaden: Springer VS.

Knoblauch, Hubert (2004): Die Video-Interaktions-Analyse. In: Sozialer Sinn 5, S. 123–138.

Knoblauch, Hubert (2011): Videoanalyse, Videointeraktionsanalyse und Videographie – zur Klärung einiger Missverständnisse. In: Sozialer Sinn 12, S. 139–145.

Knoblauch, Hubert/Tuma René/Schnettler, Bernt (2010): Interpretative Videoanalyse in der Sozialforschung. Enzyklopädie Erziehungswissenschaft Online. https://doi.org/10.3262/EEO07100074.

Knoblauch, Hubert/Tuma, René (2017): Videoanalyse. In: Gugutzer, Robert/Klein, Gabriele/Meuser, Michael (Hrsg.): Handbuch Körpersoziologie. Band 2: Forschungsfelder und methodische Zugänge. Wiesbaden: Springer, S. 409–422.

Kühl, Stefan (2020): Zwischen Präzision und Anonymisierung. Wie weit muss man bei der Verfälschung wissenschaftlicher Daten gehen? In: Soziologie 49, S. 62–71.

Küsters, Ivonne (2009): Narrative Interviews. Grundlagen und Anwendungen. 2. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

Lewandowski, Sven (2012): Die Pornographie der Gesellschaft. Beobachtungen eines populärkulturellen Phänomens. Bielefeld: transcript.

Lewandowski, Sven (2021): »Roboter mit Senf«. Ironie, Irritation, Authentifizierung und Subversion in einem (semi-)amateurpornographischen Video. In: Benkel, Thorsten/Lewandowski, Sven (Hrsg.): Kampfplatz Sexualität. Normalisierung, Widerstand, Anerkennung. Bielefeld: transcript, S. 109–143.

Lewandowski, Sven (2023): Amateurpornographie und die soziologische Erforschung privater Sexualität. In: Heim, Tino/Schrage, Dominik (Hrsg.): Sexualtechnische Konsumobjekte und Metamorphosen moderner Sexualitäten. Praktiken, Beziehungsformen, Identitäten, Sozialverhältnisse. Wiesbaden: Springer, S. 445–473.

Lewandowski, Sven (2025): Sexual media practices in private homemade amateur porn. In: Porn Studies 12, *in print*.

Lewandowski, Sven/Siemer, Tara Elena (2021a): Amateurpornographische Praktiken. Ein Beitrag zur soziologischen Erforschung des Sexuellen. In: Benkel, Thorsten/Lewandowski, Sven (Hrsg.): Kampfplatz Sexualität. Normalisierung, Widerstand, Anerkennung. Bielefeld: transcript, S. 73–108.

Lewandowski, Sven/Siemer, Tara Elena (2021b): Sex und Porno im Forschungsalltag – ein Werkstattbericht über den Umgang mit sexuell explizitem Material in der empirischen Forschung. In: Zeitschrift für Sexualforschung 34, S. 89–96.

Reichertz, Jo/Englert, Carina J. (2011): Einführung in die qualitative Videoanalyse. Eine hermeneutisch-wissenssoziologische Fallanalyse. Wiesbaden: VS Verlag für Sozialwissenschaften.

Salheiser, Axel (2014): Natürliche Daten: Dokumente. In: Baur, Nina/Blasius, Jörg (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer VS, S. 813–827.

Schäfer, Hilmar (2016): Praxistheorie. Ein soziologisches Forschungsprogramm. Bielefeld: transcript.

Schmidt, Robert (2012): Soziologie der Praktiken. Konzeptionelle Studien und empirische Analysen. Berlin: Suhrkamp.

Tuma, René (2018): Video-Interaktionsanalyse. Zur Feinauswertung von videographisch erhobenen Daten. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer, S. 423–444.

Tuma, René/Schnettler, Bernt (2014): Videographie. In Baur, Nina/Blasius, Jörg (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer, S. 875–886.

Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.

Audiovisuelles Datenmaterial in der ethnographischen (Jugend-)Szeneforschung

Rapvideo-Eigenproduktionen als Zugang zu kulturellen Wissens- und Sinnzusammenhängen

Katharina Bock

Zusammenfassung: Der Beitrag widmet sich der Bedeutung von audiovisuellen Artefakten in der ethnographischen (Jugend-)Szeneforschung. Dabei werden Musikvideo-Eigenproduktionen Gangsta-Rap-begeisterter Jugendlicher fokussiert. Diese werden als ein wichtiger Zugang zu szenespezifischen Wissens- und Sinnzusammenhängen und damit zu zentralen kulturellen Themen einer spezifischen sozialen Sinnwelt betrachtet. Das Beitragsthema wird am Beispiel einer Musikvideo-Analyse entfaltet, in deren Zusammenhang verschiedene Möglichkeiten der Datenaufbereitung, -analyse und -darstellung eröffnet werden. Dazu gehört eine ethnosemantische Analyse des Songtextes, die Videotranskription sowie eine ethnographische Beschreibung des audiovisuellen Artefakts samt Explikation und Interpretation.

Schlüsselwörter: Ethnographie, Jugendszenen, Gangsta-Rap, Musikvideos

1. Nutzungsszenario: Audiovisuelle Artefakte in der ethnographischen (Jugend-)Szeneforschung am Beispiel *Gangsta-Rap*

Ethnographische Forschungsstrategien zeichnen sich durch die Kombination vielfältiger Verfahren der Datenerhebung und -analyse aus. Als Bestandteil eines größeren Korpus' können Artefaktanalysen mit Teilnahme-, Beobachtungs- und Gesprächsprotokollen, mit Interviewtranskripten, Felddokumenten und Ergebnissen weiterführender Recherchen verwoben werden, um verschieden(artig)e Perspektiven zu generieren oder diese anzureichern und so zu deren Verdichtung beizutragen. »Gerade weil sich der Alltag und das soziale Zusammenleben nicht von Artefakten, ihrer Erzeugung, Verwendung und ihrer Bedeutung für das Denken, Fühlen und Handeln abtrennen lässt, bietet der systematische Einbezug von Artefakten in ethnographische Studien eine wichtige Bereicherung« (Lueger/Froschauer 2022, S. 527). Artefakte fungieren als »Wissens- und Kompetenzspeicher« (ebd., S. 522); sie »generieren Sinn und schaffen Muster« (ebd.,

S. 524). Hitzler und Eisewicht (2016, S. 53 f.) heben in diesem Zusammenhang die Bedeutung von im Feld hergestellten Dokumenten und Artefakten als für den Erkenntnisprozess wichtige Zugänge zu Wissensbeständen und Sinnzusammenhängen hervor, die unbedingt berücksichtigt werden sollten.

Eine ethnographische (Jugend-)Szeneforschung,[1] die audiovisuelle Eigenproduktionen der Feldteilnehmenden bewusst einbezieht, trägt nicht nur der immensen Bedeutung von Medien im Alltag junger Menschen Rechnung. Audiovisuelle Erzeugnisse vermögen komplexe Bewegungs- und Handlungsabläufe abzubilden und ermöglichen den Produzierenden/Protagonist:innen, im Zusammenspiel von Bildern, Musik und Sprache, Gefühlen und Stimmungen Ausdruck zu verleihen sowie Eindrücke und Erfahrungen symbolisch zu verarbeiten (vgl. Niesyto 2018, S. 782; Witzke 2005, S. 323). Solche selbst hergestellten Artefakte gestatten vertiefte Einblicke in die Lebens- und Medienwelten der Feldteilnehmenden, in die Wahrnehmung des eigenen Selbst und die Eindrücke aus der Umwelt (vgl. Niesyto, S. 783). Dabei bringen die Protagonist:innen bedeutsame und vielfach entwicklungsrelevante Aspekte zum Ausdruck, die sich der sprachgebundenen Artikulation entziehen (vgl. Witzke 2005, S. 323). »Jugendliche […] verfügen über ein beträchtliches […] und zum Teil nicht bewusstes mediensprachliches Wissen«, das sie »in der Lage sind […] zu ›lesen‹ und mit subjektiven Sinngebungen zu versehen. […] Damit bieten Video-Eigenproduktionen Jugendlichen nicht nur die Möglichkeit zur Verarbeitung eigener Medienerfahrung, sondern es werden neue und zusätzliche Wege des Ausdrucks und der Kommunikation eröffnet, die auch forschungsmethodisch von großem Interesse sind« (ebd., S. 325).

Einen interessanten wie gehaltvollen Anwendungskontext für audiovisuelle Artefaktanalysen im Rahmen ethnographischer (Jugend-)Szeneforschung bietet die Sinnwelt rund um das musikalische Sub-Genre *Gangsta-Rap*. Dieser ist vor allem unter jungen Menschen beliebt und verbreitet (Bundesverband Musikindustrie 2022, S. 28). Deren Perspektive auf das Genre wird jedoch eher selten thematisiert.[2] In diesem Zusammenhang sind Fragen zu Umgangs- und Aneignungsweisen des Genres sowie mit der Nutzung verbundene Sinn- und Bedeutungszusammenhänge weitgehend ungeklärt. Für die Bearbeitung dieser Forschungslücken sind ethnographische Forschungsstrategien besonders geeignet, da sich mit ihnen die situationalen und sozialen Kontexte untersuchen lassen, in denen jene Wissensbestände, Bedeutungen und Relevanzen hervorgebracht werden, die für das Verstehen des Umgangs mit der populären Musik und Kultur zentral sind (Cohen 1993, S. 135). Vor eben diesem Hintergrund widmen sich

1 Zum (Jugend-)Szene-Begriff vgl. Hitzler/Niederbacher (2010).

2 Die meisten Untersuchungen zum Genre thematisieren die Perspektive populärer Profi-Rapper:innen. Dort wird Gangsta-Rap überwiegend als Reaktion auf und die Be-/Verarbeitung von gesellschaftlichen Missständen und sozialer Benachteiligung betrachtet (u. a. Seeliger/Dietrich 2017; Seeliger 2021).

einige neuere Untersuchungen der Auseinandersetzung junger Menschen mit Gangsta-Rap im Rahmen musikalischer (Video-)Eigenproduktionen. In diesen Untersuchungen werden teilnehmende (Selbst-)Beobachtungen, ethnographische Gespräche und Interviews mit der Analyse audiovisueller Artefakte aus dem Feld kombiniert. Dabei zeigt sich u. a., *dass* und *wie* diese jungen Menschen aus einer breiten Palette von Stilisierungs- und Identifikationsangeboten populärer Protagonist:innen der (inter-)nationalen Rap-Szene auswählen und diese auf eigen(sinnig)e Weise adaptieren; sie experimentieren mit verschiedenen Themen und setzen diese mit ihrer persönlichen lebensweltlichen Situation in Zusammenhang, wobei sie zum Teil fiktive Realitäten entwerfen und sich das Genre auf vielfältige Weisen kreativ zu eigen machen (u. a. Bock 2024; Bock 2025a; Schmidt/Bock 2024).

Hier knüpft der Beitrag an und widmet sich audiovisuellen Artefakten als wichtigem Zugang zu szenekulturellen Wissens- und Sinnzusammenhängen. Am Beispiel eines selbstproduzierten Musikvideos junger männlicher Gangsta-Rap-Anhänger werden Möglichkeiten der Datenaufbereitung und -analyse eröffnet, die zum Verständnis zentraler kultureller Themen einer spezifischen sozialen Sinnwelt beitragen.

2. Erhebungskontext

2.1 Forschungsstrategie und Feldzugang

Das im Rahmen dieses Beitrags analysierte Musikvideo habe ich während meiner Mitarbeit am DFG-geförderten SFB 1472 »Transformationen des Populären« der Universität Siegen, im Teilprojekt A03 »Inszenierungen von Kriminalität: Gangsta-Rap in interaktiven Identitätspraktiken Jugendlicher« (2021–2024), gewonnen. Das Projekt verfolgte eine ethnographische Forschungsstrategie, die teilnehmende Beobachtungen, ethnographische Gespräche,[3] umfangreiche Recherchen sowie die Sammlung und Analyse kultureller Artefakte miteinander kombiniert (v. a. Dellwing/Prus 2012; Breidenstein et al. 2015; Poferl/Schröer 2022) und um audioaufgezeichnete Interviews ergänzt hat.

Mein Zugang zum Untersuchungsgegenstand erfolgte über großstädtische Einrichtungen der Offenen Jugendarbeit, in denen Hip-Hop-Kultur bzw.

3 Ethnographische Gespräche dienen dazu, das durch Teilnahme, Beobachtung, aus Recherchen und Artefaktanalysen erworbene Wissen zu hinterfragen, zu reflektieren und etwaige Verständnisfragen (spontan) klären zu können. Im Unterschied zu klassischen Interviews betonen ethnographische Gespräche den informellen Charakter von Konversationen im Feld. Solche, in Umgangssprache geführten, Unterhaltungen ohne Hierarchien und Rollenzuweisungen sind erfahrungsgemäß essenziell für das Verständnis dessen, was im Feld vor sich geht (vgl. Dellwing/Prus 2012, S. 114 ff.).

Rap-Musik bedeutsam waren. Diese wurden telefonisch kontaktiert und persönlich aufgesucht, um vor Ort belastbare Forschungsbeziehungen zu Gangsta-Rap-affinen jungen Menschen aufzubauen.

Das Video stammt von männlichen jugendlichen Insidern der Rap-Szene, die Gangsta-Rap nicht nur gerne hören, sondern diesen autodidaktisch und laienhaft[4] auch jeweils selbst herstellen; sie schreiben eigene Rap-Texte und nehmen ihre Songs mit der Unterstützung von Sozialarbeitenden in einem Jugendzentrum auf. Zu einigen ihrer musikalischen Eigenproduktionen erstellen sie auch Musikvideos, die sie dann auf YouTube und via Instagram an ihre Freundes- und Bekanntenkreise verbreiten.

Im Mittelpunkt des im Folgenden analysierten Rap-Videos stehen die Jugendlichen Eduardo und Lorenzo, die ihren Song »Keine Miene verziehn« performen. Zum Zeitpunkt der Videoveröffentlichung auf YouTube sind die beiden 14 und 16 Jahre alt. Forschungsethischen Prinzipien und datenschutzrechtlichen Bestimmungen folgend, ist das Datenmaterial anonymisiert worden. Um eine mögliche Identitätsrekonstruktion der jugendlichen Urheber zu verhindern, wird aus dem Songtext nur ausschnitthaft zitiert.

2.2 Soziokultureller Kontext: Gangsta-Rap und die Kategorie »Straße«

Bezieht man audiovisuelle Artefakte in ethnographische Forschungsstrategien ein, so gilt zu berücksichtigen, unter welchen Umständen und zu welchem Zweck diese im Feld produziert und gehandhabt werden. Adressiert ist damit ein Erfassen von Bedeutungszuschreibungen, Zeitbezügen, situationalen und sozialen Kontexten sowie motivationalen Hintergründen der Entstehung und Nutzung (vgl. Lueger/Froschauer 2022, S. 521 f.; Hitzler/Eisewicht 2016, S. 53 f.; Cohen 1993, S. 135).

Wichtige Ausgangspunkte für die Eigenproduktion von Gangsta-Rap bilden ein spezifisches Interesse, Freude und Vergnügen sowie das Bedürfnis nach kreativem und künstlerischem (Selbst-)Ausdruck. Motivierend scheint außerdem die Aussicht auf Berühmtheit, Geld und Status. Zudem erweisen sich die Rap-Eigenproduktionen als eine Möglichkeit der Verarbeitung von alltäglichen Erfahrungen auf der »Straße« (hierzu ausführlicher: Bock 2024).

Die Bedeutung der Kategorie »Straße« ist für das Verständnis des im Beitrag analysierten Musikvideos zentral, denn: »Gangsta-Rap ist Straßenrap«, wie es im Feld heißt. »Straße« ist großstädtische Öffentlichkeit und der unmittelbare Lebens- und Erfahrungsraum der Insider. Hier werden Betrachtungs- und Verhaltensweisen erworben, die (auch) in den Produktionen populärer

4 Lai:innen verfügen über keine spezifische Qualifikation und werden für ihre Tätigkeit nicht bezahlt (vgl. Schützeichel 2007, S. 547).

Profi-Rapper:innen artikuliert und bestätigt werden. Seit der Entstehung von Rap im New York der 1970er Jahre sind Ereignisse auf der »Straße« – vor allem »gewalttätige Auseinandersetzungen« sowie »Formen der Kriminalität, die für Gangs typisch sind« – ein wichtiger Bezugspunkt der [Rap-]Szene (Schröer 2012, S. 66 f.). Inzwischen hat sich in diesem Zusammenhang der Begriff »Drill-Rap« etabliert, den auch viele meiner Forschungsteilnehmenden favorisieren. Soundbezogen zeichnet sich »Drill« u. a. durch technisierte und minimalistische Beats und einen zurückgelehnten Flow[5] aus (vgl. Tonspion.de 2021). Inhaltlich steht »Drill« für Rap, der die harte und besonders gewalttätige Seite der »Straße« thematisiert (dazu ausführlich: Terms & Conditions 2020). Entsprechend werden »Drill«-Texte als »nihilistic, provocative and violent« (Monteith 2022, S. 546) erachtet.

Darüber hinaus ist »Straße«, wie es Kadeem (einer meiner Gesprächspartner im Feld) beschreibt, ein »Mindset«, das einen prägt und begleitet. Auf der »Straße« gehe es vor allem darum, »sich vor anderen zu profilieren und sozusagen ›den Macker‹ [im Sinne von Anführer, Macher] zu machen«. Es sei hauptsächlich »dieses Macho-Ego-Prinzip: ›ich bin der Alpha hier und ihr seid die Betas‹«, um das es gehe. Und auch im »Rap Game« – hier schlägt Kadeem die Verbindung von der »Straße« zum Gangsta-Rap – gäbe es »einfach diese Sache mit Status und Gesicht«. Es gehe darum, »der größte MotherFUCKer« zu sein, sich »krass« zu fühlen und das zum Ausdruck zu bringen. Hierhinter verbirgt sich eine Notwendigkeit, der »Straße« und ihren Themen gewachsen und imstande zu sein, dies glaubhaft zu verkörpern. Um die hierfür notwendigen Gefühls- und Verhaltenssicherheiten zu erlangen, gilt es, das »Mindset« zu verinnerlichen. Die Partizipation am »Rap Game« stellt für die jungen Szene-Insider eine passende Möglichkeit dar, das eigene Selbst im Zusammenhang mit den Themen der »Straße« spürend zu erfahren und dies auch für andere glaubhaft darzustellen (dazu ausführlicher: Bock 2024; Bock 2025a).

3. Analysekontext

Im Folgenden werden Aufbereitungsweisen und verschiedene analytische Zugangsmöglichkeiten zum audiovisuellen Material vorgestellt. Dies beinhaltet eine ethnosemantische Kurzanalyse (Maeder/Brosziewski 1997) des Songtextes (Abschnitt 3.1), die Videotranskription (Korte 2005) sowie Möglichkeiten zur Auswertung des Materials (Bullerjahn 2005; Bienk 2008; Denzin 2010) (Abschnitt 3.2). Die analytischen Einzelschritte werden anschließend in eine

5 *Flow*, wie ich im Feld lerne, beschreibt das Passungsverhältnis zwischen dem Rhythmus eines Beats und dem Zusammenspiel von Worten, der Betonung einzelner Silben, Pausen, dem Sprechtempo und der Sprachmelodie (dazu auch: Kautny 2009).

ethnographische Beschreibung und Explikation des Rap-Videos überführt (Abschnitte 3.3 und 3.4).

3.1 Song-Analyse

»Beim Musikvideo«, wie Bullerjahn (2005, S. 489) ausführt, »stehen Musikstil, Songtext, Image des Interpreten sowie Darstellungstypen in Wechselbeziehung mit den zentralen visuellen Darstellungsinhalten und befinden sich auch untereinander in reziproker Verbindung«. Diese Elemente werden im Folgenden entfaltet.

Zur Analyse des Songtextes, wie auch zur Analyse des übrigen ethnographischen Materials, eignet sich die Methode der Ethnographischen Semantik (Maeder/Brosziewski 1997). Sie ermöglicht es herauszuarbeiten,

> »auf welche Art und Weise die Angehörigen einer Kultur ihre Welt kategorisieren, bzw. wie sie ihre Welt mittels alltäglich-praktischer Klassifikationen definieren und sich einander vermitteln. Welche Bedeutungen sie den Handlungen und Objekten in ihrer Kultur zuschreiben (können), wird aus dem gelernten, kategorialen Haushalt der geteilten und primär in der Sprache aufgehobenen Wissensbestände der Mitglieder einer Szene, eines Milieus oder auch einer noch umfassenderen Kultur rekonstruiert. Das Produkt solcher Rekonstruktionen sind gewissermaßen ›kulturelle Landkarten‹ über verfügbare Elemente in einer bestimmten sozialen Sinnwelt« (ebd., S. 344).

Im Rahmen dieser Analyse werden schrittweise einzelne (Sprach-)Kategorien identifiziert und zu Domänen entwickelt, die den Elementen einer sozialen Situation entsprechen: *Personen*, *Orte/Räume*, *Handlungen/Aktivitäten* und *Dinge* (vgl. Maeder 2002, 8. Absatz; Maeder 2007, S. 684 f.).[6] Auf diesem Weg können ein Personenrepertoire der Sinnwelt erstellt sowie bedeutsame Bezugsräume, Aktivitäts-/Handlungsrepertoires und Dingwelten ermittelt werden. Da Dinge sowohl *physikalisch* als auch *abstrakt* sein können (vgl. Blumer 1981, S. 90), kann diesbezüglich nach Gebrauchsgegenständen, Artefakten, Konsummitteln oder abstrakten Dingen (wie Vorstellungen, Konzepte, Haltungen, Regeln usw.) kategorisiert werden. Setzt man die Domänen miteinander in Bezug, so ergibt sich ein »Image« (Goffman 1971, S. 10 ff. sowie Bullerjahn 2005, S. 489), das die beiden jugendlichen Interpreten im Drogen- und kriminellen Milieu verortet und als gewaltbereite »Drill«-Rapper ausweist. Mit ihrem Songtext demonstrieren

6 Im weiteren Forschungsprozess werden diese Domänen schrittweise immer weiter ausdifferenziert (vgl. Maeder/Brosziewski 1997, S. 351 f.). Dabei werden Zu-, Über- und Unterordnungen vorgenommen, Reihenfolgen, Verweise, Inkludierungen, Abgrenzungen und Kontrastierungen berücksichtigt sowie alle möglichen weiteren Zusatzinformationen zusammengetragen.

die beiden ihre Zugehörigkeit zu einer Sinnwelt, die von normverletzenden Verhaltensweisen (Drogenkonsum, Drogenkriminalität, Gewalthandlungen) und eigensinnigen Regeln (»Keine Miene verziehn«; keine Angst zeigen; unerkannt bleiben) geprägt ist (vgl. Tab. 1).

Tabelle 1: Domänen im Text zum Song »Keine Miene verziehn«

Personen	Orte/Räume
– Olle, Bitch, hässlige Hoes – Chicas – Du; deine Freunde – wir, uns – der Junkie – Ticker (an Ecken) – ein Mann (mit Tattoo im Gesicht) sein	– am Block – auf Telegram; bei Zalando & Co. (online) – in meiner Gegend – Straße; an [Straßen-]Ecken – Gericht – Traphouse (= Ort, an dem abgehangen wird, um Alkohol und Drogen zu konsumieren; Urbandictionary.com o. J. a)
Handlungen/Aktivitäten	**Dinge**
– am Block chillen (= entspannt abhängen) und ticken (= Drogen verkaufen) – drohen, jmdn. zu ficken (= besiegen, betrügen oder Gewalt aussetzen) – jmdn. laufen schicken (= Drogen verkaufen lassen) – Koks auf K[ommission] holen – Kartell aufbauen – Standi (= ›Standardzeug‹) rauchen – auf Xanax sein; high sein – abgehoben sein – ängstlich wie Bugs Bunny sein – einen auf Rambo machen – Gewalthandlungen in Aussicht stellen: »Du wirst meine Faust schmecken«; »Dich schlag ich K. O.«; »Leberstich« – Unterstellungen/Behauptungen: »Deine Leute wollen mein Flex«; »Deine Olle will mein Dick lecken« – in Rage sein – Courage zeigen – über Straße reden – über Lambo[rghini] und Millionen sprechen – bei Zalando & Co. bestellen	– Jogginganzug – Xanax (= Pharmazeutikum mit beruhigender, entspannender, angstlösender Wirkung) – Flex (= Kokain) – Drogen – Millionen – Lambo[rghini] – Koks **abstrakte Dinge** – Bugs Bunny – ein Stück vom Kuchen wollen – die Vorstellung, jmdn. abzustechen – keine Miene verziehn – im Jogginganzug bei Gericht erscheinen – lesh (= beschissen; Urbandictionary.com o. J. b) aussehen – Regeln (befolgen): keine Angst zeigen; unerkannt bleiben; keine Spuren hinterlassen

Der Song besteht aus zwei sogenannten »Parts« (Strophen) und einer sich mehrmals wiederholenden »Hook« (Refrain), die jeweils durch eine überwiegende Aneinanderreihung von Hauptsätzen mit Endreimen nach dem Muster *aaaa, bbbb, cccc* usw. gekennzeichnet sind. Im Rahmen ihrer Narration richten sich die beiden Insider an eine signifikante andere Person, die sie duzen. Wie nah sie dieser Person stehen, ist unklar. Dies gilt auch für Geschlecht und Alter der Person; einige wenige Hinweise lassen vermuten, dass die Person heterosexuell-männlich

und (etwas) älter ist als die beiden Jugendlichen. Eduardo und Lorenzo nehmen die Person als großspurig und angeberisch wahr. Sie mögen die Person offensichtlich nicht und wollen sich von ihr unterscheiden und abgrenzen. In diesem Zusammenhang machen sie von allerlei Genre-typischen Übertreibungen, Unterstellungen, Behauptungen, Abwertungen und Drohungen Gebrauch: So sei die Person ängstlich (wie »Bugs Bunny«), rauche (nichts Hochwertiges, sondern lediglich) herkömmliches Standardzeug (»Standi«); die Person umgebe sich mit »Ollen«, »Bitches« und »hässligen Hoes«; hinzu kommt die Unterstellung »Deine Bitch will mein Dick lecken«. Die Person hat zudem offenbar ein Tattoo im Gesicht, was die Jugendlichen wohl zwar grundsätzlich mit Männlichkeit assoziieren, die sie der Person aber zugleich absprechen (»Du bist kein Mann nur mit Tattoo im Gesicht«). Scheinbar hat die Person Gewalthandlungen gegen die Jugendlichen in Aussicht gestellt (»Rede nicht, […] dass du ein von uns haust«). Darauf reagieren Eduardo und Lorenzo in ihrem Song und drohen nun ihrerseits mit Gewalt (»Du wirst meine Faust schmecken«) sowie damit, die Person abzustechen (»Leberstich«). Dabei imaginieren die beiden Jugendlichen ein etwaiges juristisches Nachspiel ihrer Gewalthandlungen und malen sich aus, wie sie »in Jogginganzug« im »Gericht« erscheinen und dabei »keine Miene verziehen« würden. Damit verweisen sie einerseits auf für das Genre typische Normverletzungen. Andererseits erfolgt hier der Hinweis auf ein in der Rap-Szene bedeutsames Kleidungsstück: Der Jogginganzug wurde »[i]n den späten 1980ern und frühen 1990ern […] durch legendäre Musiker wie Run-DMC, Biggie Smalls und Dr. Dre« geprägt und markiert einen »Streetwear-Hype«, der »bis heute anhält« (Hainz 2022).

3.2 Videotranskription und -analyse

Die Analyse des Videos zum Song beginnt mit einer »möglichst unvoreingenommene[n] Sichtung des audiovisuellen Gesamtprodukts […], wobei diese erste gefühlsbasierte Erfahrung im Nachhinein stichwortartig schriftlich fixiert werden sollte« (Bullerjahn 2005, S. 485). Wie darüber hinaus Denzin (2010, S. 427) ausführt, geht es hierbei darum, den Wirkungen des Materials nachzuspüren und erste Empfindungen und Eindrücke zu notieren.

Da Videos zum Teil »hochkomplexe Aussagesysteme« (Korte 2005, S. 387) darstellen können, ist es sinnvoll ihrer Untersuchung eine systematische Transkription an die Seite zu stellen. Dazu wird ein Sequenz- oder Einstellungsprotokoll[7] angefertigt. »Beide Transkriptionsarten […] dienen gleichermaßen der

7 Eine »Einstellung bezeichnet die kleinste Filmeinheit, die durch zwei Schnitte oder Blenden begrenzt wird« (Bienk 2008, S. 52). Exemplarische Einstellungsprotokolle finden sich z. B. bei Korte (2005, S. 389) oder Bock (2017, S. 164–167) sowie Bock (2018, S. 161–164).

Sicherung des Untersuchungsgegenstands […]. Während der Informationsgehalt des erheblich aufwändigeren und i. d. R. sehr umfangreichen Einstellungsprotokolls im Rahmen einer qualitativen Analyse meist nur bedingt genutzt werden kann, bietet das deutlich kürzere Sequenzprotokoll (vgl. Tab. 2) eine handhabbare Übersicht« (ebd., S. 388) über die Länge und zeitliche Abfolge unterscheidbarer Informationseinheiten (Sequenzen), die z. B. räumlich, zeitlich, thematisch oder szenisch zusammengehören. Die systematische Transkription »zwingt zur genauen Beobachtung und eröffnet damit die Möglichkeit, den Untersuchungsgegenstand weitaus besser kennen zu lernen als bei einer auch mehrfach reinen Betrachtung« (ebd., S. 387 f.).

Tabelle 2: Sequenzprotokoll zum Rapvideo »Keine Miene verziehn«

Zeit	Sequenz	Wissensbestände	
		Bildebene	**Tonebene**
00:00–00:15	Erste Sequenz (E1-E4)	– Personen: männliche Jugendliche (junge, weiche Gesichtskonturen; leichter Oberlippenbart); männlicher Freundeskreis; ernste Gesichtsausdrücke – Orte/Raumelemente: (groß-)städtischer Raum – Handlungen/Aktivitäten: Begrüßungsritual; Rauchen; Rappen – Dinge: Markenkleidung und Accessoires – Pit Bull Terrier	unverständliche Worte; Xylophon-Klänge; verzerrte Sirenengeräusche; unregelmäßige, flache und leicht blecherne Beat-Schläge
		Slow-Motion; Einblendung der Namen und des Songtitels »Keine Miene verziehn«	
		Darstellungstyp: Konzept	
		leichter Sonnenschein, blau-grauer Himmel mit vielen Wolken; reduzierte Leuchtkraft der Farben	
00:16–02:48	Hauptsequenz (E5-E111)	– Personen: männliche Jugendliche; männlicher Freundeskreis; ernste Gesichtsausdrücke – Orte/Raumelemente: (groß-)städtischer Raum, größtenteils menschenleer; Wohnviertel in der Nähe einer Schnellstraße; Treppen; Betonelemente; Plattenbauten; Bolzplatz; Überführung einer Schnellstraße/Stadtautobahn; Platz mit Denkmal; Hauseingang; U-Bahn-Eingang – Handlungen/Aktivitäten: auffällig hoher Zigarettenkonsum; Körperbewegungen, die Gewaltgesten nachempfinden – Dinge: Markenkleidung und Accessoires; iPhone; Zigaretten; Graffitis – Pit Bull-Terrier	technisierte minimalistische Beats (Trap- bzw. Drill-Sound)
		Normalgeschwindigkeit	
		Darstellungstyp: Konzept	
		wolkenverhangener Himmel; Farben ohne Leuchtkraft und Kontraste	

In Abhängigkeit vom jeweiligen Erkenntnisinteresse[8] werden Auffälligkeiten und Besonderheiten protokolliert, etwa hinsichtlich *Bildinhalten* (Personen, Ort/e, Raumelemente, Handlungsabläufe, Dinge[9]), der *Bildgestaltung* (Einstellungsgrößen, Kameraperspektiven, Kamerabewegungen), der *Gestaltung von Licht und Farben* sowie hinsichtlich der *Tongestaltung* (Geräusche, Sprache, Musik) (vgl. Bienk 2008). Auch Darstellungstypen lassen sich hier festhalten. Typisch für viele Musikvideos sei eine Kombination mehrerer Darstellungstypen, wobei grundsätzlich drei Typen unterschieden werden können: bei der 1) *Performance* stehe die musikalische Aufführung und die Musizierenden im Mittelpunkt (z. B. Live-Konzert-Mitschnitt); bei einer 2) *Konzept-Performance* finde das Musizieren in Räumlichkeiten statt, die nichts mit musikalischen Darbietungen zu tun haben; beim 3) *Konzept* habe die Darstellung einen Bezug zum Song und/oder dem Image der Musizierenden, gegebenenfalls agieren diese als Schauspieler:innen und tragen zur abstrakten Visualisierung bei (Bullerjahn 2005, S. 489).

Im Gegensatz zum hier analysierten Video existieren natürlich auch erheblich *filmischer angelegte* Materialien. Diese verfügen häufig über deutlich mehr Sequenzen und eine weitaus komplexere Struktur. Je nach Erkenntnisinteresse und Komplexität der Filmsprache kann es dann sinnvoll sein, kleine Protokolle zu den identifizierten Einzelsequenzen anzufertigen und diese später zu einer detaillierten Beschreibung der Filmsprache auszubauen (vgl. Bienk 2008). Dies kann z. B. in Anlehnung an die von Norman K. Denzin vorgeschlagenen »Prinzipien einer kritischen visuellen Analyse« (Denzin 2010, S. 426) erfolgen. Dabei handelt es sich um »Leitlinien«, die sich in vier Phasen aufteilen und den jeweiligen individuellen Forscher:innenbedürfnissen anzupassen sind: Phase eins widmet sich dem »*Sehen und Fühlen« des Materials*; Phase zwei gilt der *Formulierung von Forschungsfragen*; Phase drei umfasst die »[S]*trukturierte Mikroanalyse*«, d. h. das Erfassen und detaillierte Beschreiben von Szenen und Sequenzen; Phase vier widmet sich der »*Suche nach Mustern*« (ebd., S. 426 f.; Herv. i. O.).[10]

3.3 Beschreibung des Rap-Videos

Die vorausgegangenen analytischen Einzelschritte werden nun in eine detaillierte Beschreibung des Rap-Videos überführt. Das Musikvideo ist 2:48 Minuten lang und lässt sich in eine Einführungs- und eine Hauptsequenz unterteilen. Die Hauptsequenz enthält 106 Einstellungen. Im Mittelpunkt des Videos

8 Dies gilt auch für den Grad der Differenzierung.

9 Hier zeigen sich Ähnlichkeiten zur Domänenanalyse der Ethnographischen Semantik (siehe Abschnitt 3.1).

10 Eine solche Analyse findet sich z. B. in Bock (2025b).

stehen Eduardo (rappt »Part« 1) und Lorenzo (rappt »Part« 2 und die »Hook«), die ihren Song »Keine Miene verziehn« performen. Das Video beginnt mit einer Halbtotalen in Slow-Motion. Lorenzo befindet sich in der Bildmitte. Er trägt schwarze Jeans, einen gesteppten schwarzen Parker mit Kapuze, schwarze Sportschuhe und ein schwarzes Basecap mit dem Logo des NBA-Basketball-Teams »New York Yankees«. Unter dem linken Arm hält er zwei weiße Schuhkartons. Mit seiner rechten Hand führt er einen grauen Pit Bull-Terrier an der Leine. Im Hintergrund zeigt sich ein blau-grauer Himmel mit vielen Wolken; die Sonne steht tief und wirft lange Schatten. Die Bäume im Hintergrund sind nahezu kahl – es ist vermutlich später Herbst. Die Leuchtkraft der Farben ist deutlich reduziert. Grauer Beton dominiert das Farbspiel des Bildausschnitts. Die Sequenz wirkt eher kalt und ungemütlich. Auf Tonebene sind einige unverständliche Worte eines männlichen Sprechers vernehmbar; einzelne Xylophon-Klänge und verzerrte Sirenengeräusche mischen sich unter rhythmisch unregelmäßige, flache und leicht blecherne Beat-Schläge. Die Namen der Interpreten und der Songtitel werden eingeblendet. Mit cool wirkendem Gang läuft Lorenzo sechs flache Treppenstufen hinab. Schnitt. Unten trifft er auf seinen Freund Eduardo. Die beiden begrüßen einander mit einem Handschlag, der in eine kurze Umarmung übergeht, die zwar vertraut, insgesamt allerdings eher mechanisch als herzlich wirkt. In Eduardos Mund steckt eine Filterzigarette. Auch er trägt ein schwarzes Basecap; dazu einen gesteppten schwarzen Parker mit einem auffällig großen Echtfellkragen und einem Emblem der Marke »Canada Goose«, die für hochwertige und hochpreisige Outdoor-/Funktionskleidung steht. Lorenzo präsentiert Eduardo seine Schuhkartons, die den Schriftzug der Marke »Moncler« tragen. In späteren Einstellungen der Hauptsequenz werden die Jugendlichen weitere Marken und Statussymbole präsentieren, u. a. ein neu verpacktes »iPhone« sowie weitere Schuhkartons der Marken »Versace« und »Leandro Lopes«. Schnitt. Zwei weitere jugendliche Personen werden eingeführt, Chris und David. Vermutlich sind sie mit Lorenzo und Eduardo befreundet. Die beiden rappen nicht. Sie halten sich eher im Hintergrund und sollen die Szenen vermutlich nur personell anreichern. Auch diese beiden Jugendlichen tragen schwarze Kleidungsstücke. Alle Beteiligten zeigen wenig mimisches Spiel und verkörpern damit eindrücklich das Prinzip »Keine Miene verziehn«. Schnitt. Zum erweiterten Ensemble gehört auch der erwähnte Pit Bull-Terrier. Pit Bulls sind sogenannte Listenhunde und gelten als potenziell gefährlich, weshalb für sie ein Maulkorb- und Leinenzwang in der Öffentlichkeit gilt. Der Hund trägt ein auffälliges schwarzes Lederhalsband, besetzt mit weißen Strasssteinen, die auf den ersten Blick allerdings wie Nieten aussehen. In einer sehr eindrücklichen Einstellung (E4) sind seine Vorderpfoten auf einer Art Reling platziert, wodurch es beinahe wirkt als würde das Tier stehen wie ein Zweibeiner. In dieser Körperhaltung kommt der überaus muskulöse Tierkörper besonders zur Geltung. Zunächst von der Kamera abgewandt, dreht der Hund

plötzlich seinen Kopf. Sein strenger, starrer Blick trifft direkt auf die Kamera. Eine ruckartige Zoombewegung wirkt, als würde der Hund einen Satz nach vorne machen. Dieser Eindruck wird auditiv verstärkt durch drei kraftvolle, aufeinanderfolgende Beat-Schläge. Der kalt wirkende Grauton des kurzen Fells lässt den Hund unnahbar wirken. Dieser Eindruck wird verstärkt durch die Beschaffenheit der Schnauze, die an hängende Mundwinkel grimmiger, besonders schlecht gelaunter Menschen erinnert. Auf dieses Wissen scheinen die Jugendlichen hier zu referieren. Der Hund wird als machtvolles Wesen inszeniert, das Härte, Kraft und Strenge ausstrahlt und so die Selbstinszenierung der beiden als gewaltbereit und Respekt einflößend stützt.

Timecode 00:16 – die Aufnahme wechselt von Slow Motion in die Normalgeschwindigkeit und markiert den Beginn der Hauptsequenz. Der Raptext setzt ein. Besonders eindrücklich wirkt in diesem Zusammenhang die Verkörperung der Zeile »Zack, rein, Leberstich«, die als Teil der »Hook« mehrfach wiederholt und dabei durchweg von einer Körpergeste begleitet wird, die das Zustechen in einem Messerkampf imitiert. Die Hauptsequenz ist von vielen schnellen Schnitten und etlichen, sich wiederholenden Einstellungen gekennzeichnet, deren Reihenfolge wahllos erscheint. In diesem Zusammenhang wechseln sich Bildhintergründe ab, die verschiedene Orte mit urbaner Rahmung präsentieren. Zunächst sitzen Lorenzo und Eduardo rappend auf einer Tischtennisplatte, die Teil eines Bolzplatzes ist, der sich in einer Wohngegend mit mehrstöckigen Mietshäusern (Plattenbauten) befindet. Weitere Einstellungen präsentieren den Eingangsbereich zu einem dieser Mietshäuser, einen U-Bahneingang, eine großflächige bronzene Gedenktafel sowie einen schneckenartig angelegten Betonkomplex, der Personen zur Überwindung einer innerstädtischen Schnellstraße sowie eines parallel dazu verlaufenden Flusses mit Binnenschifffahrt dient. Dieser Überführungskomplex besteht scheinbar aus drei Ebenen, womit den Jugendlichen verschiedene Einzelsettings bereitstehen, die sich ständig wiederholen. Viele Bereiche dieser Betonanlage bieten kleinen und großflächigen Graffiti-Schriftzügen Platz. Der Himmel ist grau und wolkenverhangen. Die Farbgestaltung ist wenig kontrastreich und weist nur eine geringe Sättigung auf. Sämtliche Einstellungen der Hauptsequenz sind geprägt von kalt wirkenden Farben und Materialien. Grau- und Beige-Töne dominieren.

3.4 Explikation und Interpretation des Rap-Videos

Elaborierte ethnographische Beschreibungen enthalten neben der reinen Darstellung auch weitere Explikationen sowie eine Interpretation des Gegenstands (u. a. Bock 2019). Dem folgend lassen sich die im Rap-Video präsentierten Inhalte wie folgt charakterisieren: Die gewählten Drehorte stellen vermutlich den unmittelbaren Lebens- und Erfahrungsraum der Jugendlichen dar – geprägt

von Normverletzungen und Gewaltthematiken, von Urbanität, Eintönigkeit und Tristesse. Solche Räume sind »ein immer wieder auftauchender Bezugspunkt« im Gangsta-Rap (Gruber 2017, S. 171). »Darstellungen einer lebensfeindlichen, städtischen Umgebung dienen als theatrale Kulisse, in der sich der Protagonist als den widrigen Umständen trotzender und daran gewachsener Großstadtkämpfer inszenieren kann« (ebd., S. 177). In der hier (re)präsentierten Sinnwelt ist das Zurschaustellen von teuren Marken und Statussymbolen bedeutsam; es dient »der Suggestion einer ›*Boss-Aura*‹« (ebd., S. 152; Herv. i. O.) und der Demonstration von Zugehörigkeit. Handlungen dieser Art sind identitätsrelevant. Mit ihrer Rapvideo-Eigenproduktion imaginieren und präsentieren sich Eduardo und Lorenzo als harte »Gangster« und »Drill-Rapper«, die mit den Praktiken, Kommunikations- und Ausdrucksweisen der »Straße« vertraut sind und dies souverän zum Ausdruck zu bringen vermögen. In diesem Zusammenhang adaptieren sie typische Themen und Stilmerkmale des Genres, bringen spezifische Reim- und Vertextungsstrategien zur Anwendung, (re)präsentieren Sinnwelt-interne Regeln, Wissens- und Sinnzusammenhänge und stellen so wichtige »Straße«- und Genre-bezogene Kompetenzen unter Beweis.

4. Zusammenfassung und Fazit

Dieser Beitrag widmet sich der Bedeutung von audiovisuellem Datenmaterial in der ethnographischen (Jugend-)Szeneforschung. Am Beispiel der Analyse einer audiovisuellen Eigenproduktion jugendlicher Gangsta-Rapper entfaltet der Beitrag Möglichkeiten der Datenaufbereitung und -analyse mit dem Ziel der Entdeckung zentraler kultureller Themen einer spezifischen Sinnwelt, und zwar aus der Perspektive junger männlicher Szene-Insider. In diesem Zusammenhang zeigt sich, *dass* und *wie* junge Szene-Insider wichtige Wissensbestände und Sinnzusammenhänge der »Straße« und des »Rap Games« Sound-bezogen, textlich und darstellerisch adaptieren und inszenieren, um sich selbst in einer für sie bedeutsamen Sinnwelt zu verorten. Dabei ergeben sich sowohl Einblicke in Szene-, Genre- und medienbezogene Kompetenzen jugendlicher Musikfans als auch Einblicke in die Beschaffenheit und subjektive Wahrnehmung ihrer unmittelbaren alltäglichen Lebens- und Erfahrungswelt. Mit dem Einbezug audiovisueller Artefakte in ethnographische Forschungsstrategien, so lässt sich vor dem Hintergrund dieses Beitrags zusammenfassen, eröffnen sich andersartige Einblicke in Artikulations-, Kommunikations-, Denk- und Handlungsweisen einzelner Teilnehmender des Feldes. Dies ist für die Beschreibung kultureller Wissens- und Sinnzusammenhänge von besonderem Wert, weil dadurch andersartige oder zusätzliche Perspektiven auf den Untersuchungsgegenstand generiert und das forscher:innenproduzierte Material sinnvoll angereichert und verdichtet werden kann.

Literatur

Bienk, Alice (2008): Filmsprache. Einführung in die interaktive Filmanalyse. 2. Auflage. Marburg: Schüren.

Blumer, Herbert (1981): Der methodologische Standort des Symbolischen Interaktionismus. In: Arbeitsgruppe Bielefelder Soziologen (Hrsg.): Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit. Band 54/55: 1, Symbolischer Interaktionismus und Ethnomethodologie. Opladen: Westdeutscher, S. 80–146.

Bock, Katharina (2017): Kommunikative Konstruktion von Szenekultur. Skateboarding als Sinnstiftung und Orientierung im Zeitalter der Digitalisierung. Weinheim und Basel: Beltz Juventa.

Bock, Katharina (2018): Zur Rolle online-medialer Inhalte für die Skateboardkultur. In: Schwier, Jürgen/Kilbert, Veith (Hrsg.): Skateboarding zwischen Subkultur und Olympia. Eine jugendliche Bewegungskultur im Spannungsfeld von Kommerzialisierung und Versportlichung. Bielefeld: transcript, S. 143–164.

Bock, Katharina (2019): Ethnografisches Protokollieren – Erkenntnisabsichten und sprachlich-stilistische Gestaltungsprinzipien [56 Absätze]. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 20(6). http://dx.doi.org/10.17169/fqs-20.1.2933.

Bock, Katharina (2024): Gangsta-Rap als soziales Spiel. In: Zeitschrift für Soziologie, 53(3). https://doi.org/10.1515/zfsoz-2024-2018.

Bock, Katharina (2025a) (i. E.): »Gangster« und »Gangsta-Rapper« von der »Straße« sein: Spielerisch gerahmtes Selbst-Erleben in der Auseinandersetzung mit populärer Musik und Kultur. In: Eisewicht, Paul/Eulenbach, Marcel/Heinze, Carsten/Schierbaum, Anja (Hrsg.): Jugend als zeit- und gesellschaftsdiagnostisches Phänomen. Weinheim, Basel: Beltz Juventa.

Bock, Katharina (2025b) (i. E.): Inszenierung urbaner Praktiken und szenespezifischer Wissensbestände. Skateboarding im Kurzfilm Belong. In: Rehbach, Simon (Hrsg.): Filmische Darstellung von Sport. Ansätze, Perspektiven, Analysen. Wiesbaden: Springer VS.

Breidenstein, Georg/Hirschauer, Stefan/Kalthoff, Herbert/Nieswand, Boris (2015): Ethnografie. Die Praxis der Feldforschung. 2., überarbeitete Auflage. Konstanz, München: UVK.

Bullerjahn, Claudia (2005): Analyse von Filmmusik und Musikvideos. In: Mikos, Lothar/Wegener, Claudia (Hrsg.): Qualitative Medienforschung. Ein Handbuch. Konstanz: UVK, S. 484–495.

Bundesverband Musikindustrie (2022): Musikindustrie in Zahlen 2022. www.musikindustrie.de/fileadmin/bvmi/upload/06_Publikationen/MiZ_Jahrbuch/2022/BVMI_Musikindustrie_in_Zahlen_2022_ePaper_230420_geschuetzt.pdf (09.07.2024).

Cohen, Sara (1993): Ethnography and popular music studies. In: Popular Music, 12, S. 123–138.

Dellwing, Michael/Prus, Robert (2012): Einführung in die interaktionistische Ethnografie. Soziologie im Außendienst. Wiesbaden: Springer VS.

Denzin, Norman K. (2010): Reading Film – Filme und Videos als sozialwissenschaftliches Erfahrungsmaterial. In: Flick, Uwe/Kardorff, Ernst von/Steinke, Ines (Hrsg.): Qualitative Forschung. Ein Handbuch. 8. Auflage. Reinbek bei Hamburg: Rowohlt, S. 416–428.

Goffman, Erving (1971): Interaktionsrituale. Über Verhalten in direkter Kommunikation. Frankfurt am Main: Suhrkamp.

Gruber, Johannes (2017): Performative Lyrik und lyrische Performance. Profilbildung im deutschen Rap. Bielefeld: transcript.

Hainz, Manuela (2022): Der Retro Trainingsanzug ist der neueste Hype – diese Modelle sind jetzt Trend. www.gq-magazin.de/mode/artikel/trainingsanzug-hype-modelle-trend (09.07.2024).

Hitzler, Ronald/Eisewicht, Paul (2016): Lebensweltanalytische Ethnographie – im Anschluss an Anne Honer. Weinheim, Basel: Beltz Juventa.

Hitzler, Ronald/Niederbacher, Arne (2010): Leben in Szenen. Formen juveniler Vergemeinschaftung heute. 3., vollständig überarbeitete Auflage. Opladen: Leske + Budrich.

Kautny, Oliver (2009): Ridin' the Beat. Annäherungen an das Phänomen Flow. In: Hörner, Fernand/Kautny, Oliver (Hrsg.): Die Stimme im HipHop. Untersuchungen eines intermedialen Phänomens. Bielefeld: transcript, S. 143–171.

Korte, Helmut (2005): Sequenzprotokoll. In: Mikos, Lothar/Wegener, Claudia (Hrsg.): Qualitative Medienforschung. Ein Handbuch. Konstanz: UVK, S. 387–394.

Lueger, Manfred/Froschauer, Ulrike (2022): Ethnographie und Artefaktanalyse. In: Poferl, Angelika/Schöer, Norbert (Hrsg.): Handbuch Soziologische Ethnographie. Wiesbaden: Springer VS, S. 519–528.
Maeder, Christoph/Brosziewski, Achim (1997): Ethnographische Semantik: Ein Weg zum Verstehen von Zugehörigkeit. In: Hitzler, Ronald/Honer, Anne (Hrsg.): Sozialwissenschaftliche Hermeneutik. Eine Einführung. Opladen: Leske + Budrich, S. 335–362.
Maeder, Christoph (2002): Alltagsroutine, Sozialstruktur und soziologische Theorie: Gefängnisforschung mit ethnografischer Semantik [26 Absätze]. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 3(1), Art. 15. http://nbn-resolving.de/urn:nbn:de:0114-fqs0201159
Maeder, Christoph (2007): Ethnographische Semantik. Die Ordnung der Mitgliedschaftssymbole am Beispiel des Bergsports. In: Buber, Renate/Holzmüller, Hartmut H. (Hrsg.): Qualitative Marktforschung. Konzepte–Methoden–Analysen. Wiesbaden: GWW Fachverlage, S. 681–696.
Monteith, Keir (2022): Ban rap and drill lyrics in the courtroom. In: Popular Music, 41, S. 546–557.
Niesyto, Horst (2018): Symbolverstehen von Video-Eigenproduktionen. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 781–795.
Poferl, Angelika/Schröer, Norbert (Hrsg.) (2022): Handbuch Soziologische Ethnographie. Wiesbaden: Springer VS.
Schmidt, Friederike/Bock, Katharina (2024): Virtualität(en) und Gangsta-Rap. Entwürfe möglicher Realitäten in der Jugend. In: Heinze, Carsten/Masmanidis, Konstantinos/Zirfas, Jörg (Hrsg.): Anthropologien der Virtualität. Pädagogiken der Möglichkeit, des Entwurfs und der Imagination. Weinheim, Basel: Beltz Juventa, S. 104–115.
Schröer, Sebastian (2012): »Ich bin doch kein Gangster!« Implikationen und Paradoxien szeneorientierter (Selbst-)Inszenierung. In: Dietrich, Marc/Seeliger, Martin (Hrsg.): Deutscher Gangsta-Rap. Sozial- und kultwissenschaftliche Beiträge zu einem Pop-Phänomen. Bielefeld: transcript, S. 65–83.
Schützeichel, Rainer (2007): Laien, Experten, Professionen. In: Schützeichel, Rainer (Hrsg.): Handbuch Wissenssoziologie und Wissensforschung. Konstanz: UVK, S. 546–578.
Seeliger, Martin (2021): Soziologie des Gangsta-Rap. Weinheim, Basel: Beltz Juventa.
Seeliger, Martin/Marc Dietrich (Hrsg.) (2017): Deutscher Gangsta-Rap II. Popkultur als Kampf um Anerkennung und Integration. Bielefeld: transcript.
Terms & Conditions (2020): A UK Drill Story. Großbritannien 2020. 80 Minuten. Regie: Brian Hill. www.youtube.com/watch?v=kno5T4y5SBY (09.07.2024).
Tonspion.de (2021): Trap Music im Hip-Hop: Die konfuse Findungsgeschichte eines Genres. www.tonspion.de/news/trap-music-im-hip-hop# (09.07.2024).
Urbandictionary.com (o. J. a): Lesh. www.urbandictionary.com/define.php?term=Lesh (09.07.2024).
Urbandictionary.com (o. J. b): Traphouse. www.urbandictionary.com/define.php?term=traphouse (09.07.2024).
Witzke, Margrit (2005): Jugendforschung mit Video-Eigenproduktionen. In: Mikos, Lothar/Wegener, Claudia (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 323–33.

Analyse von Videodaten

Unterschiede zwischen Feldvideos und Plattformvideos

Bernt Schnettler

Zusammenfassung: Schnettler diskutiert die methodischen Prinzipien der Analyse von zwei grundverschiedenen Typen von Videodaten, die als Feldvideos in Videografien von den Forschenden selbst erzeugt werden oder als Plattformvideos aus einem proliferativen Fundus online verbreiteter Videoclips stammen. Auf Grundlage von Feinanalysen zweier Datenauszüge aus eigenen Forschungen werden die Prinzipien der Sequenzanalyse und Interpretation von Feldvideos versus Plattformvideos exemplarisch vorgeführt. Der Autor diskutiert die Herausforderungen der jeweiligen Datentypen sowie die ihnen angemessenen methodischen Verfahrensweisen und plädiert für eine Kombination von Methoden ethnomethodologisch fundierter Korpusanalysen mit hermeneutisch begründeten Einzelfallanalysen.

Schlüsselwörter: Sequenzanalyse, Ethnomethodologie, Hermeneutik, Social Media

1. Einleitung

Videodaten sind in den vergangenen Jahren verstärkt zum Mittel und Gegenstand sozialwissenschaftlicher Analysen geworden, wobei eine wachsende Vielfalt neuer Methoden zu verzeichnen ist. Tatsächlich stellen Videodaten die sozialwissenschaftliche Forschung vor erhebliche Herausforderungen. Gründe dafür sind die enorme Menge von Videoclips und der gewaltige Formenreichtum diversester Videoformate, die kontinuierlich in die Welt strömen. Videos tauchen in immer neuen Varianten und Gattungen auf. Um diese Vielfalt analytisch in den Griff zu bekommen, ist es notwendig, sozialwissenschaftliche Methoden der Videoanalyse weiter voranzutreiben und bestehende Ansätze konsequent fortzuentwickeln.

In diesem Aufsatz werden zwei in der Videoanalyse häufig verwendete Typen audiovisueller Daten diskutiert: Zum einen Videos, die im Zuge eigener Feldforschung angefertigt werden. Zum anderen Videodaten, die von digitalen Plattformen abgerufen, über soziale Medien verteilt werden oder im Internet zirkulieren. Adressiert wird im Folgenden eine Frage, die sich auf diese verschiedenen audiovisuellen Datensorten in der qualitativen Sozialforschung bezieht. Neben Seitenblicken auf die Erhebungskontexte adressiere ich damit vor allem Probleme der

Videodaten*analyse*. Gezeigt werden soll, welcher Ertrag sich aus einer Kombination hermeneutischer, ethnomethodologischer und gattungsanalytischer Prinzipien der Sequenzanalyse von Videodaten methodisch gewinnen lässt.

Entsprechend gliedert sich der Text in folgende Abschnitte: Im zweiten werden die grundlegenden Prinzipien interpretativer Videoanalyse knapp rekapituliert. Der dritte Abschnitt widmet sich der exemplarischen Detailanalyse von zwei Hauptvarianten von Videodaten, die als *Feldvideos* und als *Plattformvideos* bezeichnet werden. Anhand detaillierter Sequenzanalysen von zwei Datenfragmenten eigener Forschungen werden die Herausforderungen veranschaulicht, die beide Typen von Videomaterial in der Analyse stellen. Schließlich wird im vierten Abschnitt eine Bewertung der beiden methodischen Herangehensweisen in Bezug auf ihre Vor- und Nachteile bei der sozialwissenschaftlichen Videoanalyse vorgenommen.

2. Grundlagen interpretativer Videoanalyse

Die Videoanalyse ist, verglichen mit Verfahren wie teilnehmende Beobachtung, Umfragen oder Fokusgruppen, eine relativ junge Methode der qualitativen Sozialforschung. Inzwischen haben sich Videoanalysen aber in zahlreichen Feldern etabliert, z. B. in der Gesprächsanalyse, der Interaktionsforschung, der Schulforschung oder der Marktforschung.[1]

Die methodischen Wurzeln der Videoanalyse liegen in Forschungen aus der Anthropologie, der Ethologie und der Kommunikationsanalyse, die sich durch ihr gemeinsames Interesse an einem sehr detaillierten Studium von Verhaltens- und Kommunikationsabläufen auszeichnen. Ebenso stark stützt sich die Videoanalyse auf die linguistische Forschung über Formen und Strukturen gesprochener Sprache. Bedeutsam war hier vor allem die Ethnographie der Kommunikation von John J. Gumperz, der in den 1970er Jahren in seinen Studien über interkulturelle Interaktionen Videoaufnahmen als Erhebungstechnik einsetzte (Gumperz/Hymes 1972). Die qualitative Videoanalyse rekurriert ebenso auf die *Kontextanalyse* von Kendon (Kendon 1990) und Goffmans urbane Ethnografien über Interaktionen im öffentlichen Raum (Goffman 1961) und seine soziolinguistischen Arbeiten zu *Forms of Talk* (Goffman 1981) sowie zur Interaktionsordnung (Goffman 1983).

1 In außerakademischen Anwendungsbereichen gibt es außerdem Praxisformen der Videoanalyse, die z. B. im Leistungssporttraining, in Therapiesettings, in der Rhetorikausbildung oder bei der polizeilichen Analyse von Videoüberwachungsmaterial zum Einsatz kommen, die Tuma (2017) in seiner instruktiven Studie untersucht. Er prägt dafür den Begriff »vernakulare« Videoanalyse und leitet methodische Prinzipien für die sozialwissenschaftliche Videoanalyse ab, womit er einen wichtigen Betrag zur Methodenentwicklung leistet.

Forschende aus der Ethnomethodologie, allen voran Charles Goodwin und Christian Heath (Heath 1986; Heath/Hindmarsh/Luff 2010) haben die qualitative Videoanalyse entscheidend vorangetrieben. Diese interaktionale Videoanalyse gründet in theoretischen Postulaten der Ethnomethodologie (Garfinkel 1967) und der Konversationsanalyse (Sacks 1992) und ist besonders nützlich zur Untersuchung »natürlicher« Interaktionen und Alltagspraktiken.[2]

Für eine Analyse von Videodaten aus *medialen Kontexten* ist es jedoch erforderlich, auf einen anderen methodologischen Ansatz zurückzugreifen, der als Videohermeneutik bezeichnet wird (Raab/Tänzler 2012; Raab/Maier 2022). Sie operiert im Rahmen der von Soeffner (2004) entwickelten Hermeneutischen Wissenssoziologie und praktiziert eine Form der Sequenzanalyse, die sich von der ethnomethodologischen deutlich unterscheidet. Vorzug der Videohermeneutik ist, dass sie auch Dimensionen der Filmästhetik systematisch berücksichtigt (vgl. unten).

Videoaufnahmen werden schon seit vielen Jahren in der Feldforschung als Ergänzung zur ethnographischen Beobachtung eingesetzt. Häufig beschränkte sich der Einsatz von Video durch die Forschenden jedoch darauf, Untersuchungsergebnisse zu präsentieren oder Schlüsselaspekte des Untersuchungsthemas audiovisuell zu dokumentieren. Dabei wird übersehen, dass Videoaufnahmen ein weitaus leistungsfähigeres Instrument für die interpretative Sozialforschung sein können. Was sind die wichtigsten allgemeinen Vorzüge von Videodaten für die interpretative Sozialforschung?

(1) Videos bieten die Möglichkeit, soziale Prozesse *als Prozesse* zu untersuchen, denn sie liefern Daten, die die Entfaltung sozialer Situationen in ihrem Ablauf dokumentieren und analysierbar machen. Angelehnt an Goffman werden soziale Situationen als diejenigen Ereignisse definiert, in denen mindestens zwei Personen von Angesicht zu Angesicht miteinander interagieren. Bergmann (1985) hat den Begriff der *registrierenden Konservierung* geprägt, was bedeutet, dass Aufnahmen eine exakte Untersuchung der zeitlichen Entwicklung von Interaktionsverläufen ermöglichen. Dementsprechend sind Videoanalysen umso leistungsfähiger, je fokussierter die aufgezeichneten Interaktionen sind. Die qualitative Videoanalyse ist deshalb besonders geeignet dafür, situative Interaktionen zu analysieren, die zeitlich nur *kurz* dauern und räumlich klar *fokussiert* sind. Diese auch als interaktional bezeichnete Variante der Videoanalyse ist eine Form der *Mikro*analyse. Genauer wäre sogar von *Nanoanalyse zu* sprechen, weil Videoaufzeichnungen es ermöglichen, in kleinste Detailtiefen des Verlaufs sozialer Interaktionen vorzudringen und Aspekte zu entdecken, die mit bloßem Auge

2 Vgl. vom Lehn (2017) für eine hervorragende Synthese der ethnomethodologisch fundierten Videoanalyse.

unbemerkt bleiben.[3] Aus diesem Grund ist die Videoanalyse auch als *soziologische Mikroskopie* bezeichnet worden (zuerst von Knoblauch 2001; vgl. dazu auch Knoblauch 2002).

(2) Außerdem dokumentieren Videoaufzeichnungen im Material gleichzeitig sowohl *Sequenzialität* als auch *Simultaneität*: Videos weisen eine klare innere Sequenzstruktur auf, die die kontinuierliche zeitliche Abfolge der in der Aufnahme festgehaltenen Interaktionen registriert. Kamerabewegungen und -schwenks oder Schnitt und Bearbeitung verleihen dem Video weitere Dynamik und erzeugen zusätzliche zeitliche und räumliche Modulationen. Damit wird der primären Bedeutungsstruktur der Handlungs- und Interaktionsprozesse eine weitere *mediale* Sinnschicht hinzugefügt. Die dynamische Entwicklung beider Sequenzialitäten können in Feinanalysen Schritt für Schritt untersucht werden. Simultaneität bezeichnet demgegenüber die Eigenschaft von Videos, die in Gleichzeitigkeit dauerhaft anwesenden Situationsaspekte abzubilden. So hält das Video sichtbare Elemente der räumlichen Umgebung, z. B. Möbel, Utensilien, Geräte etc. fest, sodass in der Aufzeichnung simultan unveränderliche, über den gesamten Interaktionsverlauf sichtbare Situationsaspekte visuell festgehalten werden. Ebenso gehören dazu potenziell bedeutungstragende Aspekte der in der Situation interagierenden Personen, z. B. deren körperliche Erscheinung, Statur, Größe oder Bekleidung. Alle diese statischen Elemente können und sollten ebenfalls interpretiert werden. Methodisch gesehen muss die interpretative Videoanalyse die sich bietenden Interpretations*potenziale* berücksichtigen. Wenn wir uns im Alltag ein Bild von einer Person machen, assoziieren wir üblicherweise diese Auslegungen von ihr mit denjenigen, die uns gesellschaftlich verfügbare Interpretationsrahmen liefern. Besonders anschauliche Beispiele dafür sind unsere Deutungen von Geschlechtsmerkmalen, Alter oder Klassenzurechnungen. Diese Alltagsdeutungen werden hochroutiniert vorgenommen und nur in Ausnahmefällen problematisiert, etwa wenn Irritation oder Verunsicherungen auftreten und die üblichen Deutungen gestört werden. Die interpretative Sozialforschung schließt an diese Alltagsdeutungen an. Anders als bei Alltagsdeutungen ist es in der wissenschaftlichen Auslegung jedoch unerlässlich zu explizieren, *wie* gesellschaftliche Deutungs- und Zuschreibungsprozesse ablaufen. Die Hermeneutische Wissenssoziologie in der an Soeffner (2004) anknüpfenden Analysetradition hat dazu fundamentale Beiträge geleistet. Hervorzuheben ist, dass eine exhaustive Auslegung aller erdenklicher Lesarten zu leisten ist, wobei zugleich deren jeweilige historische und soziale Prägung mitreflektiert und gegebenenfalls hinterfragt werden muss. Um das zu gewährleisten ist es ratsam, Videoanalysen mit einer Interpretationsgruppe in gemeinsamen Datensitzungen durchzuführen. Auf

3 Den Hinweis auf diese »nanoskopische« Dimension verdanke ich Thomas Eberle. Ein hierfür höchst anschauliches Beispiel liefert die Feinanalyse des Zuschlags mit dem Auktionshammer in der Studie von Heath (2015).

diese folgen jeweils Phasen stärker von Solitär- oder Teamarbeit geprägter Arbeit, die hauptsächlich der Dokumentenerstellung dienen wie das Schreiben von Analyseprotokollen, Forschungsberichten, Textentwürfen für Artikel usw.[4]

3. Typologien von Videodaten nach Herkunft: Feld versus Plattform

Kehren wir zu den Herausforderungen zurück, die aus der großen Vielfalt von Videoformaten resultieren. Bei »natürlichen« Videos im Sinne von in der Forschung selbst angefertigten Aufnahmen, die dem Grundsatz des Datennaturalismus folgen, werden soziale Situationen exakt so aufgezeichnet, wie sie sich ereignet haben. Diese Art von Videos nennen wir »*Feldvideos*«. Im Gegensatz dazu bezeichnen wir als »*Plattformvideos*« alle Arten von Clips, die über das Internet oder soziale Medien verbreitet werden. Beides sind zweifellos Idealtypen im Weber'schen Sinne: konzeptionelle Varianten, die aus einer dynamisch wachsenden Mannigfaltigkeit empirischer Formen von Videoclips gebildet werden. Denn es existieren zahlreiche Übergangsformen, Abwandlungen und Modulationen. Angesichts der Fülle von Videodatentypen stellt sich die Frage, welches die geeignete Methodik für die jeweilige Variante ist.

Das Qualitätskriterium für Feldvideos ist die Wiedergabetreue, mit der die Aufnahme die spezifische soziale Situation, die Gegenstand unseres analytischen Interesses ist, dokumentarisch fixiert. Wiedergabetreue bedeutet, dass die im Fokus stehenden Interaktionen möglichst gut im Video festgehalten werden. Das stößt auf Probleme. Zum Beispiel sind Videoanalysen von Brennholzschneiden mit der Kettensäge schwierig, wenn sich das Forschungsinteresse auf die konkreten Praktiken richtet, der Umgang mit den Gerätschaften aber nicht gut gefilmt werden kann, weil das nur von hinten möglich ist.[5] Ähnliche Schwierigkeiten stehen der Videografie häufig bei Aufnahmen in Klassenzimmern oder Kindergärten im Wege, wo in hochdynamischen Umgebungen komplexe multifokale Interaktionen mit zahlreichen Teilnehmenden stattfinden, die äußerst schwer zu videografieren sind (vgl. Carvalho; Hamer/Helbig/Urban; Jehle; Herrle/Proske/Puzicha/Zimmer; Mempel/Winterscheidt i. d. B.).

Bei Plattformvideos (vgl. Bock i. d. B.) handelt es sich demgegenüber um audiovisuelle Produkte, die zu höchst unterschiedlichen Zwecken aufgenommen, geteilt oder online hochgeladen werden. Sie sind typischerweise für ein Publikum bestimmt, das häufig aus einzelnen bzw. wenigen Adressatinnen und Adressaten

4 Für eine ausführliche Darstellung der qualitativen Videoanalyse siehe auch Tuma/Knoblauch/Schnettler (2013) sowie Schnettler/Knoblauch/Baer (2012).

5 Für zwei besonders gelungene Videointeraktionsanalysen vgl. die Studie von Sing (2019) zum Nachwuchstraining im Sport und die Studie von Wilke (2022) zum Group-Talk in der Wissenschaft.

besteht, mitunter auch von einem größeren Kreis von Rezipierenden betrachtet wird. Deshalb ist diesem Material, anders als Feldvideos, ein *recipient design* eingeschrieben. Während sich die Qualität von Feldvideos an den Sachanforderungen der Wissenschaft orientiert, prägt Plattformvideos der Kommunikationszuschnitt auf mehr oder weniger bekannte Adressaten und Adressatinnen in pragmatischen alltagsweltlichen Kontexten. Deren Fülle ist ebenso groß wie die Streubreite ihrer Formen und Funktionen. In derartigen Videoclips fließen Informationsvermittlung mit Aufklärungswillen und Wissensverbreitung zusammen. Sie oszillieren zwischen Kommunikationsbedürfnis, Selbstinszenierungswunsch, Überzeugungswillen und Manipulationsabsicht. Folglich lassen sich im Material informative, alltagsästhetische, persuasive und unterhaltungsbezogene Funktionen und Formen oft nicht leicht voneinander abgrenzen. Im Zuge des Florierens digitaler Kommunikationstechnologien nehmen Plattformvideos verständlicherweise starken Einfluss auf den Wandel der kommunikativen Interaktionsweisen, sowohl die *one-to-one-* als auch die *one-to-many*.

Plattformvideos umfassen eine Vielzahl von Untertypen und Varianten, die von kurzen, vergleichsweise frei erstellten Videobotschaften, bis hin zu standardisierten Formaten reichen, für die sich bereits eigene Termini gebildet haben und deren Gattungsmerkmale leicht identifiziert werden können. Zu letzteren gehören z. B. Tutorials und How-To-Videos, DIY-Videos, ASMR-Videos, Life Hacks, Pranks, Let's Plays oder die von Reinbold (2023) untersuchten Coronawarnvideos. Die Dynamik, mit der immer neue Varianten auftauchen, ist erstaunlich hoch, was für die Forschenden große Anstrengung erfordert, mit dem Aufkommen neuer Videoformate Schritt zu halten. Und während die Methoden der Videoanalyse für Feldvideos bereits recht fortgeschritten sind, befindet sich die qualitative Analyse von Plattformvideos methodisch erst im Anfangsstadium.

Die Fragen lauten: Können Plattformvideos auf ähnliche Weise qualitativ untersucht werden wie Feldvideos? Welche methodischen Ansätze lassen sich aus der Analyse von Feldvideos auf Plattformvideos übertragen und wo wäre eine Erweiterung oder Modifizierung des analytischen Ansatzes erforderlich? In den nächsten Abschnitten werden wir versuchen, diese Fragen zu beantworten, indem wir zwei Kontrastbeispiele analysieren, die die jeweiligen Merkmale der Analyse von Feldvideos und Plattformvideos veranschaulichen.

Beispiel 1: Eisengießen (Feldvideo)

Das erste Fragment stammt aus einem Feldvideo, das 2015 in Ghana in einer handwerklichen Eisengießerei aufgenommen wurde, in der Mahlscheiben für Maismühlen hergestellt werden. Das Datum entstand im Rahmen einer Studie, die einerseits die Tradition der an der Universität Bayreuth betriebenen Anthropologie der Arbeit aufgreift, deren zentrale Erkenntnis Spittler folgendermaßen

formuliert hat: »Arbeit ist universell, aber sie existiert in vielen Formen. Wenn wir die Einheit und die Vielfalt verstehen wollen, genügt es nicht, Arbeit in unserer eigenen Gesellschaft zu untersuchen, sondern wir müssen unseren Blick darüber hinaus lenken« (Spittler 2016, S. 2). Andererseits knüpft die Studie an die interdisziplinäre Forschungsrichtung an, die als *Workplace-Studies* (WPS) bekannt geworden ist. Dieser Forschungsansatz, der sich in detaillierten Untersuchungen mit Arbeit, Technologie und Interaktion in komplexen Organisationen befasst, entstand in den 1980er Jahren. Die WPS sind international weit verbreitet und aus der Konvergenz von Forschungsinteressen in den Bereichen Mensch-Computer-Interaktion (HCI), künstliche Intelligenz (AI), computergestützte kollaborative Arbeit (CSCW), ethnomethodologische Gesprächsanalyse, Ethnographie sowie entsprechende soziologische Ansätze hervorgegangen (Luff/Hindmarsh/Heath 2000; Heath/Luff 2008). Die WPS konnten beträchtliche Forschungsbefunde akkumulieren, die zu einem vertieften Verständnis derjenigen interaktiven und kommunikativen Prozesse geführt haben, die für die Ausführung von Arbeitsaktivitäten entscheidend sind. Die meisten dieser Studien rücken komplexe, hochtechnisierte Arbeitsumgebungen in industriellen oder sogar postindustriellen Gesellschaften in den Fokus. Handwerkliche bzw. »low-tech« Arbeitsumgebungen wurden bisher kaum erforscht, obwohl gerade diese eine ideale Gelegenheit dafür bieten, die für den Arbeitsvollzug wesentlichen interaktiven Prozesse zu durchleuchten. Eine der wenigen Ausnahmen ist Dant (2005, 2007), der in seinen ethnographischen Studien den Umgang der in einer Autowerkstatt Tätigen mit Werkzeugen analysiert und die pragmatische Bedeutung der Verwendung von Arbeitsgeräten eindrücklich herausgearbeitet hat.

Die folgenden Auszüge stammen aus eigenen Videoaufzeichnungen. Angefertigt wurden sie während einer Forschungsreise, an der der Autor als Teil eines dreiköpfigen Teams teilnahm, zu dem ein sehr erfahrener Ethnologe und Afrikaexperte sowie eine auf Verfahrenstechnik spezialisierte Ingenieurin gehörten. Das Projekt war als Feldforschung konzipiert, bei der neben anderen qualitativen Methoden auch die Videointeraktionsanalyse zum Einsatz kam. Kurzgefasst bestand das Forschungsziel darin, die Prozesse wechselseitiger Wissensflüsse zwischen den Handwerkern im Suame-Magazin und den Metallurgen und Metallbauingenieuren der Nkrumah University of Science and Technology (KNUST) in Kumasi (Ghana) zu untersuchen. In Ghana wurden seit den 1980er Jahren mehrfach Versuche unternommen, vorhandene handwerklich-vorindustrielle Eisenverhüttungstechniken zu optimieren, um Importabhängigkeiten zu verringern. Dem Know-how der Handwerker räumte man dabei große Bedeutung ein.

Das Suame Magazine in Kumasi wird von manchen als Afrikas größter industrieller Slum bezeichnet, von anderen als Zentrum handwerklichen Erfindungsreichtums in Afrika. In Ghana genießt es den Ruf, dass dort alle erdenklichen technischen Probleme gelöst werden können – wenn nicht förmlich, dann informell. Als größtes metallverarbeitendes Cluster in Westafrika erstreckt sich das

Suame Magazine über mehr als 50 km² und beherbergt rund 1.000 Unternehmen im informellen handwerklichen Fertigungssektor (Jaarsma et al. 2011). Die meisten Betriebe in Suame Magazine arbeiten rein handwerklich und sind gleichzeitig in der Ausübung ihrer Funktion hoch spezialisiert (Kensok 1987; Powell 1995). In jedem Fall beherrscht das »Wissen der Hand« hier eine ganz spezifische Materialität, wobei Kenntnisse und Fähigkeiten eng mit der handwerklichen Praxis verbunden geblieben sind. Dabei geht es nicht um formalisierte, sondern um hochgradig personalisierte Wissensbestände, die nur in individualisierten Lehrbeziehungen von Person zu Person weitergegeben werden. Es gibt weder Diplome noch Lehrpläne oder formale Ausbildungswege, mithilfe derer man sich die entsprechenden Kenntnisse aneignen könnte. Aus wissenssoziologischer Perspektive interessierte sich unsere Forschung vor allem für jene präprädikativen Wissensformen, die als Fertigkeiten oder verkörpertes Wissen bezeichnet werden und die für jeden Umgang mit Objekten und Arbeitsinstrumenten grundlegend sind (Beck 1997; Harper 1992; Dant 2005; Keller/Keller 1996).

Dank lokaler Kontakte konnte unser Forschungsteam eine der typischen Eisengießereien besuchen, die im Suame Magazine produzieren. Feldzugang erhielten wir durch einen kompetenten Ingenieur mit langjähriger Ortserfahrung, der als Direktor der »Intermediate Technology Transfer Unit« (ITTU) an der Schnittstelle zwischen Handwerk und Wissenschaft tätig war. Mit seiner Hilfe gelang es, eine der Werkstätten zu betreten. Beim zweiten Besuch wurden uns Videoaufnahmen erlaubt. Der folgende Auszug aus meinen Feldnotizen hält meine damaligen Eindrücke vor Ort fest:

> Er führte uns schließlich zu der Schmelze, wo Gusseisen hergestellt wurde und wir beeindruckende Aufnahmen von den archaisch anmutenden Arbeitsprozessen machen konnten. Die Hitze war erheblich. Die Arbeitenden ohne Schutzkleidung und mit selbsthergestellten Geräten ausgestattet. Ich konnte mehrere Durchgänge aufnehmen, die mit dem Anstich begannen. Zwei Männer stellen eine Art eiserne Bottichtrage, offenkundig Marke Eigenbau (…), vor den Hochofen, ein dritter sticht mit einer Stange den Lehmklumpen durch, der den Ofen verschließt und in einem hellglühenden Strahl schießt das geschmolzene Gusseisen in die Trage. Ist dieser gefüllt, stopft der Stecher mit der Hand, die oft nur mit einem dicken Handschuh ausgestattet ist, mit einem Lehmpfropfen das Loch wieder zu und bringt so den Strahl zum Stoppen.

In den Feldnotizen spiegeln sich bereits einige der wichtigsten Schritte dieser mühsamen metallurgischen Arbeit wider.[6] Um jedoch eine genaue mikroanalytische Studie durchzuführen, werden die Videoaufzeichnungen feinanalytisch untersucht. Eine kleinteilige Sequenzanalyse der Arbeitsvollzüge gibt Aufschluss

6 Aus Platzgründen musste der Beitrag hier stark gekürzt werden.

darüber, wie jeder Schritt dieses komplizierten Produktionsprozesses gestaltet ist. Dabei erkennen wir, wie die Koordination der Arbeit am Hochofen und dem Befüllen der Gussformen durch zwei Werkzeuge vermittelt wird, nämlich: Eisenstange und Tragebottich (siehe Abb. 1).

Abbildung 1: Sequenzanalyse der Phasen der untersuchten Arbeitstätigkeit (eigene Zusammenstellung von Standbildern aus dem Videomaterial)

Die folgende Interaktionsanalyse dieses Materials von Pach (2018) kombiniert die Ethnologie der Arbeit mit der Arbeits-, der Wissens- und der Kommunikationssoziologie und integriert ihr linguistisches Wissen in die Videodatenanalyse. Pach beforscht die »Praxis des Eisengießens im […] Suame Magazine aus der Perspektive der Workplace Studies« (Pach 2018, S. 6), wobei sie die interaktionelle Nutzung von Werkzeugen zur Koordination der Arbeitsprozesse zur Leitachse der Analyse macht. Adressierte das Projekt zum »Wissen der Hand« den gesamten Prozess des Wissensaustauschs zwischen Handwerkern und Metallurgen, konzentriert sich Pach auf die mikroskopische Detailanalyse der Arbeitsprozesse selbst. Die Konzepte »embodiment« (Streek/LeBaron 2013) und »professional Vision« (Goodwin 1994) aufnehmend, die entwickelt wurden, um in den Körper eingeschriebene, schwer verbalisierbare Wissensformen zu bezeichnen, konzediert Pach, dass das Konzept verkörperter Kompetenzen einen wichtigen konzeptionellen Ausgangspunkt liefert, zieht es aber vor, mit den in der Multimodalitätsforschung entwickelten Konzepten Kooperation und Koordination zu arbeiten. Zur Fundierung ihrer »Performance-Rekonstruktions-Perspektive« stützt Pach sich auf das von Goffman (1983) entwickelte Konzept der Interaktionsordnung sowie auf Folgekonzepte der Interaktionsforschung (Deppermann/Schmidt 2007), die nicht nur sprachliche, sondern auch körperliche und räumliche Ressourcen in der Analyse berücksichtigen.

In der Perspektive einer exakten Interaktionsanalyse wird »Koordination« als Proto-Phänomen der Interaktion definiert, das keinen »vollwertigen Handlungscharakter« besitzt, sondern die »organisatorische Voraussetzung für Kooperation« (Pach 2018, S. 30) darstellt.

Der Einwand, die Vernachlässigung der Verbalkommunikation aufgrund unserer mangelnden Kenntnisse der von den Arbeitenden gesprochenen Sprache Twi stelle eine ernsthafte Einschränkung dar, kann relativiert werden. Denn in diesem Kontext spielt die verbale Modalität eine untergeordnete Rolle. Zwar reden die Arbeiter miteinander. Für die Koordination ihrer einzelnen Arbeitsvollzüge sind physische Kommunikationsmodalitäten jedoch weitaus wichtiger. Denn die Arbeit findet in einer Umgebung statt, die durch enorme Hitze, große Betriebsamkeit und hohen Lärmpegel gekennzeichnet ist. Als primäre Koordinationsressourcen für die erfolgreiche Durchführung kommen deshalb paraverbale Ausdrucksformen wie Blickorientierung, Gestik, Mimik oder Körperhaltungen und -bewegungen zum Einsatz, und nicht Vokalisierungen.

Pachs Analyse deckt mikroskopische Details der Koordinierung der Arbeitsvorgänge sowie deren Kontextualisierung in großer Nuanciertheit auf. Weil es sich um harte körperliche Arbeit unter extremen Bedingungen von Hitze, Lärm und Staubbelastung mit hohem Risikopotenzial handelt, ist eine gut eingespielte und zuverlässig funktionierende Koordination der Tätigkeiten unerlässlich. Darüber hinaus beschreibt sie die Werkzeuge, die zur Herstellung der Mahlscheiben verwendet werden: Hochofen, Trog und Sandgussformen spielen eine zentrale Rolle. Genaue Beobachtung und sorgfältige Sequenzanalyse lassen sie die wesentlichen Merkmale der Art und Weise entdecken, mit der die Teilnehmer den Arbeitsprozess ausführen. Pach (2018, S. 45) hebt die Besonderheit der verwendeten Werkzeuge hervor, die das außergewöhnliche Improvisationsgeschick der in dieser Werkstatt Tätigen eindrücklich demonstriert. Mit ihren detaillierten Sequenzanalysen der praktischen Struktur dieses Arbeitsprozesses stellt sie die analytische Kraft der mikroskopischen Videoanalyse unter Beweis und rekonstruiert mit ihrer Studie akribisch jede der Phasen des Eisengießens – vom Anbohren des Ofens über das Befüllen der Formen und dem anschließenden Rücktransport der Gerätschaften an den Ausgangspunkt der Arbeitskette.

Pachs videografische Analyse zeichnet sich durch größere Genauigkeit und Detailliertheit aus als eine nur auf Feldnotizen gestützte Untersuchung. Um ihre Ergebnisse zu veranschaulichen, fertig sie Skizzen an, die die einzelnen Arbeitsschritte deutlicher sichtbar machen als Videostandbilder. Auf diese Weise bricht Pach nicht nur den Arbeitsprozess bis ins kleinste Detail herunter. Es gelingt ihr außerdem, in der Videoanalyse die zentrale Bedeutung der *Intermaterialität* aufzuzeigen (siehe Abb. 2). Intermaterialität ist eine Form haptischer Kommunikation, mit der die Arbeiter die Mikrokoordination der einzelnen Schritte ihrer gefährlichen Arbeit vollziehen. Diese

Interaktion entpuppt sich als Form wechselseitigen intermateriellen Handelns, das physisch und haptisch vollzogen wird.

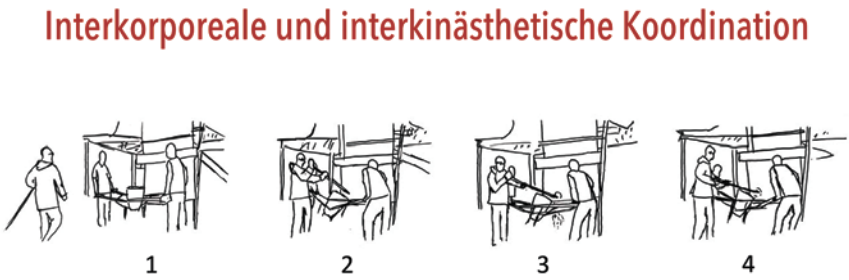


Abbildung 2: Wechselseitige Koordination mithilfe von Körper und Arbeitsinstrumenten (Pach 2018)

Diese Analysen konvergieren im zentralen analytischen Konzept des »Wissens der Hand«. Damit wird zum Ausdruck gebracht, dass der Schwerpunkt auf der Intensität händischer Koordination liegt und die unverzichtbare Bedeutung des Einsatzes von Händen (und Armen) der beteiligten Arbeiter für die erfolgreiche Ausführung ihrer Arbeit unterstrichen. Pach zeigt die interaktive Relevanz dieser Details eindrucksvoll, indem sie Routineabläufe und problematische Abläufe miteinander vergleicht. Damit die beiden Träger ihre Arbeit erfolgreich ausführen können, ist es unerlässlich, sich ständig gegenseitig zu beobachten. Ihre Koordination mit Hand- und Armbewegungen wird von wechselseitiger Körperwahrnehmung begleitet, z. B. durch über die Griffstangen des Trogs vermitteltes Spüren der Gewichtsbelastung des anderen. Daher erweist sich die exakte zeitliche Koordination der Wahrnehmungsfokussierung als eine der kritischen Aufgaben, die kontinuierlich und präzise durchgeführt werden muss, um die notwendigen körperlichen Aktivitäten in der erforderlichen Synchronität ausführen zu können. Pach verwendet zur Vertiefung der Analyse die körperphänomenologischen Konzepte »Inter-Korporealität« und »Inter-Kinästhesie« (Meyer/von Wedelstaedt 2017), womit sie zeigt, dass die sichtbare Konzentration der Träger als Ausdruck einer erfolgreichen interaktionellen »Verschmelzung« zwischen den beiden Arbeitern durch das Gerät verstanden werden muss.

Videosequenzanalyse ermöglicht also, bis in die kleinsten Details der untersuchten Interaktionen vorzudringen. Videodaten dieser Art eröffnen privilegierten Zugang zum *unmittelbaren* Interaktionskontext und den sie umlagernden Kontexten: Auf der ersten Ebene können wir mit diesem Video die fragliche Arbeitstätigkeit selbst im Detail analysieren. Weil es sich um von den Forschenden selbst angefertigte Aufnahmen im Rahmen eigener Feldforschung handelt, können diese Daten, auf einer zweiten Ebene, verwendet werden, um

allgemeinere Fragen zu adressieren. Wir können Fragen beantworten zu den weiteren lokalen Kontexten und sozialen Welten, jenseits dessen, was im Videofragment selbst dokumentiert ist. So wissen wir aus unserer damaligen Ethnographie, dass die untersuchte Aktivität Teil eines Produktionsprozesses des Hüttenbetriebs ist, der wiederum Teil des Suame Magazins ist, das zum Metallhandelsnetzwerk in Ghana gehört, das weit bis in die umliegenden Nachbarländer in Westafrika reichende Produktions- und Lieferbeziehungen unterhält usw. Diese Außenkontexte erweitern sich bis in globale Verflechtungszusammenhänge, berücksichtigt man, dass die für die Gießerei verwendeten Schrottteile aus dem Abfall ausrangierter Autos und Kleinbusse aus verschiedenen europäischen Ländern stammen und nach Ende ihres zweiten Lebens zum Futter für die Schmelzöfen werden. Der lokale Arbeitsplatz (vor dem Hochofen) ist Teil eines bestimmten Arbeitsumfelds (des Schmelzwerks), der wiederum in den größeren Komplex (das Suame Magazine) integriert ist und an lokalen, überregionalen und globalen Wirtschaftsverflechtungen partizipiert. Diese Zusammenhänge verweisen auf die »globalen Mikrostrukturen« (Knorr Cetina/Bruegger 2002; Heath 2015), die Verbindungen zwischen der im Video untersuchten situativen Ordnung und der Arbeitsteilung mit den globalen Strukturen der »kapitalistischen« Warenproduktion und -distribution auf internationaler Ebene herstellen.

Selbst wenn die Videoaufzeichnungen in methodisch kontrollierter Weise angefertigt und Teil eines ausgeklügelten Forschungsdesigns sind, wohnt ihnen fraglos eine Spur an Serendipität (Merton/Barber 2004) inne. Sie bilden in diesem Fall die Grundlage für Anschlussstudien, wobei die Analyse der obigen Sequenzen, die ihrem Wesen nach »natürliche« Interaktionsdaten darstellen, erste wichtige Erkenntnisse liefern, denen weitere Forschungen folgen müssen, sei es zur Frage der interaktiven Konstitution von Arbeitsprozessen, zum lokalen Kontext der Produktionsweise oder zum Wissensfluss zwischen Handwerkern und Spezialisten.

Dieses erste Beispiel illustriert die Besonderheiten der Erfassung und Analyse von Feldvideodaten. Wenden wir uns nun einer vollkommen anderen Art von Videodaten zu, die neue methodische Herausforderungen mit sich bringt und einen modifizierten Analyseansatz erfordert.

Beispiel 2: Schrubb-Schrubb (Plattform-Video)

Das zweite Videobeispiel ist ein Auszug aus einem Projekt in der Anfangsphase. Bei dieser Forschung untersuchen wir, wie Wissen durch *Plattformvideos* vermittelt wird. Dabei geht es, wie der folgende Ausschnitt zeigt, um sehr banale Alltagskenntnisse (siehe Abb. 3).

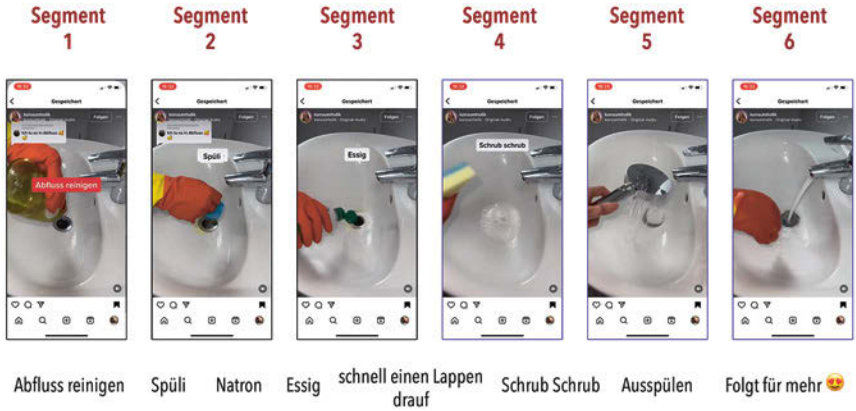


Abbildung 3: Segmentstruktur des Reinigungsvideos (Instagram, Kanal: konsumholik, übernommen von TikTok | Arya.lifestyle | letzter Zugriff 16. Juli 2023)

Das Video enthält eine Kurzanleitung zur Reinigung eines Handwaschbeckens. Es stammt von der Plattform Instagram und ist Teil eines sog. *Reel*, das sich Reinigungstipps und Hilfen zur Alltagsorganisation widmet. Offensichtlich handelt es sich um ein gewöhnliches Problem, für dessen Lösung keine Spezialkenntnisse erforderlich sind. Das Video von etwa elf Sekunden Länge hat eine beschleunigte Abspielgeschwindigkeit. Es zeigt eine Hand in einem gelbroten Gummihandschuh, die mit verschiedenen Reinigungsmitteln ein Waschbecken säubert (Abb. 3). Die erste Szene wird von Instrumentalmusik begleitet und präsentiert die Nahaufnahme eines weißen Waschbeckens. Im Zentrum ist ein Abfluss zu sehen und in der oberen rechten Ecke Teile einer Einhebelmischbatterie. In weißer Schrift auf rotem Hintergrund prangt in der Bildmitte die Zeile »*Abfluss reinigen*«. Am linken oberen Bildrand erscheint das Signet des Bloggers bzw. der Bloggerin mit dem Namen »Komsumholik«, mit 95.600 Followern (Stand 16.07.2023). Veröffentlicht wird dieser Inhalt als Antwort auf einen Kommentar von »nilomed«, wobei der von zwei Smiley-Emojis begleitete Satz »*Ich tu es* in *Abfluss*« hinzugefügt wird. Es handelt sich um einen Repost bzw. Reupload von TikTok, wo er zuerst von einer laut Profilbeschreibung auf »Reinigung und Haushalt« spezialisierten Nutzerin namens »Arya.lifestylee« mit 300.700 Follower gepostet wurde. Der Clip folgt einer klaren Rezeptstruktur und ist in acht Segmente unterteilt:

1. Abfluss reinigen [= Titel]
2. Spüli
3. Natron
4. Essig
5. schnell einen Lappen drauf
6. Schrubb-Schrubb
7. ausspülen
8. folgt für mehr 😍

Zur detaillierten Kompositionsanalyse unterscheiden wir zwischen den *Abschnitten*, aus denen der Clip besteht, d. h. den technischen und ästhetischen Elementen der Filmaufnahme, und den *Segmenten* der dargestellten Reinigungshandlung (Abb. 4). Beide sind eng miteinander verbunden, aber nicht identisch. Die Abschnitte sind Teil der filmischen Darstellung, d. h. sie gehören zur ästhetischen bzw. medialen Ebene. Die Segmente hingegen sind Elemente, die sich auf die im Clip dargestellte Reinigungsaktion beziehen, d. h. sie gehören zur Ebene der Handlung selbst.

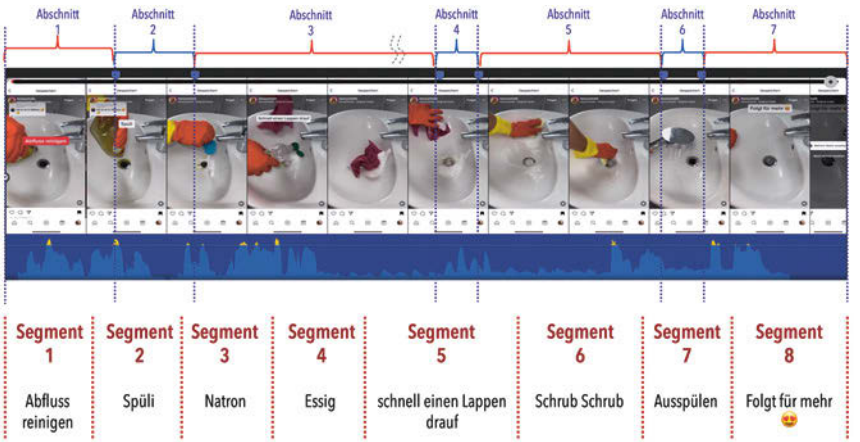


Abbildung 4: Komposition der Abschnitte und Segmente des Videos (eigene Darstellung)

Im ersten Segment sehen wir eine Hand in einem gelbroten Handschuh, die eine Flasche mit Spülmittel hält. Die Hand richtet deren Düse zunächst auf den Abfluss, bevor in lebhaft kreisenden Bewegungen ein dünner Strahl Spülmittel in das Waschbecken gespritzt wird. Die Einblendung »*Abfluss reinigen*« verschwindet, sobald das Deonym »*Spüli*« im Untertitel dieser Eröffnungssequenz erscheint.

Das zweite Segment beginnt mit einem Schnitt und unveränderter Kameraeinstellung. Nun sehen wir die Hand in einer anderen Position, jetzt aus der linken Bildmitte kommend, einen kleinen durchsichtig-blauen Messbecher haltend, ähnlich dem für Waschmittel. Er enthält ein weißes Pulver, das in Segment drei in kreisenden Bewegungen in den Abfluss gestreut wird.

Erneut durch einen direkten Schnitt getrennt, erscheint im vierten Segment die Hand – jetzt aus einer etwas niedrigeren Position – in der linken Bildmitte und umfasst nahezu vollständig eine Essigflasche mit geöffnetem Deckel. Die Flasche wird auf den Abfluss gerichtet und ihr Inhalt in identischen kreisförmigen Bewegungen ausgegossen; sofort kommt Schaum aus dem Abfluss. Gleichzeitig als die behandschuhte rechte Hand die Flasche anhebt und aus der Ebene verschwindet, kommt die linke Hand mit einem lila Tuch zum Einsatz. Es wird zusammengeknüllt sofort auf den Abfluss gelegt, um ihn vollständig zu bedecken.

Im fünften Segment sind keine Hände im Bild. Stattdessen sehen wir Schaum, der aus dem Überlauf entweicht und seitlich über das Tuch gleitet. Dieses Segment enthält einen Schnitt, was sich an der Geschwindigkeit und Menge des austretenden Schaums zeigt. Nach einem weiteren Schnitt kommt die behandschuhte Hand wieder aus der oberen linken Ecke der Bildebene, ergreift das Tuch und entfernt es zusammen mit einem Teil des anhaftenden Schaums aus dem Becken.

In Segment sechs erscheint die Hand wieder und zeigt mit der Handfläche nach oben einen von drei Fingern gehaltenen gelben Schwamm mit blauem Scheuerpad in die Kamera. Die Hand führt den Schwamm schnell in die Mitte, wo sich der Abfluss befindet, und reinigt das Becken mit energischen Zickzack-Bewegungen. Die Aktion wird von Musik begleitet und mit dem lautmalerischen Ausdruck »schrubb-schrubb« untertitelt, der an Refrains gängiger Kinderlieder über das Zähneputzen oder humorvolle Anweisungen zur Handhygiene erinnert. Man könnte sagen, dass der insgesamt als recht lehrreich zu bezeichnende Clip hier einen scherzhaften und infantilisierenden Akzent erhält.

In Segment sieben kommt die nackte linke Hand zum Einsatz, die einen Brausekopf hält, aus dem Wasser strömt, womit das Waschbecken in konzentrischen Bewegungen linksseitig und um den Abfluss herum ausgespült wird. Anschließend wird der Brausekopf nach links aus dem Bild herausgezogen. Nach dem nächsten Schnitt sehen wir, wie sich die rechte, behandschuhte Hand zurückzieht. Diesmal fließt Wasser aus dem Wasserhahn und das Waschbecken ist nun mit einem Stöpsel versehen. Die Hand korrigiert die Position des Stopfens und drückt die Einhebelmischbatterie schnell herunter, sodass der Wasserfluss stoppt. Die Hand zieht sich schnell zurück und es wird der Satz »Folgt für mehr« eingeblendet, der auch verbalisiert wird.

Das letzte Segment acht ist mit den Instagram- typischen Serialitätsmarkern »Folgt für mehr« und dem Emoji »Lächelndes Gesicht Mit Herzförmigen Augen« versehen. Syntaktisch dient es als Abschlusszeichen, mit dem der Clip nach elf Sekunden endet. Nun erscheinen zwei vom System generierte Linkbuttons, mit den Optionen »weitere Reels ansehen« und »noch einmal ansehen«.

Der Sequenzanalyse des Clips sind folgende Beobachtungen hinzuzufügen: Es handelt sich um ein *sauberes* Waschbecken ohne offensichtliche Verschmutzungen, was den demonstrativen Lehrcharakter des Videoclips unterstreicht. Darüber hinaus zeichnet sich der Clip in filmästhetischer Hinsicht durch eine Reihe audiovisueller und textueller *Redundanzen* aus: Handlungen werden (1.) visuell sichtbar gemacht, gleichzeitig (2.) mit Untertiteln in Textform präsentiert und (3.) synchron dazu verbal artikuliert, was eine dreifache Darbietung identischer Informationen darstellt.

4. Methodische Anschlussüberlegungen

Für dieses zweite Beispiel haben wir methodisch auf die Hermeneutik zurückgegriffen, die sich von der ethnomethodologisch fundierten Interaktionsanalyse deutlich unterscheidet und in der Lage ist, die in der Analyse des ersten Feldvideobeispiels hinausgehende Aspekte des Videos zu berücksichtigen. Denn bei Plattformvideos ist es methodisch nötig, nicht nur den gefilmten Kontext zu analysieren, sondern auch die filmästhetischen Ebenen. Sie sind für die Interpretation von zentraler Bedeutung. Im Waschbeckenvideo fallen z. B. seiner Repräsentationseinheit zum Trotz verschiedene Repräsentationsebenen auf, die nacheinander analysiert werden müssen: Die erste bezieht sich auf die Analyse der Abfolge der *im* Video festgehaltenen *Handlungen*: Die einzelnen Schritte der Waschbeckenreinigung, die als Reinigungstätigkeit an sich die erste Analyseebene bildet. Auf der zweiten Ebene wird diese Sequenzialität des Gezeigten von einer weiteren Sequenzialität überlagert. Erforderlich wird hier eine detaillierte Aufschlüsselung der Verwendung filmästhetischer Mittel, wie Kameraeinstellung, Schnitt, Einblendungen, Musik, Voice-over usw. Wenngleich sie in diesem Clip in sehr schlichter Form verwendet werden, gehören diese Ausdrucksmittel zur Gruppe kinematografischer Stilelemente. Daher ist es in diesem Fall unerlässlich, kinematografische Analysedimensionen einzubeziehen, wie Bildausschnitt, Fokus, Planimetrie, Bildkomposition, Dramaturgie, und weitere filmstilistische Merkmale. In dieser Erweiterung des Analysefokus gegenüber Feldvideos liegt der methodische Hauptunterschied bei der Interpretation des von Plattformen stammenden Videomaterials.

Wie festgestellt wurde, weist die für beide Videodatensorten verwendete Methodik sowohl große Ähnlichkeiten als auch deutliche Unterschiede auf: In beiden Fällen wird eine Form der *Sequenzanalyse* verwendet, die gründliches Vorgehen erfordert und jede Einheit in seine kleinsten Ausdrucksbestandteile herunterbricht. Bei Feldvideos konzentriert sich die ethnomethodologischen Grundsätzen folgende Sequenzanalyse ausschließlich auf den Video*inhalt*, d. h. auf die Sequenzabfolge der im Video dokumentierten Interaktionen. Bei der Analyse von Plattformvideos hingegen kommt eine Sequenzanalyse der filmischen Darstellungs*formen* hinzu, für die an die Hermeneutische Wissenssoziologie (Soeffner 2004) angeknüpft wird.

Sequenz im weitesten Sinne des Wortes bezieht sich auf eine zeitliche Abfolge von einander nachgeordneten Elementen. Die unterschiedlichen qualitativen Analysemethoden verwenden den Begriff jedoch für sehr verschiedene Sinneinheiten. Maiwald (2005) erläutert die diversen Analyseprinzipien ethnomethodologischer und hermeneutischer Varianten der Sequenzanalyse. Beide führen minutiöse *Schritt-für-Schritt-Analysen* des Videomaterials durch, allerdings mit zwei wesentlichen Differenzen: Diese betreffen zum einen die Berücksichtigung des Kontexts, zum anderen das Sequenzverständnis.

Während die ethnomethodologische Version sich allein auf den *faktischen* Interaktionskontext beschränkt, hebt die hermeneutische Version ihn auf, indem sie multiple *imaginäre* Lesarten vorstellbarer Kontexte entwirft, um damit nicht offenkundige Sinnpotenziale im Material aufdecken zu können. Die ethnomethodologisch fundierte Sequenzanalyse konzentriert sich konsequent auf die Analyse der schrittweisen Entfaltung der im Video abgebildeten Interaktionen und lässt deren videografische Darstellung unberücksichtigt. Man geht davon aus, dass die aufgezeichneten Interaktionen in bestimmten, von den Forschenden zu analysierenden Abfolgestrukturen vonstattengehen. Die ethnomethodologische Analyse untersucht also die sequenzielle Organisationsstruktur der im Video festgehaltenen aufeinanderfolgenden Interaktionsketten und konzentriert sich darauf, jedes kleine Detail der ablaufenden Interaktionen zu identifizieren. Diese Analyse setzt voraus, dass die Akteure ihre Handlungen systematisch ausführen, was eine gewisse ›Methodizität‹ impliziert. Die miteinander Interagierenden bringen in und durch ihre Alltagshandlungen eine interaktionale ›Ordnung‹ hervor, die interaktiv erzeugt, einander angezeigt und zugleich verständlich gemacht wird. Die sequenzielle Analyse bezieht sich in diesem Fall einzig auf die Rekonstruktion der zeitlichen Abfolge der gefilmten Interaktionsabläufe selbst. Videos sind in der ethnomethodologischen Videoanalyse kein Forschungs*gegenstand*, sondern dienen ausschließlich als Interaktionsanalyse*instrument*.

In hermeneutisch orientierten Methodenansätzen dominiert hingegen ein anderes Sequenzkonzept. Sequenzanalyse bezieht sich dort auf das Vorgehen bei der Dateninterpretation. Genau wie in der ethnomethodologischen Analysepraxis werden kleinste Datenfragmente einer eingehenden Analyse unterzogen. Aber die Interpretationsweise unterscheidet sich in zweierlei Hinsicht: Erstens werden Sinneinheiten und deren *potenzielle* Sinndimensionen identifiziert. Das Deuten stützt sich dabei auf Alltagswissen, angereichert durch die Perspektive relevanter Expertinnen und Experten, wobei die Aufgabe darin besteht, alle *möglichen* Lesarten zu entfalten, einschließlich der unwahrscheinlichsten. Man kann dies als angewandte Form »soziologischer Vorstellungskraft« (Mills 1959) begreifen, die sensibel ist für die Historizität und kulturelle Überformung der Sinnkonstitution. In diesem Fall bezeichnet »Sequenz« also eine Sinneinheit im Video als medialem Kulturprodukt. Die strikte Regel, das Material in der Reihenfolge seiner Entwicklung zu betrachten, hilft zu vermeiden, frühere Passagen im Lichte späterer zu interpretieren, bevor nicht alle möglichen Interpretationen dieser ersten Sinneinheit entfaltet wurden. Erst dann wird zur folgenden Einheit übergegangen, deren Inhalt Aufschluss darüber gibt, welche unserer Lesarten durch das Material bestätigt, und welche verworfen werden.

Zweitens bezeichnet Sequenz in diesem Ansatz nicht eine Reihe von Einzelbildern des Videomaterials, sondern die in den »produzierten« Bildern vorhandenen Sinneinheiten. Bei den untersuchten audiovisuellen Daten handelt es sich um edierte Medienprodukte, also nachbearbeitetes Videomaterial aus

Filmen, Fernsehaufnahmen oder Internetvideoclips, die Teil einer breiten Palette von Videoformaten sind und als Kommunikationsmaterial kulturell geprägter Ausdrucksformen verstanden werden. Trotz ihrer bemerkenswerten Unterschiede in Länge, Stil oder Format sind sie alle durchstrukturiert und aus multiplen Sinneinheiten komponiert. Dieser Perspektive folgend, reflektiert »Sequenzialität« die von den Produzierenden in das mediale Kulturprodukt eingebettete Sinnstruktur. Ziel der Analyse ist es, die verschiedenen Bedeutungsebenen aufzuzeigen, die dem Videomaterial durch bestimmte filmische Handlungen bei Aufnahme, Schnitt und Nachbearbeitung »eingeschrieben« wurden. Die Untersuchung von Kameraeinstellungen, Kadrierung, Schwenks oder anderen filmischen Stilmitteln verweist dabei auf das Handeln der Produzierenden. Sequenzialität ist damit nicht nur methodisches Prinzip der Sozialforschung im Sinne eines Konstruktes zweiter Ordnung; vielmehr zeigt sich die »produzierte Sequenzialität« in den Analyseeinheiten als Konstrukt erster Ordnung (Schütz 2004).

Hinsichtlich der Unterschiede, die bei der Analyse beider Videotypen zu beachten sind, wird klar, dass die Untersuchung von Plattformvideos im Gegensatz zu Feldvideos zusätzliche empirische Untersuchungsdimensionen nötig macht. Dieser Datentypus erfordert, das gesamte Potenzial des im Video Sicht- und Hörbaren in Bezug auf seine eigene semiotische und kommunikative Struktur, aber auch bezüglich kultureller Referenzen, gesellschaftlicher Semantiken und Intertextualitäten auszulegen. Weitere markante Vor- und Nachteile beider Videomaterialtypen lassen sich wie folgt systematisieren:

(a) *Funktionalität*: Während Feldvideos eine sehr klare, aber eingeschränkte Funktionalität aufweisen, decken Plattformvideos ein breites Spektrum von Funktionen ab, das von Informationsübermittlung, Selbstdarstellung und Selbstinszenierung bis hin zu gezielter Werbung, Verhaltensbeeinflussung und Propaganda (vgl. die Studie über russische Meme-Propaganda von Kaden 2021) reicht.
(b) *Stand der Technik und Forschungserfahrung*: Der Einsatz audiovisueller Mittel wie Film und Video kann auf eine lange Tradition in der visuellen Ethnographie und den Sozialwissenschaften im Allgemeinen zurückblicken (vgl. Knoblauch/Wilke i. d. B.). Demgegenüber weist die Forschung zu Plattformvideos noch erhebliche Lücken auf, wobei die hohe Geschwindigkeit und Volatilität mit der neue Formen von Clips auftauchen und wieder verschwinden eine Konsolidierung methodischer Verfahren erheblich erschweren.
(c) *Kontrolle über die Daten*: Als Produkt der Forschenden selbst unterliegen Feldvideos einer hohen Datensouveränität, im Gegensatz zu Plattformvideos, die von anderen produziert und in diversen Bearbeitungsprozessen nachbearbeitet wurden.

(d) *Kenntnis des Kontexts*: Feldarbeit ermöglicht aufgrund der Nähe zur Situation und zu den Akteuren ein hohes Maß an Wissen über den Kontext, in dem die zu analysierenden Aktivitäten stattfinden. Bei Plattformvideos ist hingegen, abgesehen von Credits oder wenigen im Video enthaltenen Informationen, fast nichts über die Kontexte bekannt, in denen das Material produziert, eingesetzt und rezipiert wird.
(e) *Rechtliche und technische Einschränkungen*: Beide Arten von Aufnahmen unterliegen rechtlichen Restriktionen, die von der Notwendigkeit von *informed consents* bei Videoaufnahmen in der Feldforschung bis hin zur genehmigten Verwendung von Plattformvideos reichen. Das Gleiche gilt für die Technologie, denn der Einsatz von Kameras, Mikrofonen und anderen technischen Geräten nimmt Einfluss auf den Forschungsprozess. Ganz zu schweigen von Speicherwerkzeugen und Spezialsoftware für die Digitalisierung, Transkription und Verarbeitung von Daten. Letztere sind für die Analyse von Plattformvideos unerlässlich.
(f) *Selektivität*: Beide Ansätze werfen Probleme bei der Auswahl von Fragmenten für die Feinanalyse innerhalb eines bestimmten Korpus auf. Dies liegt an der Fülle des Materials, das in den meisten Fällen aus der Feldforschung stammt, und an der großen Menge an Plattformvideos, die im Internet zirkulieren. Im Falle von Plattformvideos ist die Auswahllogik zudem algorithmisch geprägt und nicht ausschließlich das Ergebnis menschlichen Handelns.

Die Analyse von Plattformvideos erzeugt methodische Herausforderungen und verlangt neue technische und methodische Mittel. Eine systematische Kombination von Korpus- und Fallanalyse kann die Stärken der analytischen Prinzipien aus der Ethnomethodologie mit denen der Hermeneutischen Wissenssoziologie verbinden. Plattformvideos bieten der Videoanalyse aus methodischer Sicht die Chance, eine derartige Kombination von Korpus- und Fallanalyse zu testen, für die ich an anderer Stelle plädiert habe (Schnettler 2019). Diese Kombination reagiert auf die besonderen Herausforderungen, die sich aus der Fülle, Volatilität und Modulation von Inhalten der Plattformkommunikation ergeben und kann außerdem als Versuch gesehen werden, neue Wege in der qualitativen Forschung in Bereichen zu beschreiten, die stark von der durch Big Data und künstliche Intelligenz getriebenen Methodenentwicklung geprägt sein werden.

Literatur

Beck, Stefan (1997): Umgang mit Technik. Kulturelle Praxen und kulturwissenschaftliche Forschungskonzepte. Berlin: Akademie.

Dant, Tim (2005): Materiality and society. Maidenhead: Open University Press.

Dant, Tim (2007): The ›Pragmatics‹ of Material Interaction. In: Journal of Consumer Culture 8, H. 1, S. 11–33.

Garfinkel, Harold (1967): Studies in Ethnomethodology. Englewood Cliffs: Prentice Hall.
Goffman, Erving (1983): The Interaction Order. In: American Sociological Review 48, H. 1, S. 1–17.
Goodwin, Charles (1994): Professional Vision. In: American Anthropologist 96, H. 3, S. 606–633.
Gumperz, John J./Hymes, Dell H. (Hrsg.) (1972): Directions in Sociolinguistics. The Ethnography of Communication. New York: Holt, Rinehart, and Winston.
Heath, Christian (1986): Body Movement and Speech in Medical Interaction. Cambridge: Cambridge University Press.
Heath, Christian (2015): The Dynamics of Auction: Social Interaction and the Sale of Fine Art and Antiques. Cambridge: Cambridge University Press.
Heath, Christian/Hindmarsh, Jon/Luff, Paul (2010): Video in Qualitative Research. London: Sage.
Heath, Christian/Luff, Paul (2008): Video and the Analysis of Work and Interaction. In: Alasuutari, Pertti/Bickman, Leonard/Brannen, Julia (Hrsg.): The Sage Handbook of Social Research Methods. London: Sage, S. 493–505.
Jaarsma, Thomas/Maat, Harro/Richards, Paul/Wals, Arjen (2011): The Role of Materiality in Apprenticeship: The Case of the Suame Magazine, Kumasi, Ghana. In: Journal of Vocational Education and Training 63, H. 3, S. 439–49.
Kaden, Tom (2021): Autoritative Macht und politische Einflussnahme. In: Gostmann, Peter/Merz-Benz, Peter Ulrich (Hrsg.): Macht und Herrschaft. Zur Revision zweier soziologischer Grundbegriffe. 2. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 107–131.
Keller, Charles/Keller, Janet Dixon (1996): Cognition and Tool Use. The Blacksmith at Work. Cambridge: Cambridge University Press.
Kensok, Peter (1987): Fitter. Entwicklung aus der Werkzeugkiste. Informelle Ausbildung von Kraftfahrzeughandwerkern in Ghana. Saarbrücken: Breitenbach.
Knoblauch, Hubert (2001): Fokussierte Ethnographie. In: Sozialer Sinn 2, H. 1, S. 123–141.
Knoblauch, Hubert (2002): Fokussierte Ethnographie als Teil einer soziologischen Ethnographie: Zur Klärung einiger Missverständnisse. In: Sozialer Sinn 3, H. 1, S. 129–136.
Knoblauch, Hubert/Schnettler, Bernt/Tuma, René (2018): Videography. In: Flick, Uwe (Hrsg.): The Sage Handbook of Qualitative Data Collection. London: Sage, S. 362–377.
Knorr Cetina, Karin/Bruegger, Urs (2002): Global Microstructures: The Virtual Societies of Financial Markets. In: American Journal of Sociology 107, H. 4, S. 905–995.
Luff, Paul/Hindmarsh, Jon/Heath, Christian (Hrsg.) (2000): Workplace Studies: Recovering Work Practice and Informing System Design. Cambridge: Cambridge University Press.
Maiwald, Kai-Olaf (2005): Kompetenz und Praxis: Sequenzielle Analyse in der deutschen Soziologie. In: FQS 6, H. 3, Art. 31. https://doi.org/10.17169/fqs-6.3.21.
Merton, Robert K./Barber, Elinor (2004): The Travels and Adventures of Serendipity: A Study in Sociological Semantics and the Sociology of Science. Princeton: Princeton University Press.
Meyer, Christian/von Wedelstaedt. Ulrich (2017): Intercorporeality, Interkinesthesia, and Enaction. New Perspectives on Moving Bodies in Interaction. In: Meyer, Christian/von Wedelstaedt, Ulrich (Hrsg.): Moving Bodies in Interaction – Interacting Bodies in Motion. Intercorporeality, Interkinesthesia, and Enaction in Sports. Amsterdam: John Benjamins, S. 1–26.
Mills, C. Wright (1959): The Sociological Imagination. Oxford: Oxford University Press.
Pach, Helen (2018): Technik, Interaktion und heißes Eisen. Eine African Workplace Study, unveröffentlichte Masterarbeit, Universität Bayreuth.
Powell, John (1995): The Survival of the Fitter. Lives of some African Engineers. London: Intermediate Technology Publications.
Raab, Jürgen/Maier, Kevin (2022): Sinnstiftung durch Sinnüberschuss: Wissenssoziologische Videohermeneutik als methodisches Verfahren zur Analyse multimodaler Daten. In: Kondratjuk, Maria/Dörner, Olaf/Tiefel, Sandra/Ohlbrecht, Heike (Hrsg.): Qualitative Forschung auf dem Prüfstand. Leverkusen: Barbara Budrich, S. 187–208.
Raab, Jürgen/Tänzler, Dirk (2012): Video Hermeneutics. In: Knoblauch, Hubert/Schnettler, Bernt/Raab, Jürgen/Soeffner, Hans-Georg (Hrsg.): Video Analysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology. 3. Auflage. Frankfurt am Main: Peter Lang, S. 85–97.
Reinbold, Judith (2023): Persuasive Mittel in Corona-Präventionsvideos. Eine gattungsanalytische Untersuchung globaler Gesundheitskommunikation. Wiesbaden: Springer VS.
Sacks, Harvey (1992): Lectures on Conversation. Oxford: Blackwell.

Schnettler, Bernt (2019): Die konträren Logiken von Korpus und Fall: Plädoyer für eine Integration. In: Hitzler, Ronald/Reichertz, Jo/Schröer, Norbert (Hrsg.): Kritik der Hermeneutischen Wissenssoziologie. Weinheim und Basel: Beltz Juventa, S. 126–135.

Schnettler, Bernt/Knoblauch, Hubert/Baer, Alejandro (2012): Videoanálisis interpretativo de situaciones sociales: Etnografía, análisis secuencial y hermenéutica. In: Arroyo Menéndez, Millán/Sabada Rodríguez, Igor (Hrsg.): Metodología de la investigación social. Tecnicas innovadoras y sus aplicaciones. Madrid: Editorial Síntesis, S. 251–269.

Schröer, Norbert (Hrsg.) (1994): Interpretative Sozialforschung. Auf dem Wege zu einer hermeneutischen Wissenssoziologie. Opladen: Westdeutscher.

Schütz, Alfred (2004): Common-Sense und wissenschaftliche Interpretation menschlichen Handelns. In: Strübing, Jörg/Schnettler, Bernt (Hrsg.): Methodologie interpretativer Sozialforschung. Klassische Grundlagentexte. Konstanz: UVK/UTB, S. 157–197.

Singh, Ajit (2019): Wissenskommunikation im Sport. Zur kommunikativen Konstruktion von Körperwissen im Nachwuchstraining. Wiesbaden: Springer VS.

Soeffner, Hans-Georg (2004): Auslegung des Alltags – Der Alltag der Auslegung. Zur wissenssoziologischen Konzeption einer sozialwissenschaftlichen Hermeneutik. 2. Auflage. Konstanz: UVK.

Spittler, Gerd (2016): Anthropologie der Arbeit. Ein ethnographischer Vergleich. Wiesbaden: Springer VS.

Spittler, Gerd (2018): Anthropologie du travail: des classiques à la comparaison ethnographique. In: Musso, Pierre/Supiot, Alain (Hrsg.): Qu'est-ce qu'un régime de travail réellement humain. Paris: Hermann, S. 149–160.

Streeck, Jürgen/Goodwin, Charles/LeBaron, Curtis (2013): Embodied Interaction in the Material World. An Introduction. In: Streeck, Jürgen/Goodwin, Charles/LeBaron, Curtis (Hrsg.): Embodied Interaction. Language and Body in the material World. Cambridge: Cambridge University Press, S. 1–28.

Tuma, René (2012): The (Re)Construction of Human Conduct: »Vernacular Video Analysis«. In: Qualitative Sociology Review 8, H. 2, S. 152–163.

Tuma, René (2017): Videoprofis im Alltag. Die kommunikative Vielfalt der Videoanalyse. Wiesbaden: Springer VS.

Tuma, René/Knoblauch, Hubert/Schnettler, Bernt (2013): Videographie. Einführung in die interpretative Video-Analyse sozialer Situationen. Wiesbaden: Springer VS.

vom Lehn, Dirk (2018): Ethnomethodologische Interaktionsanalyse. Weinheim und Basel: Beltz Juventa.

Wilke, René (2022): Wissenschaft kommuniziert. Eine wissenssoziologische Gattungsanalyse des akademischen Group-Talks am Beispiel der Computational Neuroscience. Wiesbaden: Springer VS.

III. Interaktionsanalysen mit Ethnographischen Videos

1. Interaktion in Paarbeziehungen, (Klein-)Gruppen, Kollektiven und der Öffentlichkeit

Videodaten bei der Analyse der Kommunikation mit Menschen mit der Diagnose Demenz

Jo Reichertz und Anna-Eva Nebowsky

Zusammenfassung: In dem Beitrag wird über die Feldforschung in dem DFG-Projekt ›Kommunikation mit Menschen mit der Diagnose Demenz‹ berichtet. Im Rahmen dieser Feldforschung wurden (nach einer Phase des Miteinanderbekannt-Werdens und des Wachsens von Vertrauen) Videos von alltäglichen Handlungen erstellt. Das Besondere der Videos ist, dass i. d. R. auch die Feldbeobachter*innen und deren Interaktion mit den Beteiligten zu sehen sind. Die Feldbeobachter*innen sind, da sie die Familien über einen langen Zeitraum regelmäßig besucht haben, keine fremden Wissenschaftler*innen hinter der Kamera, sondern aktiv in die Handlungen eingebundene Personen. Alle handeln also auch vor dem Hintergrund einer zurückliegenden Interaktionsgeschichte, also einem Kontext, welcher mit der laufenden Interaktion fortgesetzt wird. In dem Beitrag wollen wir neben der Darstellung der Erhebung und der Auswertung solcher Daten auch diskutieren, welche Bedeutung diese Eingebundenheit der Forscher*innen für die Forschung, aber auch für die Nachnutzung der Daten hat.

Schlüsselwörter: Demenz, beobachtende Teilnahme, Kommunikation, Beziehungszeichen

1. Das Projekt und seine Fragestellung

In einem von der DFG geförderten Projekt beobachteten wir über drei Jahre (von Oktober 2019 bis Oktober 2021) Menschen, die damit leben, dass Mediziner*innen sie mit der Diagnose ›Demenz‹ versehen haben.[1] Diese Diagnose verändert das Leben der so Diagnostizierten und das Leben aller, die mit ihnen eine Beziehung haben, gravierend. Allerdings nicht mit einem Mal und nur einmal, sondern die Diagnose setzt einen komplexen sozialen Prozess in Gang, in dem alle

1 Siehe Reichertz/Keysers/Nebowsky (2020), Roth/Reichertz (2020), Keysers/Nebowsky (2020), Nebowsky (2024), Reichertz (2020b, 2021a, 2024). Eine ausführlichere Beschreibung des Projekts findet sich in Reichertz (2024).

Beteiligten immer wieder neu für sich Wege finden müssen, ihr Leben und ihre Handlungen aufeinander abzustimmen (vgl. Frewer-Graumann 2020).

Aktuell scheint die Diagnose *Demenz* das Schlimmste zu sein, was einem widerfahren kann, sowohl als Selbst-Betroffene*r als auch als Angehörige*r.[2] Mit der Diagnose werden durch Medien und Erzählungen von Freund*innen und Bekannten genährte Erwartungen aktiviert, die davon ausgehen, dass die jetzt offiziell zugeschriebene Krankheit *irreversibel* ist, über kurz oder lang zu massiven Beeinträchtigungen der kognitiven und kommunikativen Fähigkeiten und damit zu einer eingeschränkten Teilnahme am Leben führt, aber vor allem den Verlust eines erfüllten und glücklichen Lebens zur Folge hat. Alle Betroffenen müssen dann erst einmal und vor allem ihren Alltag neu praktisch bewältigen. *Während* und *indem* sie dies tun, müssen sie darüber hinaus vor allem folgende Fragen immer wieder kommunikativ mit- und gegeneinander bearbeiten (siehe hierzu auch: Bryden 2011; Swaffer 2016; Steffens 2023; Taylor 2008): *Verläuft das Leben nach der Diagnose ›Demenz‹ bei mir/uns anders als bei anderen? Wer bin ich (noch)? Wer bist du (noch)? Wer bin ich (noch) für dich? Wer bist du (noch) für mich? Was kann ich (noch)? Wer werde ich sein, wenn ich nicht mehr ich bin? Wer wirst du sein, wenn du nicht mehr du bist?* Natürlich stellen sich neben diesen Fragen noch ökonomische und soziale – und auch die Frage nach dem Sinn des (Weiter-)Lebens, nach dem Tod und was nach dem Tod zu erwarten ist – unabhängig davon, ob man konfessionell gebunden ist oder nicht.

Bei allen diesen Prozessen ist das kommunikative Handeln der Beteiligten entscheidend: Denn indem sie immer wieder miteinander über sich, ihre Beziehung und die Welt kommunizieren, *erschaffen* sie diese immer auch wieder (Reichertz 2009 und 2021b; Keller/Knoblauch/Reichertz 2013). Damit soll nicht gesagt werden, dass Demenz allein das Ergebnis einer Zuschreibung ist, sondern dass es neben einem biologischen Prozess der Degeneration von Neuronen (den die Mediziner*innen als ursächlich für die Demenz ansehen[3]) auch immer einen sozialen und kommunikativen Prozess gibt, der die Ausprägung des biologischen Prozesses massiv überformt. Die hier vorgenommene Betonung der sozialen Seite dieses Prozesses (so sprechen wir durchgängig von Menschen mit der Diagnose Demenz) soll nun keinesfalls die Wirklichkeit neurodegenerativer

2 *Angehörige* sind alle die Personen, die mit dem Menschen mit Demenz existenziell verbunden sind, die sein Leben bisher geteilt und mitgestaltet haben, die mit diesem Menschen eine Welt aufgebaut, mit ihm gelebt, ihn geliebt, möglicherweise auch unter ihm gelitten haben, kurz: alle die Personen, die zu dem Menschen mit Demenz eine unaufgebbare Beziehung haben, die zwar durch Liebe, oft auch durch ambivalente Gefühle gekennzeichnet ist.

3 Der Einfachheit halber sprechen wir durchgehend nur von neurodegenerativen Prozessen. Allerdings ist diese Formulierung so nicht richtig, da es verschiedene Formen der Demenz gibt (z. B. vaskuläre Demenz), die nicht auf neurodegenerative Veränderungen des Gehirns zurückgehen. Zu dem Krankheitsverlauf, dessen Phasen und möglichen präventiven Maßnahmen siehe Kruse 2017 und hier insbesondere 326 ff.

Prozesse in Abrede stellen noch deren Auswirkungen auf die kognitiven und kommunikativen Fähigkeiten des*der so Diagnostizierten. Unstrittig ist, dass die direkt Betroffenen schrittweise ihre sprachlich gebundene kommunikative Kompetenz (Taylor 2008; Steffens 2023) verlieren. Das heißt, schrittweise verlieren sie die biographisch erworbene Fähigkeit, die Grammatik, die Semantik und die Pragmatik einer Sprach- und Interaktionsgemeinschaft regelgerecht einzusetzen. Damit einher geht oft der Verlust der Fähigkeit, Worte zu artikulieren, und der Fähigkeit, die Perspektive des anderen zu übernehmen mit dem Ziel, kommunikative Handlungen aufeinander abzustimmen. Zudem verändert sich massiv die Kommunikationsmacht gegenüber den jeweils relevanten Anderen (Bryden 2011; Bryden 2018; Swaffer 2016).

In dem Projekt untersuchten wir, wie die Prozesse der kommunikativen Handlungsabstimmung in Familien aussehen, in denen eine*r mit der Diagnose Demenz lebt und wie sich diese Prozesse im Laufe der Zeit verändern. Dabei untersuchten wir immer die gesamte *Konstellation*. Damit ist nicht nur das Ehepaar gemeint und die gesamte *Figuration* der Personen, die zu dem betroffenen Paar in Beziehung standen, sondern dazu gehören auch die Wohnsituation (Wohnung oder Eigenheim), die Haustiere, dazu gehört die finanzielle Lage, die Anbindung an versorgungsrelevante Institutionen und medizinische Einheiten, dazu gehört aber auch die gesamte Infrastruktur mit Kultur- und Sportangeboten. Zu dieser Ordnung im Raum gehört auch die Betrachtung der Ordnung der Räume und der Ordnung in den Räumen.

2. Erhebungskontext: Das Projekt und seine Vorgehensweise

Um dieses Ziel zu erreichen, nahmen die Projektmitarbeiter*innen über einen Zeitraum von drei Jahren an einem kleinen Ausschnitt des Alltags dieser Menschen teil – betrieben also teilnehmende Beobachtung oder besser: *beobachtende Teilnahme*. Während der Beobachtung sprachen wir mit *allen* Beteiligten in unterschiedlichen Formaten: Small Talk, Unterhaltung, Gespräche mit unterschiedlichen Schwerpunkten. Wir führten *Gespräche*, also keine Interviews – immer wieder, in unterschiedlichen Situationen, in unterschiedlichen Konstellationen. Wir schauten uns zusammen *Fotoalben* und *Filme* an, machten auch Fotos – und wenn das Vertrauen gefestigt war, auch *Videos* von alltäglichen Handlungen – wie: zusammen essen, sich unterhalten, Einkaufen gehen, eine Ausstellung besuchen, den Tag zusammen verbringen (zu den Möglichkeiten und Problemen des Filmens in dem Zuhause von anderen siehe auch Silverman 2016[4]). Von allen

4 »I had also planed to be out of sight, retreat to a corner and discreetly moving the camera should the action change location. Yet during the actual visits this strategy proved impossible. Filming lived experience required dynamic active interaction and movement on

Besuchen fertigten die Projektmitarbeiter*innen teils ausführliche Memos an, oft mehrere.

Die Videos nahmen bei unseren Besuchen, aber vor allem bei unseren späteren Datenauswertungen eine besondere Stellung ein – insbesondere, weil in den von uns beobachteten Familien die verbale Kommunikation immer weniger von Bedeutung war, sondern stattdessen die körperliche Kommunikation zwischen den Ehepartner*innen zunahm und zentral für die Handlungsabstimmung wurde. Deshalb war von entscheidender Bedeutung, den Prozess der Handlungsabstimmung genauer zu beobachten – d. h. nicht nur Fragen zu stellen und Interviews führen, sondern genau hinzuschauen und Videos zu analysieren und hier vor allem den Mikro- und Nanobereich der Interaktion in den Blick zu nehmen. Mit *Nanobereich* ist die Ausdrucksebene sozialer Interaktion gemeint, bei der die einzelnen bedeutungstragenden Einheiten entweder von so *kurzer Dauer* sind oder aber sich in *minimalen Veränderungen* zeigen, dass sie entweder für die normale wissenschaftliche Beobachtung (Auge, Ohr, Gespür) nicht wahrnehmbar, aber auf jeden Fall kaum erinnerbar und damit auch nicht für die Analyse verfügbar sind. Die Entdeckung und Vermessung des Nano-Bereichs ist aus unserer Sicht für die qualitative Sozialforschung im Allgemeinen und für die Untersuchung der Kommunikation von Menschen mit der Diagnose Demenz im Besonderen von enormer Bedeutung und sie bringt einen enormen Wandlungsschub mit sich, der unseres Erachtens nur mit dem Schub zu vergleichen ist, der mit der Erfindung und Einführung der Tonbandgeräte einherging (vergleichbare Videoanalysen finden sich bei Döttlinger 2017; Welling 2014; Welling 2018).

Die Feldbeobachter*innen waren, da sie die Familien über einen langen Zeitraum regelmäßig besuchten, keine fremden Wissenschaftler*innen hinter der Kamera, sondern aktiv in die Handlungen eingebundene Personen. Alle handelten also auch vor dem Hintergrund einer zurückliegenden Interaktionsgeschichte, also einem Kontext, welcher mit der laufenden Interaktion fortgesetzt wurde.

Teilnehmende Beobachtung und beobachtende Teilnahme

In der Feldforschung gibt es einen teilweise erbittert geführten Streit darüber, ob Sozialwissenschaftler*innen besser mit *teilnehmender Beobachtung* arbeiten oder mit *beobachtender Teilnahme – also der lebensweltlichen Ethnografie* (ausführlich Honer 2011; Reichertz 2016). Für beide Verfahren gibt es gute Gründe: Für die teilnehmende Beobachtung spricht, dass man im Feld einen kühlen Kopf und ein

---

my part […]. For example, my embodied awareness of heat, smells and sounds affected the way I moved with the camera. This embodied awareness led to and enhanced visceral understanding of the environment that in turn contributed to deeper analysis, as well as to the moment to moment choices of where I set up the camera and how I filmed« (Silvermann 2016, S. 577 ff.).

klares Auge bewahrt – neutral bleibt und mit freischwebender Aufmerksamkeit alles wahrnimmt, was sich ereignet –, fast so wie eine Kamera. Kritiker*innen der teilnehmenden Beobachtung (Honer 2011; Hitzler/Eisewicht 2016) wenden zurecht gegen diese ›kühle‹ Art der Teilnahme ein, dass man so zwar alles sehe, was es an den Oberflächen der Anderen zu sehen gebe, dass einem aber gerade das Systematische entginge, was den Reiz z. B. eines Ortes und des Handelns an diesem Ort ausmache. Nur der*die Beobachtende (so das Argument weiter), der*die in sich das Gefühl auslösen könne, das die Feldangehörigen durch ihr Tun tatsächlich erleben, könne dem wirklichen Reiz des Tuns und damit den echten Motiven für dieses Tun auf die Spur kommen. Aber es gibt noch einen weiteren Grund, sich auf die Anderen im Feld mit all seinen Sinnen einzulassen: denn sich einzulassen und es nicht nur zu beobachten zeigt den Anderen, dass man sie ernst nimmt, dass man *Respekt* vor ihnen und ihrer Welt hat. Ohne diesen Respekt gibt es kein Vertrauen und auch keine wertvolle Auskunft, keine ›tiefe‹ Erzählung und somit auch keinen echten Einblick in das Feld. Beobachtende Teilnahme verbessert also die Datenqualität: Einerseits erlebt der/die Forscher*in *in sich und an sich*, was die Beobachteten bewegt, andererseits lassen die Beobachteten die Forschenden nur dann in die inneren Bereiche ihres Feldes, wenn sie sich auf das Feld wirklich radikal einlassen, dem *Feld also Respekt erweisen* – auch indem sie sich in die Welt der Anderen existenziell verstricken lassen und somit *zum Teil* selbst erleben, was die Anderen erleben.

Die Frage, welche Art von Beobachtung wir im Projekt vornehmen sollten, war schnell beantwortet. Schon allein deshalb, weil man in einer Privatwohnung nicht einfach schweigen und nur den anderen Anwesenden zuschauen kann. Man kann sich dort nicht unsichtbar machen. Man ist deutlich sichtbar anwesend. Man ist gerade keine *Fliege an der Wand*, sondern ein Mensch, der im Weg steht, der Fragen stellt, den man bittet, etwas zu tun, der einen anschaut, der reagiert, kurz: einer, der dabei ist und mitmacht. Man kann nicht jemanden, in dessen Wohnung man sich stundenlang aufhält, sagen: »Bitte tun Sie so, als sei ich nicht da!« Denn dieser kann nur sagen: »Wie sollte ich das können, Sie sind ja in meiner Wohnung«. Hier gilt der Satz von Watzlawick ohne Einschränkung: Unter der Bedingung der gegenseitigen Wahrnehmung kann man nicht *nicht* kommunizieren: man wird zur relevanten Interaktionspartner*in, um den/die sich beide Betroffene bemühen. Von daher war die Frage nach teilnehmender Beobachtung und beobachtender Teilnahme schnell akademisch: Praktisch verwandelte sich jede unserer Beobachtungen in eine *beobachtende Teilnahme* – ob man das wollte oder nicht: Man war immer mittendrin! *Beobachten*, *Bezeugen* und *Mitwirken* flossen schnell und ununterscheidbar ineinander. Die beobachtende Teilnahme war also unausweichlich bzw. es gab überhaupt keine Alternative dazu. Dies nicht nur wegen des häufigen und engen Kontakts, sondern vor allem deshalb, weil *einige* der beobachteten Menschen mit der Diagnose Demenz mit der Normalform teilnehmender Beobachtung und deren Konsequenzen für die Interaktion

entweder nicht oder nicht mehr vertraut waren. Das führte regelmäßig dazu, dass die wissenschaftlichen Beobachter*innen wie normale Besucher*innen behandelt wurden und entsprechend in die Interaktion und die damit einhergehenden Interaktionsverpflichtungen hineingezogen wurden.

3. Analysekontext: Beobachtung ist immer auch Intervention

Unser Beobachten war also immer mehr als Hinsehen und Beobachten und Registrieren und Merken, sondern unsere Art der Beobachtung war immer auch *Intervention*. Dies deshalb, da unser Projekt erklärtermaßen das kommunikative Handeln fokussierte und deshalb alle Beobachteten nicht mehr selbstverständlich wie gewohnt kommunizierten, sondern bewusst(er), also anders. Die Anwesenheit der Beobachter*innen war auch deshalb Intervention, weil deren Anwesenheit und die Kommunikation mit ihnen die Lebenssituation der Beobachteten (und der Beobachter*innen) teils maßgeblich änderte. Wenn man so will: Der/die Beobachter*in ist als ›neue Dritte im Leben‹ der Beobachteten: Er/sie wurde in jeder Hinsicht *signifikante Andere* im Sinne Meads (Mead 1973).

Auch wenn die Beobachter*innen es nicht wollen, die Beobachteten werden sich um sie bemühen, werden Koalitionen erwünschen oder Abneigungen ausleben. So bildeten die Ehepartner*innen manchmal eine Koalition gegen die Beobachter*in, aber es konnte auch sein, dass der Mann und die Beobachter*in eine Koalition gegen die Ehefrau bildeten. So konnten Gefühle der Vernachlässigung oder Zurückgewiesenheit aufkommen oder ein Angenommensein von dem/der Besucher*in. Es gab Unmut und Verbundenheit, Zurückweisung oder Zustimmung – auf jeden Fall schaffte der oder die neue ›Dritte‹ eine neue Beziehungskonstellation mit einer Dynamik, die schwer zu kontrollieren und noch weniger gezielt einzusetzen war.

Die ›Intervention‹ durch die Beobachtenden war jedoch auch noch aus anderen Gründen unserer Forschung zweckdienlich: Sie ermöglichte uns nicht nur, auf Basis des aufgebauten Vertrauensverhältnisses vergleichsweise intime Daten zu gewinnen. Sie ermöglichte uns auch, genau die Daten zu gewinnen, die unsere Forschungsfrage beantworten sollten. Die Anwesenheit des/der Beobachter*in provozierte immer wieder Koalitionen *der Partner*innen selbst*. Da sie wussten, dass der/die Forscher*in sie insbesondere in ihrer Funktion als Ehepaar beobachtete, versuchten die Ehepartner*innen dem ›Bild‹ eines Ehepaares zu entsprechen. Anders ausgedrückt: Die Partner*innen produzierten laufend *Beziehungszeichen*, sogenannte ›tie signs‹ (Goffman 1982), mittels derer sie ihre verankerte Beziehung vor dem/der Forscher*in evident bzw. beobachtbar machten. Beziehungszeichen sind alle Beziehungsbekundungen eines Paares, die nicht »wörtlich gemeinte, explizite dokumentarische Aussagen« sind, sondern »Gegenstände, Handlungen oder Expressionen« (Goffman 1982, S. 262). Beziehungszeichen

werden Goffman zufolge vor allem dann hergestellt, »wenn sich das Paar einem generalisierten oder tatsächlichen Dritten gegenüberstehen sieht oder sich unter Beobachtung eines solchen wähnt« (Nebowsky 2024, S. 143).

Gerade bei Demenz, wo die verbalsprachlichen Fähigkeiten häufig beeinträchtigt werden und die sozialen Partizipationsmöglichkeiten entsprechend auch, kann die *beobachtende Teilnahme* des/der Forscher*in die Interaktionsdynamik bestärken, und so auch der Person mit Demenz eine bessere Teilnahme an der Situation ermöglichen (vgl. Nebowsky 2024).

Ein solcher Nutzen wird in der folgenden Videosequenz sichtbar. Die (enge) Anwesenheit der Forschenden führt dazu, dass uns die pflegende Angehörige zeigen möchte, dass ihr Mann (mit fortgeschrittener vaskulärer Demenz) noch mit ihr zusammen singen kann und dies gerne tut. Sie geht ein Liederbuch mit uns durch. In der Videosequenz liest sie den Titel eines Lieds vor: »Vater Unser im Himmel«, was bei ihr sowie bei der Forscherin AN (Co-Autorin dieses Beitrags) für Verwirrung sorgt. Herr Esselmann (Pseudonym) und Forscher TH (Kürzel) jedoch sind darüber nicht verwundert und bilden eine Koalition. In dieser Konstellation fanden die Teilnehmenden bereits das zweite Mal zusammen. Etwa ein halbes Jahr zuvor waren die Forschenden schon einmal zu zweit zur Videoaufnahme bei dem Paar. In der Sequenz wartet die Person mit Demenz, Herr Esselmann, darauf, dass seine Frau (Frau Esselmann) das nächste Lied zum gemeinsamen Singen aus dem Liederbuch heraussucht. Er schaut sie dabei an. Als sie den Titel »Vater unser im Himmel« nennt, lehnt sich Herr Esselmann leicht nach vorne. Die Hände hält er gefaltet vor seinem Gesicht. Sie dienen ihm in unterschiedlichen Kontexten auf verschiedene Art und Weise als Ausdrucksmittel, da er sich verbalsprachlich nicht mehr artikulieren kann (siehe Abb. 1).

Abbildung 1: Herr Esselmann lehnt sich mit vor dem Gesicht gefalteten Händen leicht nach vorne, während Frau Esselmann in das Liederbuch schaut

Die Hände weiterhin vor dem Gesicht gefaltet, dreht sich Herr Esselmann langsam nach links, wo Forscher TH sitzt und schaut ihn an (Abb. 2).

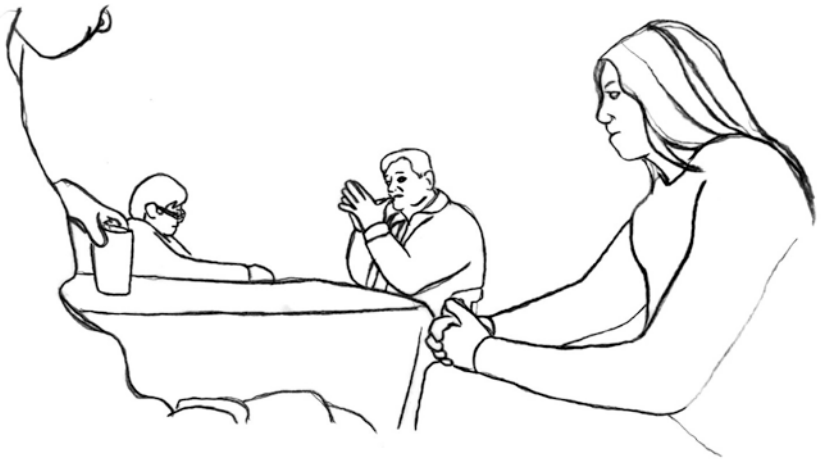


Abbildung 2: Herr Esselmann dreht sich mit vor dem Gesicht gefalteten Händen nach links, wo Forscher TH sitzt

Als Herr Esselmann bemerkt, dass TH ihn ebenfalls anblickt, beginnt er mit den gefalteten Händen auf Höhe seines Munds zu klatschen. TH klatscht daraufhin ebenfalls, während er und Herr Esselmann sich weiterhin anblicken. Herr Esselmanns Hände bewegen sich dabei klatschend, aber in einer Art zeigenden Geste, auf Forscher TH zu. Herr Esselmann öffnet anschließend den Mund, so als wollte er etwas zu TH sagen (Abb. 3).

Abbildung 3: (v. l. n. r.) Herr Esselmann beginnt zu klatschen während er Forscher TH anguckt; er ist sichtlich erheitert und öffnet den Mund so als wollte er etwas sagen

Herr Esselmann und alle anderen beginnen, zu lachen. Frau Esselmann schaut zuerst ihren Mann an, während dieser noch TH anblickt. Schließlich wendet Herr Esselmann den Blick von TH ab und schaut seine Frau ebenfalls an (Abb. 4).

Abbildung 4: Die Anwesenden lachen; Herr Esselmann schaut zu seiner Frau

Frau Esselmann übersetzt anschließend für TH, was die Bedeutung der Gestik und Mimik Herrn Esselmanns war: »Das heißt, Sie müssen mitsingen«.

Die beschriebene Sequenz verweist auf die laufende Ko-Konstruktion von Beziehungszeichen des Paar Esselmanns: Zur Unterhaltung der Anwesenden, oder auch um ihnen zu zeigen, wozu Herr Esselmann trotz seiner fortgeschrittenen Demenzerkrankung noch fähig ist, geht Frau Esselmann ein ihrem Mann bekanntes Liederbuch mit ihm und den Forschenden durch. Herr Esselmann wartet geduldig auf die Liederwahl seiner Frau, während Frau Esselmann seine Gestik und Mimik aufmerksam verfolgt.

Das Paar zeigt damit, dass es aneinander interessiert ist und sich ko-orientiert. Die Anwesenheit der Forschenden erzeugt hier jedoch nicht nur die Aktivität des gemeinsamen Singens, sondern auch ein Moment der initiativen sozialen Partizipation Herrn Esselmanns, der Forscher THs Aufmerksamkeit aktiv sucht und schließlich gestisch-mimisch mit ihm interagiert. Zuletzt verweist Frau Esselmann zusätzlich verbalsprachlich auf die Zusammengehörigkeit des Paares, indem sie Herrn Esselmanns leibkörperliche Aussage bzw. Aufforderung TH gegenüber verbalsprachlich übersetzt. Sie vermittelt dadurch, dass sie ihren Mann besonders gut kennt und einen exklusiven Wissensvorrat mit ihm teilt.

4. Nachnutzung: drei für die Weiterverwendung relevante Besonderheiten der Fälle

Die in der Projektarbeit erstellten Videos weisen im Hinblick auf die spätere Weiterverwendung durch Dritte einige Besonderheiten auf, welche die Weiternutzung erschweren – insbesondere, weil die Anonymität der aufgezeichneten Personen nur schwer bzw. nicht wirklich gesichert werden kann. Und zwar betrifft dies sowohl die von uns beobachteten Menschen als auch die wissenschaftlichen Beobachter*innen.

Erst einmal und vor allem sind auf den Videos Menschen zu sehen, die aufgrund ihrer Erkrankung nicht mehr oder nicht mehr hinreichend in der Lage sind, ein Image-Management erfolgreich durchzuführen. Sie können ihre

Identität, also den Blicken der Anderen, nicht mehr durch ›Gesichts- und Körperarbeit‹ (Goffman 1967) entziehen. In diesem Sinne sind sie (und ihre Identität) also besonders leicht verletzbar – weshalb die Kommunikation mit ihnen in einem sicheren Raum stattfinden sollte, in dem sie auch vor den Blicken anderer Unbekannter geschützt sind. Das gilt auch für wissenschaftliche Beobachter*innen.

Die zweite Besonderheit besteht darin, dass aufgrund der beobachtenden Teilnahme auf den Videoaufzeichnungen nicht nur die beobachteten Personen zu sehen sind, sondern immer auch deren Interaktion mit den wissenschaftlichen Beobachter*innen. Diese sind immer in das Geschehen selbst verstrickt, sind eigene Akteure, die das Geschehen aktiv mitgestalten, die in den Prozess eingreifen, die intervenieren. Daraus folgt, dass neben den Beobachteten auch die wissenschaftlichen Beobachter*innen ihre Einverständniserklärung für die weitere Nutzung der Daten geben müssen, was keineswegs gesichert ist. Das strukturell größere Problem ist allerdings, dass die Anonymität der wissenschaftlichen Beobachter*innen nicht gewährleistet werden kann, da aus dem Projektunterlagen und den Veröffentlichungen ihre Identitäten sehr schnell zu rekonstruieren sind. Eine Verpixelung des Bildes würde also wertlos sein.

Eine dritte Besonderheit der Videodaten ist, dass der Mensch mit der Diagnose Demenz oft nur über ein geringes sprachliches Ausdrucksvermögen verfügt, sodass die Kommunikation mit anderen mithilfe des gesamten Körpers stattfindet. Hier sind der Augenkontakt, die Mimik und die Gestik, aber auch jede Art von Körperbewegung von zentraler Bedeutung. Diese Art der Körperkommunikation kompensiert die ausgefallene oder nur noch rudimentäre verbalsprachliche Handlungsabstimmung. Da also für die Rekonstruktion der Kommunikationsprozesse zwischen dem Menschen mit Demenz und den jeweiligen Beziehungspartner*innen oder Besucher*innen der Körperausdruck und hier insbesondere der Gesichts- und Augenausdruck zentral sind, ist es von essenzieller Bedeutung, diese möglichst genau erfassen und analysieren zu können. Diese möglichst genaue Aufzeichnung der gesamten Körperkommunikation macht eine Anonymisierung unmöglich. Würde man eine Anonymisierung jedoch durch *Verpixelung* vornehmen, dann würde für die Analyse das Wichtigste, nämlich die Körperkommunikation verloren gehen. Auch dies widerspricht einer Anonymisierung durch Verpixelung. Ähnliche Vorbehalte gelten natürlich auch, wenn man statt der Verpixelung die Transformation der gezeigten Person in *Comicfiguren* vornimmt. Denn bei dieser Transformation geht die individuelle Besonderheit, also die kommunikative Bedeutung einer spezifischen Ausdrucksbewegung gerade verloren. Wie bei jeder Comiczeichnung wird nur das Typische sichtbar gemacht, nicht jedoch das Besondere. Kurz: Videodaten von der Art, wie wir sie in dem DFG-Projekt über die Kommunikation mit Menschen mit Demenz erhoben haben, sind für eine spätere Nutzung durch Dritte nicht geeignet.

Literatur

Bryden, Christine (2011): Mein Tanz mit der Demenz. Bern: Huber.

Döttlinger, Beatrix (2017): Gestisch-kommunikatives Handeln als Bindeglied zwischen Sprache und Handeln bei Menschen mit Demenz: Beziehungs- und Interaktionsgestaltung. Weinheim und Basel: Beltz Juventa.

Frewer-Graumann, Susanne (2020): »Es ändert sich alles« – der Alltag mit Demenz aus der Perspektive der Angehörigen. In: Zeitschrift für Gerontologie und Geriatrie (1). https://doi.org/10.1007/s00391-019-01643-y.

Goffman, Erving (1967): Interaction Ritual: Essays on Face-to-Face Behavior. New York: Anchor Books.

Goffman, Erving (1982): Das Individuum im öffentlichen Austausch. Mikrostudien zur öffentlichen Ordnung. Frankfurt a. M.: Suhrkamp.

Hitzler, Ronald/Eisewicht, Paul (2016): Lebensweltanalytische Ethnographie-im Anschluss an Anne Honer. Weinheim und Basel: Beltz Juventa.

Honer, Anne (2011): Kleine Leiblichkeiten. Erkundungen in Lebenswelten. Wiesbaden: Springer.

Keller, Reiner/Knoblauch, Hubert/Reichertz, Jo (Hrsg.) (2013): Kommunikativer Konstruktivismus. Wiesbaden: Springer.

Keysers, Verena/Nebowsky, Anna-Eva (2020): Über Salutozentrismus und Fallstricke der Deutung ethnographischer Situationen. In: Hitzler, Ronald/Kreher, Simone/Schröer, Norbert/Poferl, Angelika (Hrsg.): Ethnographie der Situation. Essen: Oldib, S. 455–470.

Kruse, Andreas (2017): Lebensphase hohes Alter. Wiesbaden: Springer.

Mead, George H. (1973): Self, Language and the World. University of Texas.

Nebowsky, Anna-Eva (2024): Ethnografie bei Demenz – ein Perspektivwechsel. Person-Sein durch Paar-Sein. In Teupen, Sonja/Serbser-Koal, Jonathan/Laporte Uribe, Franziska/Dinand, Claudia/Roes, Martina (Hrsg.): Qualitative Forschung mit Demenz. Beiträge aus dem MethodenForum Witten 2021. Weinheim und Basel: Beltz Juventa, S. 137–156.

Nebowsky, Anna-Eva/Spiekermann, Nils (2021): Kommunikativer Tod durch Demenz? Wie die Beziehungsidentität Kommunikationsmacht erhält. In: Bidlo, Oliver/Keysers, Verena/Roslon, Michael/Schröer, Norbert (Hrsg.): Warum entfaltet kommunikatives Handeln auch ohne Gewalt und Herrschaft Macht? Weinheim und Basel: Beltz Juventa, S. 180–196.

Reichertz, Jo (2009): Kommunikationsmacht. Wiesbaden: VS Verlag für Sozialwissenschaften.

Reichertz, Jo (2020a): Grenzen der Kommunikation – Kommunikation an den Grenzen. Weilerswist: Velbrück.

Reichertz, Jo (2020b): »Auch wenn die geschriebenen Worte nicht von mir sind«. Depersonalisierung oder Wertschätzung? In: Pflege & Gesellschaft (4), S. 365–370.

Reichertz, Jo (2021a): »Man ist immer mittendrin«. Chancen und Risiken einer beobachtenden Teilnahme von Menschen mit der Diagnose ›Demenz‹. In: Nover, Sabine/Panke-Kochinke, Birgit (Hrsg.): Handbuch Qualitative Pflegeforschung. München: Nomos, S. 137–150.

Reichertz, Jo (2021b): Kommunikationsmacht, soziale Macht, Körpermacht. In: Bidlo, Oliver/Keysers, Verena/Roslon, Michael/Schröer, Norbert (Hrsg.): Facetten der Kommunikationsmacht. Weinheim und Basel: Beltz Juventa, S. 289–331.

Reichertz, Jo (2024): »Die Krankheit geht ihren Weg und nimmt sie mit«. Beobachtende Teilnahme von Paaren, von denen eine Person mit der Diagnose Demenz lebt. In: Teupen, Sonja/Serbser-Koal, Jonathan/Laporte Uribe, Franziska/Dinand, Claudia/Roes, Martina (Hrsg.): Qualitative Forschung mit Menschen mit Demenz. Weinheim und Basel: Beltz Juventa, S. 26–48.

Reichertz, Jo/Keysers, Verena/Nebowsky, Anna-Eva (2020): Einander ein Gerüst geben. In: Reichertz, Jo (Hrsg.): Grenzen der Kommunikation – Kommunikation an den Grenzen. Weilerswist: Velbrück, S. 219–237.

Roth, Michael/Reichertz, Jo (2020): »Coercive Care« or »Ur-wir [Great-we]«: Communication and Cooperation in Couples Where One Partner Has Been Diagnosed with Dementia (zusammen mit Michael Wolf Roth). In: Human Arenas (3), S. 552–574. https://doi.org/10.1007/s42087-020-00108-8.

Silverman, Marjorie (2016): Filming in the home: A reflexive Account of microethnographic data collection with family caregivers of older adults. In: Qualitative Social Work 15(4), S. 570–584.

Steffens, Gerd (2023): Ein Lebensversuch mit Demenz. Stuttgart: Kohlhammer.

Swaffer, Kate (2016): What the hell happened to my brain? Jessica Kingsley Publishers: London.
Taylor, Richard (2008): Alzheimer und Ich. Bern: Huber.
Welling, Karin (2014): »Sichtbar machen was sonst verborgen bleibt«. Annäherung an die videobasierte interpretative Mikroanalyse in der Interaktion zwischen Menschen mit Demenz und Pflegenden. In: Moritz, Christine (Hrsg.): Transkription von Video- und Filmdaten in der Qualitativen Sozialforschung. Multidisziplinäre Annäherung an einen komplexen Datentypus. Wiesbaden: Springer VS, S. 29–65.
Welling, Karin (2018): »Sich aneinander orientieren« – Feinfühligkeit und Engagement in der beziehungsorientierten Interaktion zwischen Menschen mit weit fortgeschrittener Demenz und Bezugspersonen. Eine mikroanalytische Interaktionsstudie. https://doi.org/10.26092/elib/28.

Konversationsanalyse transsituativer Kollektionen

Die Erforschung projektiver Gattungen in ihren temporalen Zusammenhängen

Jonas Kramer und Sarah Hitzler

Zusammenfassung: Der Beitrag zeigt, wie transsituative Prozesse im Rahmen einer longitudinalen konversations- und gattungsanalytischen Untersuchung von videoethnographischen Daten identifiziert, gesammelt und analysiert werden können. Exemplarisch wird erstens aufgezeigt, wie sich Videodaten so erheben lassen, dass transsituative Prozesse in alltäglichen Settings eingefangen werden können. Zweitens wird demonstriert, wie durch Verschlagwortung mit Atlas.ti das entstandene Datenkorpus so organisiert werden kann, dass Kollektionen von Trajektorien einzelner Handlungsvorhaben entstehen. Beides wird anhand einer Untersuchung illustriert, in der wir eine Abmachung zwischen Akteur*innen von ihrer Entstehung über ihre kommunikative Reproduktion und die Orientierung im Handeln bis hin zu ihrer Auflösung verfolgt haben. Analytisch spielt dabei die Orientierung an den Relevanzen der Teilnehmenden als Ressource zur Bildung und reflexiven Überprüfung der Kollektion eine zentrale Rolle. Das Vorgehen wird mit Blick auf etablierte Verfahren und weiteren Nutzungsszenarien diskutiert.

Schlüsselwörter: projektive Gattungen, longitudinale Konversationsanalyse, Transsituativität, QDA-Software, Kollektionsbildung

1. Einleitung

Die Umsetzung eines komplexen Handlungsvorhabens (Schütz 1971), z. B. die Renovierung eines Badezimmers, erstreckt sich oft über mehrere Interaktionen: Das Badezimmer muss geplant und die Renovierung vorbereitet werden. Es müssen Einkäufe getätigt werden. Schließlich muss die Renovierung selbst verrichtet werden. Die Teilschritte, die zum Vollzug des größeren Handlungsvorhabens durchgeführt werden, sind von kommunikativen Prozessen begleitet: Ziele müssen gesetzt, Aufgaben verteilt und Planänderungen kommuniziert werden. Die Umsetzung eines gemeinsamen Vorhabens ist somit ein transsituativer, kommunikativer Herstellungsprozess, und den Teilnehmenden muss es

über die Zeit hinweg gelingen, sich gegenseitig zu versichern, dass ihre Handlungen noch immer Teil eines intersubjektiven Plans sind. Ebenso wie sich dieses Problem den Teilnehmenden stellt, stellt es sich den Forschenden, die diese Prozesse untersuchen: Auch Forschenden muss es in der Analyse gelingen, aufzuzeigen, dass und wie verschiedene Situationen miteinander verbunden sind, wie sich Pläne und Projekte verändern, und schließlich wie kommunikativ erzeugte Erwartungen für die Zukunft zu handlungsanleitenden Ressourcen in der neuen Gegenwart werden – oder aber vergessen und ignoriert werden. Die Bildung von Kollektionen, welche die für einen sozialen Prozess relevanten Situationen darstellen, stellt eine zentrale Aufgabe in diesem Zusammenhang dar. Ziel des folgenden Beitrags ist es, ein methodisches Vorgehen zur Bildung von Kollektionen audiovisueller Daten zu beschreiben, welche Trajektorien (Strauss et al. 1985) komplexer Handlungsvorhaben abbilden. Das Vorgehen zeichnet sich dadurch aus, zunächst sehr offen eine Vielzahl möglicher Datenstücke in einer »Proto-Kollektion« zu sammeln und diese anschließend methodisch und analytisch abzusichern, und um weitere relevante Stellen zu erweitern. Hierbei spielt insbesondere die Teilnehmendenperspektive eine wichtige Rolle, die berücksichtigt werden muss, um eigene Annahmen über die Vollständigkeit der Kollektion mit der sozialen Wirklichkeit, wie sie von den Teilnehmenden hergestellt wurde, abzugleichen. Das Vorgehen wurde ursprünglich entwickelt, um projektive Gattungen und Gattungsreihen (Ayaß 2021) zu erforschen, eignet sich unserer Auffassung nach aber auch für die allgemeine Untersuchung longitudinaler Prozesse anhand von Audio- oder Videodaten. Die Daten, anhand derer diese Methode der Kollektionsbildung entwickelt wurde, wurden im Rahmen des Forschungsprojekts »Planning-in-Action: Die kommunikative Verfertigung von Zukunft in projektiven Gattungen« (2021–2024; Leitung Prof. Dr. Ruth Ayaß, Universität Bielefeld) generiert. Das Projekt wird finanziert von der Deutschen Forschungsgemeinschaft. Es widmet sich den Formen des Sprechens über künftige Handlungen in alltäglichen Kontexten, also den kommunikativen Formen, mit denen Interagierende im Alltag ihre Handlungen vorbereiten, antizipieren und planen. In diesem Rahmen wurden über mehrere Wochen hinweg über 200 Stunden videographische Daten in zwei Feldern erhoben: Einer dreiköpfigen Familie und einem solidarischen Landwirtschaftsbetrieb.

Wir möchten im Folgenden anhand der idealisierten Darstellung unseres Forschungsprozesses aufzeigen, wie wir methodisch Kollektionen von Trajektorien erzeugen, mithilfe derer wir die sich wandelnden Wissensbestände und Pläne der Teilnehmenden untersuchen. Als erstes werden wir das Ergebnis eines solchen Forschungsprozesses aufzeigen, welches im weiteren Verlauf als wiederkehrendes Beispiel dienen soll. Anschließend werden wir etablierte Methoden der Erforschung transsituativer Prozesse vorstellen und aufzeigen, worin sich unser Vorgehen von ihnen unterscheidet. Im vierten und fünften

Abschnitt werden wir beschreiben, wie wir unser Datenkorpus stattdessen erhoben und aufbereitet haben. Dabei werden wir zunächst auf den Erhebungskontext, insbesondere auf die Entscheidungen in der Aufnahmesituation eingehen, welche zu dem großen Umfang unseres Datenkorpus führten. Anschließend klären wir, wie wir im Kontext unserer Analyse das Datenkorpus aufgebrochen, iterativ inventarisiert und Kollektionen in ihm gebildet haben, welche die Grundlage für unsere Gattungs- und longitudinalen Konversationsanalysen wurden. Abschließend möchten wir andere Anwendungsfälle und Nutzungsszenarien unseres Vorgehens diskutieren.

2. Die Ergebnisse einer transsituationalen Konversationsanalyse: Zur Trajektorie objektifizierten Wissens

Zunächst werden wir einige Ergebnisse des zu beschreibenden Forschungsprozesses vorwegnehmen, welche kürzlich unter dem Titel »Trajectory of an Agreement. Tracing Objectivated Knowledge Across a Series of Mundane Encounters« (Hitzler/Kramer 2024) erschienen sind. Im restlichen Beitrag werden wir immer wieder exemplarisch auf die Daten und den ihm zugrundeliegenden Forschungsprozess eingehen.

Im Zuge unserer longitudinalen Forschung zu Gattungsreihen (op. cit.) erkannten wir die Relevanz objektifizierter Wissensbestände (Luckmann 1983, S. 63), die von den Teilnehmenden genutzt werden, um vorangegangene Phänomene, z. B. Entscheidungen, zu referenzieren. Im Artikel verfolgen wir einen solchen Wissensbestand über mehrere Situationen und zeigen, wie er objektifiziert wurde, wie sich im Handeln der Teilnehmenden darauf bezogen wird und wie er schließlich, als er nicht mehr handlungsanleitend ist, in den Hintergrund rückt. Bei dem objektifizierten Wissensbestand handelt es sich um eine Abmachung zwischen der Familienmutter und einem befreundeten Handwerker. Die Abmachung ist Teil eines größeren Handlungsvorhabens (der Renovierung eines Bades) und lautet: Frank, ein befreundeter Handwerker, wird am Freitag kommen, um das Bad zu renovieren. Kerstin, die Mutter, wird ihm zwar nicht helfen, ihn aber dafür versorgen.

Dieser Abmachung folgten wir über mehrere Situationen hinweg: Von ihrer kommunikativen Hervorbringung, über ihre kommunikative Reproduktion, Handlungen, die von dieser Abmachung bestimmt wurden, Krisensituationen, in denen sie bedroht wurde, bis hin zu Situationen, in denen sie in den Hintergrund rückte. Hieran arbeiteten wir u. a. heraus, dass das Verhältnis von Plan und Handeln rekursiv wie reflexiv ist (Hitzler/Kramer 2024, S. 24). Reflexiv ist das Verhältnis insofern, als dass Handeln und Plan wechselseitig und jeweils durch das andere auf sich selbst verweisen und sich so gegenseitig *accountable*

(Garfinkel 1967) machen. Geplante Handlungen bekommen ihre spezifische Bedeutung durch den Plan, zu dem sie gehören. Und Pläne werden dadurch als Pläne erkennbar, dass sie sich durch Handlungen realisieren. Rekursiv sind sie wiederum, weil Pläne sich über Zeit hinweg realisieren und sich dabei der sich verändernden Welt anpassen. Durch Planung und das geplante Handeln über die Zeit hinweg wird jedoch sichergestellt, dass der Plan trotz Veränderungen identisch mit sich selbst bleibt.

Datengrundlage dieser Ergebnisse war ein umfangreiches Datenkorpus, welches wir derart inventarisierten, dass es uns gelang, ein Handlungsvorhaben in dessen Trajektorie über Situationen hinweg zu verfolgen. Darüber hinaus gelang es uns in dieser Kollektion, sicher zu stellen, dass sie Situationsverläufe abbildete, die in der Alltagswirklichkeit der Teilnehmenden als zum selben Projekt zugehörig hergestellt wurden. Die Bildung der hierzu nötigen Kollektion erforderte ein sorgfältiges methodisches Vorgehen. Im Folgenden möchten wir zwei etablierte Verfahren zur Analyse längsschnittlicher Interaktionsdaten vorstellen, allerdings auch aufzeigen, warum sie sich in unserem spezifischen Fall als nicht angemessen für unseren Gegenstand und unser Interesse an ihm erwiesen (siehe zum Gütekriterium der Gegenstandsangemessenheit Strübing et al. 2018) und worin sie sich von unserem eigenen Vorgehen unterscheiden.

3. Etablierte Analyseverfahren längsschnittlicher Interaktionsdaten

Transsequentielle Analyse: Entwickelt von Thomas Scheffer, zielt die Transsequentielle Analyse (TSA) auf die Untersuchung formativer Objekte, vor allem auf ihre Hervorbringungs- und Verwertungspraktiken. Formative Objekte sind als materielle Gegenstände definiert, die formbar sind, zu formen sind und formierend wirken (Scheffer 2013, S. 90), etwa Akten, Blaupausen, oder Gutachten. Die TSA kombiniert Konversationsanalyse, Ethnografie und Diskursanalyse, um praktische Interaktionszusammenhänge mikrosoziologisch zu untersuchen (Scheffer 2013, S. 110). Trans-sequentiell ist die Analyse insofern, als dass sie dezidiert darauf zielt, über die Analyse einzelner Sequenzen hinaus zu gehen. Die TSA wurde anhand von hochformalen, organisationalen Kontexten entwickelt und auch neuere Untersuchungen teilen diesen Gegenstandsbereich (siehe den kürzlich erschienenen Sammelband von Kolanoski et al. 2023). Ebenso fokussiert die TSA materielle formative Objekte, um die Verkettung einzelner Situationen zu untersuchen, auch wenn es hier jüngst Bemühungen gibt, nicht-materielle Objekte in den Blick zu nehmen (Löffler 2022). Trotzdem werden diese nur innerhalb institutioneller Produktionszusammenhänge betrachtet. Aus unserer Alltagserfahrung wissen wir jedoch,

dass es uns auch gelingt, Situationen zu verketten, wenn uns keine formativen Objekte vorliegen. Die Herstellung und Verwendung formativer Objekte sind also nicht die einzige Lösung für das Problem der Herstellung von Kontinuität zwischen Situationen.

Scheffer kritisiert mit seinem Verständnis von Trans-Sequenzialität den Situationalismus, welcher der ethnomethodologischen Konversationsanalyse auch von anderer Seite oftmals vorgeworfen wird (siehe z. B. Hardesty 1982; Gibson 2021). In den letzten Jahrzehnten hat sich der Phänomenbereich der Konversationsanalyse jedoch immer weiter vergrößert: Die Untersuchung von Multimodalität oder multiplen Temporalitäten (Mondada 2007; Au-Yeung/Fitzgerald 2023), wie auch der Wandel von Praktiken über die Zeit hinweg in Form von longitudinaler Konversationsanalyse wird vermehrt in den Blick genommen.

Longitudinale Konversationsanalyse: Erste längsschnittliche Arbeiten der Konversationsanalyse gehen auf Woottons Untersuchungen frühkindlicher Sprachentwicklung (Wootton 1997) und die vergleichende Studie von Clayman und anderen zu journalistischen Fragen in präsidialen Pressekonferenzen (Clayman et al. 2007) zurück. Mit einem Sammelband (Wagner/Pekarek Doehler/González-Martínez 2018) wurden erstmals verschiedene längsschnittliche Untersuchungen gebündelt und unter dem Begriff »longitudinale Konversationsanalyse« diskutiert. Es werden drei verschiedene Ansätze unter dem Sammelbegriff longitudinaler Konversationsanalyse subsumiert: Entwicklungsstudien, welche die längsschnittliche Betrachtung der Veränderung von Sprachkompetenz untersuchen, Studien soziohistorischen Wandels, welche die Unterschiede spezifischer Sprachformen zu bestimmten historischen Zeitpunkten herausarbeiten und die Untersuchung von Interaktionsgeschichten, durch welche der Wandel sprachlicher Praktiken in Praxisgemeinschaften (»Communities of Practice«) untersucht wird (Deppermann/Pekarek Doehler 2021, S. 128 ff.). Der für unsere eigene Forschung relevanteste Bereich longitudinaler Konversationsanalyse ist die Untersuchung von Interaktionsgeschichte (siehe z. B. Mondada 2018; Schmidt 2018). Hierbei werden spezifische kommunikative Praktiken in ihren jeweiligen Ausprägungen über die Zeit hinweg betrachtet und in ihrem Wandel analysiert.

Unser Ansatz unterscheidet sich jedoch in einigen grundlegenden Aspekten von den Untersuchungen zur Interaktionsgeschichte. Dies liegt vor allem an unserer soziologischen Fragestellung, die nicht den *Wandel sprachlicher Praktiken* zentral setzt, sondern darauf fokussiert wie diese Praktiken *Wandel in der Lebenswelt* der Teilnehmer*innen von Gesellschaft erzeugen: Wir untersuchen, wie Gesellschaftsmitglieder ihre Umwelt und ihr Wissen von ihr verändern, also wie Pläne entstehen, verändert werden und sich im Handeln an ihnen orientiert wird (Hitzler/Kramer 2024, S. 4). Anders als in den Studien zur Interaktionsgeschichte, die zu erklären versuchen, warum sich im Laufe

einer Interaktionsgeschichte bestimmte Sprechweisen ändern, betrachten wir, wie in Planungsgesprächen Erwartungswissen erzeugt wird – und wie dieses Erwartungswissen wiederum zu späteren Zeitpunkten den Teilnehmenden zur Lösung anderer Probleme als Ressource zur Verfügung steht. Wir folgen in diesem Sinne einer streng emischen, temporalen Perspektive, die sequenziell vorgeht und dabei sowohl vorwärts- als auch rückwärtsgewandt argumentiert, allerdings auch transsituativ über den unmittelbar nächsten Turn hinaus. Aus diesem Vorgehen speisen sich auch die Ergebnisse unserer längsschnittlichen Untersuchungen, deren Fundament in der Erzeugung und Aufbereitung eines longitudinalen Datenkorpus lag. Daher soll nun geklärt werden, wie wir unsere Daten generiert und warum wir hierzu insgesamt ca. 200 Stunden an Videodaten erzeugt haben.

4. Unser Erhebungskontext: Trajektorien von Handlungsvorhaben und Typisierungen von Gattungen im Alltag erforschen

Während unserer videographischen Feldarbeit hatten wir zwei Ziele: Erstens, möglichst viele verschiedene Formen projektiver Kommunikation aufzunehmen. Zweitens, die Realisierung eines oder mehrerer Handlungsvorhaben möglichst vollständig zu registrieren. Die vorgestellten Ergebnisse basieren auf Daten, die im ersten Feld, der dreiköpfigen Familie, entstanden sind.

Da wir versuchten, einen Einblick in den Alltagsrhythmus der Familie zu erhalten und auch, um einige längere Vorhaben über ihren Verlauf aufzunehmen, entschieden wir uns für vergleichsweise lange Feldaufenthalte, die wir am Stück durchführten. Einer der Gründe dafür war, dass wir uns aus unserer gattungsanalytischen Perspektive auch für die Sozialstruktur der Felder interessierten. Diese gibt Aufschluss über die Außenstruktur (Luckmann 1986, S. 204) der Gattungen, und ist auch relevant für die Planungsaktivitäten als solche. Entsprechend wichtig war es, ein gutes Verständnis über die Alltagsstrukturen der untersuchten Felder zu erlangen, innerhalb derer sich die untersuchten Handlungsvorhaben und deren Kommunikation realisieren. Hieraus ergab sich, dass nicht aus einem rein konversationsanalytischen Interesse heraus gefilmt wurde, sondern z. B. auch ethnographische Gespräche zur Vergangenheit laufender Vorhaben geführt wurden (zum Vergleich von konversationsanalytischem und ethnographischem Filmen siehe Heath/Hindmarsh/Luff 2010; vgl. Knoblauch/Wilke i. d. B.). Auch die Intimität des familialen Kontextes trug dazu bei, dass der Videograph oft selbst in Gespräche verwickelt wurde und teilnehmend beobachtete bzw. filmte.

Ein weiterer Grund dafür, am Stück zu filmen, war, dass sich früh zeigte, dass über Zukunft jederzeit und überall gesprochen werden kann. Oftmals aus dem Nichts, z. B. etwa nach einer guten Stunde schweigsamen Fernsehens, fragte ein Familienmitglied ein anderes, was morgen zu tun sei, wie ein Familienmitglied zur Arbeit komme oder ob noch etwas eingekauft werden müsse. Ein fokussiertes Filmen war daher nicht möglich. In den vielen Stunden Material wird dementsprechend nicht immer über Zukunft gesprochen, es wird z. B. auch vom Tag berichtet, ferngesehen, gekocht oder die Hunde gestreichelt – aber die Verfertigung eben dieses Alltags bildet den Kontext alltäglicher Planung. Auch deshalb war es für uns nützlich, möglichst viel Alltag der Familie zu dokumentieren.

Die zwei Feldphasen (eine erste im Winter, eine zweite im Sommer) erstreckten sich auf insgesamt drei Wochen. Der Videograph war wochentags jeden Nachmittag anwesend, sobald das erste Familienmitglied von der Arbeit bzw. der Schule nach Hause kam, sowie jeweils samstags. An den Sonntagen hat die Familie zum Teil selbst gefilmt, hatte hier aber auch einen Tag für sich, um unbeobachtet zu bleiben. Beim Filmen wechselten sich zwei Kameras ab: Während die eine lief, konnte die andere gewartet werden. Der Videograph begleitete die Familie auch auf Spaziergängen, bei Autofahrten, zu befreundeten Personen nach Hause und beim Einkauf. Wo es datenschutzrechtlich z. B. aufgrund fehlender Einwilligungen unbeteiligt passierender Personen nicht möglich war zu filmen, wurden Beobachtungen durchgeführt.

Da nur ein Videograph anwesend war und dieser oft mit nur einer Kamera filmte, mussten immer wieder Entscheidungen getroffen werden, welche Aktivitäten gefilmt wurden, und welche Informationen ethnographisch zusätzlich eingeholt werden mussten (wodurch der familiale Alltag durch kurze oder längere Gespräche unterbrochen wurde). Hierbei folgte der Videograph den Fragestellungen und Relevanzen des Projektes, wie das folgende Beispiel illustriert.

Eine dieser Relevanzsetzungen war die Planung und Umsetzung eines komplexen Handlungsvorhabens, die genannte Renovierung eines Badezimmers. Die Badrenovierung erstreckte sich während der Erhebungsphase über drei Tage (Mittwoch/Donnerstag/Freitag) und die Planung erfolgte sowohl *face-to-face*, wie auch transmedial über Messaging-Apps.

Die Vorgeschichte des Bades wurde über ein ethnographisches Gespräch in Erfahrung gebracht und die Planungs- und Arbeitsschritte gefilmt. Schon während der Erhebung wurden auf verschiedenen Ebenen die Erfahrungen der Forschenden analytisch bearbeitet. Dies geschah u. a. dadurch, dass während der Aufnahme strukturierende Notizen und Beobachtungen angefertigt wurden (siehe Abb. 1). Solche Beobachtungen und Notizen bildeten auch die Grundlage für die Auswahl erster Fokussequenzen im Projektverlauf.

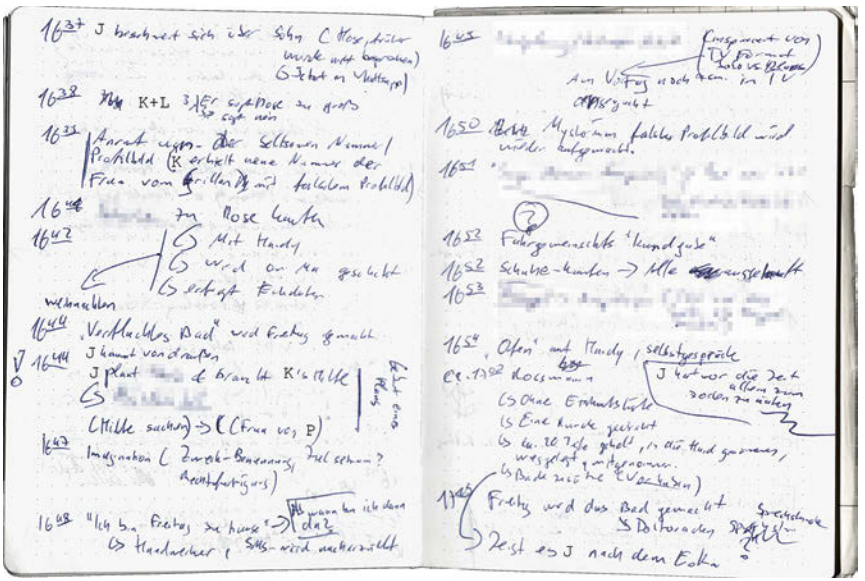


Abbildung 1: Auszug aus dem Feldtagebuch[1]

Insgesamt entstand auf diese Weise im ersten Feld ein Datenkorpus, das aus 140 Stunden Videomaterial, 30 Stunden Audiomaterial, ethnographischen Notizen, mehreren Messenger-Verläufen, einigen Fotografien sowie weiteren Artefakten bestand. So sehr dieses Datenkorpus als Antwort auf die verschiedenen Bedarfe verstanden werden kann, die sich aus unseren Fragestellungen, Methoden, datenschutzrechtlichen Gegebenheiten und den Spezifika dieses Feldes und Gegenstandes ergaben, so stellte es uns in der Auseinandersetzung mit dem Korpus vor forschungspragmatische Probleme: Im Rahmen der Projektlaufzeit wäre es uns weder möglich, das vollständige Korpus zu transkribieren und feinanalytisch zu betrachten, noch wäre das überhaupt zielführend gewesen, da nicht jedes Datenstück für unser Projekt den gleichen Stellenwert einnahm. Wir mussten also eine informierte Fallauswahl treffen. Doch auch hier stellte uns der Umfang des Datenkorpus vor Probleme, denn dessen Größe sorgte dafür, dass wir Datenstücke nicht ohne Weiteres wiederfinden konnten. Kurzum, wir hatten einige Probleme, die wir nun durch Datenaufbereitung und Analyse lösen mussten.

1 Personenbezogene und sensible Daten wurden unkenntlich gemacht, Namenskürzel wurden ausgetauscht.

5. Analysekontext: Kollektionsbildung als Bestandteil iterativer Forschungsprozesse

Um spezifische Fokussequenzen auszuwählen und feinanalytisch zu betrachten, mussten wir zunächst die umfangreichen videographischen Datenmengen aufbrechen und aufbereiten. Hierfür haben wir das Rohmaterial inventarisiert, verschlagwortet und auf Grundlage dieses Prozesses Kollektionen gebildet. Auf diese Weise verfügten wir noch immer über das unveränderte audiovisuelle Material und konnten es zu späteren Zeitpunkten unseres iterativen Vorgehens heranziehen.

Wir inventarisierten das Datenkorpus mithilfe von Atlas.ti. Hierbei verschlagworteten wir die Videodaten mit dem, was diese QDA-Software »Codes« nennt. In der Art, wie wir diese Funktion verwendeten, handelte es sich allerdings nicht um einen Kodierungsprozess im Sinne der Grounded Theory, sondern um »Tagging« bzw. um »Verschlagwortung« (Friese 2019, S. 284).

Die Vorarbeit hierzu begann bereits im Feld, denn schon dort hatte der Videograph Projekte und Gesprächsformen identifiziert, welche sich sowohl in seinem ethnographischen Wissen ablagerten wie auch in seinen Notizen materialisierten (siehe Abb. 1, z. B. 16:52 »Kundgabe« und 16:44 »Verfluchtes Bad«). Dieses Wissen schrieb sich auch in den Verschlagwortungsprozess ein.

Abbildung 2 zeigt exemplarisch, wie fünf Minuten eines bereits verschlagworteten Dokuments aussehen: Es wird ein Teil des horizontalen Zeitstrahls markiert (hier die ersten fünf Minuten des Videos), welcher in der Mitte des Bildschirms dann als vertikaler Zeitstrahl angezeigt wird. Auf diesem vertikalen Zeitstrahl werden wiederum Zitate gebildet – dies sind die vertikalen Balken auf der rechten Bildschirmseite, die wiederum einen eigenen Zeitraum abbilden. Wie an dem Bild zu sehen ist, gehen dabei einige Zitate zeitlich über den hier dargestellten Bereich hinaus (etwa das Schlagwort »F/K/L«), erstrecken sich also über einen längeren Zeitraum. An diese Zitate werden verschiedene Tags angeheftet (zu erkennen als die verschiedenfarbigen horizontalen Balken im rechten Bildschirmbereich). Hierbei verschlagworteten wir u. a. die Teilnehmendenkonstellation (gelb, z. B. »F/K/L«), Gesprächsthemen (rot, z. B. »Das Verfluchte Bad«), Praktiken (grün, z. B. »berichten«), aber auch andere Phänomene, wie z. B. die Verwendung von Objekten, die zu Planungszwecken hinzugezogen wurden.

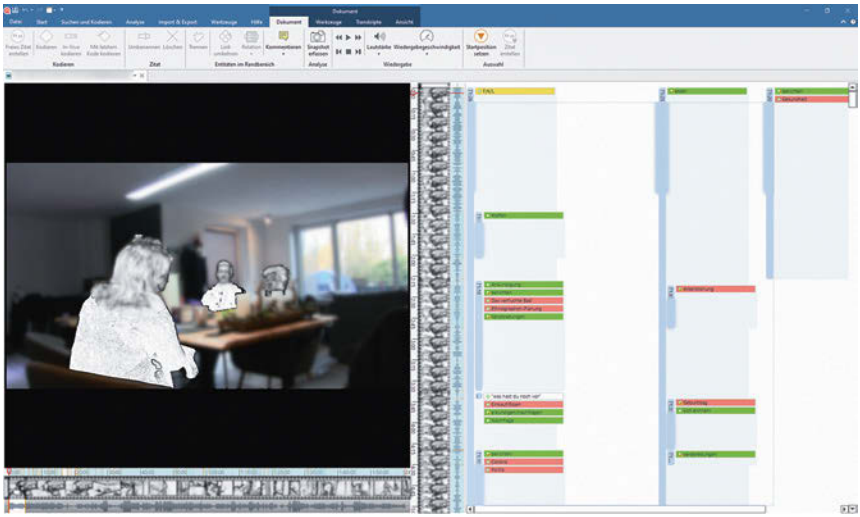


Abbildung 2: Fünf Minuten mit Atlas.ti verschlagwortetes Videomaterial (anonymisiert)

Sind mehrere Dokumente verschlagwortet, kann das Korpus gezielt durchsucht werden.[2] Trajektorien und Handlungsverläufe lassen sich beispielsweise identifizieren, indem nach Stellen gesucht wird, die mit ein und demselben Thema verschlagwortet sind. Diese Stellen können gebündelt im Atlas.ti-Projekt betrachtet werden und bei Bedarf weiterbearbeitet werden (feiner verschlagwortet, transkribiert, analysiert etc.). Dies taten wir u. a. mit einem Schlagwort namens »Das verfluchte Bad«. Dabei handelte es sich um einen Begriff der Teilnehmenden, der so bereits im Feldtagebuch notiert worden war (siehe Abb. 1, 16:44) und bei der Verschlagwortung entsprechend zur Benennung des Themas verwendet wurde.

Für die Videodaten des Mittwochs und des Donnerstags, an welchen die Badrenovierung vorbereitet wurde, wurde das Schlagwort »Das verfluchte Bad« insgesamt 15-mal vergeben. Dabei umfassen die Situationen, die so verschlagwortet wurden, konkrete Gespräche über das Bad, aber auch Vorarbeiten, wie etwa das Aufräumen des Bades, oder aber das Bereitstellen von Renovierungsmaterial. Die erste »Proto-Kollektion« des »verfluchten Bades« schränkte so das Material von ca. 15 Stunden Audio- und Videomaterial, das an beiden Tagen generiert wurde, auf ca. 30 Minuten ein. Problematischer war der Freitag, an welchem das Bad tatsächlich renoviert wurde. An diesem Tag wurden die Renovierungsarbeiten, die

2 Für einen Vortrag, den wir auf dem DGS-Kongress 2022 in Bielefeld hielten, suchten wir beispielsweise nach einem Datenstück, in dem alle drei Familienmitglieder über eine gemeinsame Zukunft sprachen (Kramer/Singh 2023). Unser Ziel war es, die Herstellung von Familie durch und bei Planung zu untersuchen. Dementsprechend suchten wir nach Datenstücken, in denen die Teilnehmendenkonstellation »J/K/L« bestand, sowie Themen, die alle Familienmitglieder betrafen (in dem Fall des Vortrags: Urlaub). Mit diesen Suchkriterien reduzierte sich das 140-stündige Datenmaterial auf vier Ausschnitte.

sich von morgens bis abends erstreckten, den ganzen Tag über gefilmt. Dementsprechend wurden auch alle Renovierungsarbeiten mit »Das verfluchte Bad« gekennzeichnet, sodass das Schlagwort uns kaum reduzierend unterstützen konnte. Trotzdem begannen wir schon einmal mit der Analyse der ersten beiden Tage. Die für unsere Fragestellung besonders relevante Vorbereitung und Planung des Vorhabens konnte so parallel zur weiteren Inventarisierung in den Blick genommen werden. Die verschlagworteten Datenstücke wurden transkribiert und zusammen mit den Videos detaillierter analytischer Betrachtungen unterzogen – gemeinsam in Datensitzungen wie auch in alleiniger Arbeit.[3] Sowohl kommunikative Formen, Teilprojekte, wie auch die Verteilung verschiedener Aufgaben ließen sich so identifizieren.

Auf Basis unserer Erkenntnisse ergab sich eine zweite Schleife durch das Datenmaterial, insbesondere durch den Freitag: Beispielsweise erschien uns das Teilprojekt des »Wasser An- und Abstellens« klein genug, um es mithilfe dieses Schlagworts festhalten zu können. Dieses umfasste die Planung, Kundgabe und Tätigkeit jeder Situation, in der das Wasser an- oder abgestellt wurde. Dieses Phänomen war mit einem anderen verbunden, nämlich der Versorgung des Handwerkers u. a. mit Kaffee: War das Wasser aus, konnte u. a. kein Kaffee gekocht werden. Auch deshalb informierte der Handwerker die Familienmitglieder, wenn er das Wasser abstellte. Aus dieser Beobachtung ergab sich ein gesteigertes Interesse daran, wie im Rahmen der Badrenovierung Aufgaben allgemein verteilt und die Arbeitsteilung praktisch umgesetzt wurde. In einer dritten Schleife betrachteten wir das Datenkorpus aus dieser Perspektive und markierten mehrere Stellen mit dem Schlagwort »Arbeitsteilung«. Wir identifizierten Stellen, in denen Aufgaben verteilt oder die Arbeitsteilung im Handeln ersichtlich wurde. Es entstand eine umfassende und unterschiedlich abstrakte Kollektion, in welcher das Phänomen der Arbeitsteilung mal konkreter oder expliziter, mal indirekter verhandelt wurde. Eine solche diverse Proto-Kollektion ist vorteilhaft für einen theoriebildenden Prozess, da das Phänomen in vielen Facetten und aus vielen Blickwinkeln betrachtet werden kann, anstatt durch eine im Vorhinein theoretisch festgelegte Definition den Phänomenbereich einzuschränken.

Durch den Prozess dieser Kollektionsbildung entdeckten wir die Vereinbarung zwischen Gastgeberin und Handwerker wieder, die wir bereits in der Feinanalyse der Transkripte vom ersten Tag identifiziert hatten und von der wir nun erstmals sahen, wie sie transsituativ immer wieder aufschien: »Ich kann dir nicht helfen, aber ich kann dich versorgen«. Ohne erneut zu verschlagworten, durchsuchten wir die Kollektion »Arbeitsteilung« sowie alle bereits transkribierten

3 Im Rahmen dieses Beitrags ist es uns leider nicht möglich, das feinanalytische Vorgehen der ethnomethodologischen Konversationsanalyse zu beschreiben. Unser Fokus hier liegt auf der Arbeit der Inventarisierung und Kollektionsbildung, über die wir unsere Daten für die Feinanalyse auswählen.

Stellen vom Mittwoch und Donnerstag und begannen die Transkription und Feinanalyse aller Stellen, in denen die Versorgung des Handwerkers verhandelt oder umgesetzt wurde. Dieses Phänomen war klein genug, um seine transsituative Ausprägung in Gänze nachvollziehen und analysieren zu können und gleichzeitig explizit und relevant genug, um zu zeigen, dass und wie sich die Teilnehmenden an der Vereinbarung orientierten.

Transkript 1[4]

```
1   K:  <<kauend>dass ja gut wenn clara kommt (.)
2       ich wollte heut nachmittag nämlich au:ch ä:hm>
3       (schluckt)) ((räuspert sich)) croissants (.) frische machen
4       (.) mit nutella [gefüllt.      ]
5   F:                  [<<:-)>ah cool>]
6       das_s cool.
7   K:  [ich hab ja gesagt ich verpflege dich;]
8   F:  [((lacht))                            ]
9       ((lacht))
10  K:  K: ich halte mein wort;
11      (2)
12  K:  ich kann dir nicht helfen aber ich kann dich
13      ernäh[ren]
14  F:        [ge ]nau
```

In diesem Transkriptauszug, welcher Teil der Kollektion zum verfluchten Bad war, ist z. B. zu sehen, wie Kerstin, die morgens am Frühstückstisch mit dem Handwerker sitzt (siehe Abb. 2), explizit die Vereinbarung aufruft. Dabei erfüllt sie ihren Teil der Abmachung bereits in der Situation selbst, in welcher sie gemeinsam frühstücken. Gleichzeitig erinnert sie ihn retrospektiv »Ich hab ja gesagt, ich verpflege dich« (Z. 7) und projiziert diese Vereinbarung auch in die Zukunft, indem sie ein Vorhaben benennt »Ich wollte heute Nachmittag auch

4 Bei diesem Transkript handelt es sich um ein Verbaltranskript nach GAT 2 (Selting et al. 2009; vgl. Meyer/Meier zu Verl i. d. B.), in dem nur wenige multimodale Phänomene dargestellt sind. In unseren Analysen – seien diese konversationsanalytisch oder eher ethnographisch geprägt – arbeiten wir jedoch eng an den Videodaten, betrachten Transkripte immer in ihrem audiovisuellen Kontext und ziehen diese z. B. als semiotische Felder hinzu Goodwin 2000, S. 1494). Aus unserem gattungsanalytischen Interesse, welches die kommunikativen Planungsprozesse zentral setzt, ist jedoch nicht von der Hand zu weisen, dass ein Großteil dieser Planung im untersuchten Alltag verbal von statten geht, sodass eine Fokussierung auf verbale Phänomene naheliegt. Ein Beispiel für die analytische, ethnographische Arbeit an den Videos ist in Fußnote 5 zu sehen. Für das hier dargestellte Argument sind jedoch nur die inhaltlich, verbalen Bezüge in ihrer sequenziellen Einbettung nötig, weshalb weitere Phänomene ausgeklammert sind.

ähm Croissants frische machen, mit Nutella gefüllt« (Z. 2–4) (für eine detaillierte Analyse des Datenstücks s. Hitzler/Kramer 2024, S. 14f.).

Datenstücke wie dieses, in welchen explizit von den Teilnehmenden Bezug auf vergangene Situationen genommen wird, sind für die Bildung und Prüfung der Kollektion sehr wertvoll: Fänden wir in unserer Proto-Kollektion keinen Moment, in dem Kerstin verspricht, Frank zu versorgen, wäre das ein starker Hinweis auf eine fehlende Interaktion in unserer Kollektion. Dafür müssen die Bezüge auch gar nicht so konkret sein, wie es hier der Fall ist: Aus der Perspektive der Ethnomethodologie stellt sich der situativ relevante Kontext interaktiv her (Garfinkel 1967, S. 7). Indem wir die Analyse der sozialen Wirklichkeit, wie sie von den Teilnehmenden interaktiv hervorgebracht wird – z. B. durch ihr *recipient design* (Sacks/Schegloff/Jefferson 1974) wiederum methodisch nachvollziehen, können wir identifizieren, ob eine bestimmte Information als geteilt hergestellt wurde oder als Neuigkeit. Darauffolgend können wir uns fragen, ob das als geteilt behandelte Wissen im Erhebungszeitraum hergestellt wurde (oder ob es z. B. gesellschaftlich geteiltes Allgemeinwissen ist oder bereits vor unserem Aufnahmezeitraum hergestellt wurde). Sollte die Herstellung eines solchen Wissensbestands nicht Teil der Proto-Kollektion sein, ist sie noch unvollständig und es kann nach der entsprechenden Stelle gezielt gesucht werden.[5]

Bis auf einen kurzen Moment, für den wir zeitlich ziemlich genau bestimmen können, wann und was in ihm passiert sein musste (siehe Fußnote 5), und der Vorgeschichte des Bades, welche durch ein ethnographisches Gespräch rekonstruiert wurde, konnten wir so eine in sich selbst kohärente Kollektion der Trajektorie erzeugen. Diese ist insofern empirisch gesättigt, als dass wir keine weiteren Hinweise auf unbekannte Handlungen und Situationen in ihr finden können.

Tatsächlich waren unser Vorgehen und unsere Ergebnisse so eng miteinander verzahnt, dass sie sich gegenseitig bedingten. In der gleichen Weise, wie wir versuchten, anhand der sich wandelnden oftmals objektifizierten Wissensbestände herauszuarbeiten, ob unsere Kollektion vollständig war, zeigten sich auch die Teilnehmenden gegenseitig an, dass ihr Handeln Teil eines transsituativen Handlungsvorhabens war, indem sie eben diese objektivierten Wissensbestände aufriefen, teilten und als Ressource der Verständnissicherung nutzen.

5 Ähnlich verhält es sich auch auf der materiellen Ebene: An einer Stelle nimmt sich der Handwerker ein Getränk von einem Tisch, welches anscheinend dort für ihn bereitgestellt wurde. Aus unserer Kollektion ging aber nicht hervor, wie es dort positioniert wurde. Die Sichtung der Videodaten ergab, dass dies in einer kurzen Aufnahmepause geschehen sein musste. In dieser Zeit wurde die Kamera im Wohnzimmer ausgestellt, um einem neu erschienenen Gast der Familie die Einwilligungserklärung zu erläutern und unterzeichnen zu lassen. In diesem etwa zehnminütigen Aufnahmestopp muss Kerstin den Tisch gedeckt haben, wobei der Kaffee dort für den Handwerker bereitgestellt wurde.

6. Ausblick: Weitere Nutzungsszenarien und Anwendungsfälle

Das hier beschriebene Vorgehen ist zwar langwierig, aber aufgrund der häufigen iterativen Schleifen dafür auch gründlich und zielführend. Es gäbe zwei Alternativen zu unserem Vorgehen. Zunächst ist die vollständige Transkription des Datenkorpus zu nennen. Textdokumente lassen sich deutlich leichter durchsuchen als Videodaten und bei Bedarf auch verschlagworten. Dafür wird allerdings auch viel Zeit oder Geld in die Verschriftlichung der Daten investiert und bei großen Datenmengen wird so mehr transkribiert, als später feinanalytisch betrachtet werden kann. Die andere Alternative ist, die Inventarisierung nur grob durchzuführen und die gebildeten Kollektionen nicht auf ihre Teilnehmendenrelevanz oder Vollständigkeit zu prüfen. Ein solches Vorgehen läuft Gefahr, die eigenen theoretischen Annahmen dem Datenmaterial aufzustülpen und entscheidende Datenstücke zu übersehen. Durch eine Triangulation von Verschlagwortung und Feinanalyse kann es jedoch gelingen, den Inventarisierungsaufwand auf ein Minimum zu beschränken und dennoch Aussagen über die Güte der Fallauswahl zu treffen. Wir halten ein solches Vorgehen daher nicht nur aus forschungspragmatischen Gründen für ökonomisch, sondern auch für äußerst erkenntnisreich: Die Feinanalyse offenbart transsituative Phänomene, die noch nicht verschlagwortet sind, die Verschlagwortung offenbart die Zusammenhänge verschiedener Situationen.

Der große Aufwand, der in die Erhebung und Aufbereitung einer solchen longitudinalen videographischen Studie fließt, macht die Frage nach Möglichkeiten der Nachnutzung des Materials relevant. Um an die zum Teil sehr sensiblen, in jedem Fall jedoch privaten Daten aus dem Feld der dreiköpfigen Familie zu kommen, haben wir ein enges Vertrauensverhältnis aufgebaut. Hierzu gehörte u. a. die datenschutzrechtliche Vereinbarung, dass nur Mitglieder des Projekts ungeschützte Einblicke in die private Lebenswelt der beforschten Familie erhalten. Der Ausschluss der Nachnutzung war in unserem Fall damit notwendige Vorbedingung für den Feldzugang und die Datenerhebung (vgl. Lewandowski; Reichertz/Nebowsky i. d. B.).

Betrachten wir die vielfältigen Schlagworte, mit denen wir unser videoanalytisches Datenkorpus aufgebrochen haben, liegt der Gedanke nahe, dass ein solches Vorgehen eine Nachnutzung auf hervorragende Weise (und ohne beträchtlichen Zusatzaufwand) vorbereiten würde. Wir denken jedoch, dass unser Vorgehen in dieser Hinsicht auch eine Reihe von Problemen aufwirft, für die zunächst noch methodische, technische oder rechtliche Lösungen erarbeitet werden müssten.

Zunächst sind die Daten in unserem Atlas. ti Projekt nicht anonymisiert und wir halten es auch für ausgeschlossen, dass Audio- und Videomaterial in dem Umfang, in dem wir es erhoben haben, gut genug anonymisiert werden kann, dass eine Identifikation der Personen unmöglich wäre und es dennoch noch als

Interaktionsdaten uneingeschränkt nutzbar wäre.[6] Es würde genügen, ein Orts- oder Nummernschild auf einer der gefilmten Autofahrten, oder den Namen einer anrufenden Person auf einem Handybildschirm zu übersehen, um die Anonymität der Gefilmten zu gefährden.

Zudem hat sich unsere Forschungsfrage in den kompletten Verschlagwortungsprozess eingeschrieben. So offen wir zu Beginn der Verschlagwortung waren, so fokussiert waren wir gleichzeitig auf unsere Forschungsfrage: Wir suchten nach Handlungsvorhaben und Aktivitäten, die irgendetwas mit der Bearbeitung von Zukunft zu tun hatten. Andere soziologisch interessante Phänomene des Familienlebens, wie etwa Geschlechterrollen, Erziehungsgespräche usw. wurden dementsprechend nicht mit derselben Gründlichkeit verschlagwortet. Phänomene, die für viele andere Forschungsfragen irrelevant sind (wann und wie der Handwerker ankündigt, das Wasser abzustellen), sind dafür umso detaillierter inventarisiert. Forschende, die an einer Nachnutzung interessiert sind, müssten dies also entsprechend reflektieren und gegebenenfalls selbst inventarisieren, wobei ihnen jedoch die ethnographischen Felderfahrungen fehlen würden, die bei der Verschlagwortung des Materials oftmals das Kontextwissen bereitstellten, um die Bedeutung konkreter Situationen für das Gesamtmaterial einzuschätzen. Die Verschlagwortung von Beginn an so vorzunehmen, dass eine Durchsuchbarkeit des Materials für andere, hypothetische Fragestellungen möglich würde, würde nicht nur den Verschlagwortungsaufwand massiv vervielfachen, sie würde auch von der Fokussierung auf das eigene Forschungsinteresse ablenken und so die Präzision der Aufbereitung zur Bearbeitung der eigenen Forschungsfrage gefährden.

Stattdessen halten wir unser Vorgehen allerdings für äußerst nützlich, um nicht nur Pläne, sondern Prozesse und Trajektorien im Allgemeinen abzubilden. Insbesondere der Schritt, die situative Herstellung von zeitlichen Kontexten der Teilnehmenden zu betrachten und diese zur Grundlage der (jeweils situativ relevanten) Vollständigkeit der Kollektion zu machen, halten wir für ein gutes Vorgehen und für eine Lösung grundsätzlicher Probleme, welche longitudinale Vorgehen oft aufweisen: Wie können wir annehmen, dass Teile der Kollektion tatsächlich Teil der untersuchten Trajektorie sind? Wie können Aussagen über deren Vollständigkeit getroffen werden? Wie kann sichergestellt werden, dass es sich bei dem untersuchten Phänomen um dasselbe handelt wie in der ersten Situation?

6 Selbst wenn es möglich wäre, hunderte Stunden Audio- und Videomaterial zu anonymisieren, müssten wir dies nach derzeitigem technischem Stand tun, ehe wir die Videos in Atlas.ti verschlagworten, denn diese lassen sich nicht nachträglich durch anonymisierte Videos austauschen. Atlas.ti prüft verschiedene Metadaten, sodass sich die Videos später nicht durch anonymisierte Versionen austauschen lassen. Alternativ könnte am Ende des Projekts eine komplette Kopie des Projekts händisch anhand von anonymisierten Daten angefertigt werden.

Indem die sich verändernden Wissensbestände der Teilnehmenden vermittelt durch ihre jeweils situative Hervorbringung in den Blick genommen werden, können diese Fragen beantwortet werden: Es kann aufgezeigt werden, dass und wie Teilnehmende Bezüge zu anderen Situationen herstellen. Hierdurch kann die Vollständigkeit der Kollektion geprüft und die hergestellte Kollektion mit den Relevanzen der Teilnehmenden abgeglichen werden. Es kann auch aufgezeigt werden, ob ein Phänomen über Situationen hinweg als mit sich selbst identisch behandelt wird, oder aber als etwas, das sich so weit gewandelt hat, dass es zu etwas anderem wurde. Schlussendlich lässt sich so der gesellschaftliche Wandel untersuchen, wie er alltäglich von Gesellschaftsteilnehmenden aktiv und geordnet verhindert oder hergestellt wird.

Literatur

Au-Yeung, Terry S. H./Fitzgerald, Richard (2023): Time structures in ethnomethodological and conversation analysis studies of practical activity. In: The Sociological Review 71, H. 1, S. 221–242.

Ayaß, Ruth (2021): Projektive Gattungen. Die kommunikative Verfertigung von Zukunft. In: Weidner, Beate/König, Katharina/Imo, Wolfgang/Wegner, Lars (Hrsg.): Verfestigungen in der Interaktion. Berlin: De Gruyter, S. 57–82.

Clayman, Steven/Elliot, Marc N./Heritage, John/MacDonald, Laurie L. (2007): Historical Trends in Questioning Presidents 1953–2000. In: Presidential Studies Quarterly 36, H. 4, S. 561–582.

Deppermann, Arnulf/Pekarek Doehler, Simona (2021): Longitudinal Conversation Analysis – Introduction to the Special Issue. In: Research on Language and Social Interaction 54, H. 2, S. 127–141.

Friese, Susanne (2019): Grounded Theory Analysis and CAQDAS: A Happy Pairing or Remodeling GT to QDA? In: Byrant, Anthony/Charmaz, Kathy (Hrsg.): The Sage Handbook of Current Developments in Grounded Theory. London: Sage, S. 282–313.

Garfinkel, Harold (1967): Studies in Ethnomethodology. Englewood Cliffs, NJ: Prentice-Hall.

Gibson, David R. (2021): Repetition Acknowledgment Prefaces 44, H. 4, S. 819–842.

Goodwin, Charles (2000): Action and embodiment within situated human interaction. In: Journal of Pragmatics 32, H. 10, S. 1489–1522.

Hardesty, Monica J. (1982): Ethnomethodology and Symbolic Interactionism: A Critical Comparison of Temporal Orientations. In: Symbolic Interaction 5, H. 1, S. 127–137.

Heath, Christian/Hindmarsh, Jon/Luff, Paul (2010): Video in Qualitative Research. Analysing Social Interaction in Everyday Life. London: Sage.

Hitzler, Sarah/Kramer, Jonas (2024, online first) The Trajectory of an Agreement. Tracing Objectivated Knowledge Across a Series of Mundane Encounters. In: Symbolic Interaction, S. 1–24. DOI: 10.1002/symb.1226.

Kolanoski, Martina/Löffler, Marlen S./Küffner, Carla/Terjung, Clara (Hrsg.) (2023): Trans-Sequentiell Forschen. Neue Perspektiven und Anwendungsfelder. Wiesbaden: Springer VS.

Kramer, Jonas/Singh, Ajit (2023): Gemeinsames Planen und die kommunikative Verfestigung sozialer Beziehungen. Zur kommunikativen Konstruktion des familialen Alltags. In: Villa, Paula-Irene (Hrsg.): Polarisierte Welten. Verhandlungen des 41. Kongresses der Deutschen Gesellschaft für Soziologie 2022.

Löffler, Marlen S. (2022): Niederschwellig arbeiten. Eine problemsoziologische Study of Work in Anlaufstellen für männliche Sexarbeitende. Wiesbaden: Springer.

Luckmann, Thomas (1983): Common Sense, Science and the Specialization of Knowledge. In: Phenomenology + Pedagogy, S. 59–73.

Luckmann, Thomas (1986): Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 27, S. 191–211.

Mondada, Lorenza (2007): Multimodal resources for turn-taking. In: Discourse Studies 9, H. 2, S. 194–225.
Mondada, Lorenza (2018): Controversial Issues in Participatory Urban Planning: An Ethnomethodological and Conversation Analytic Historical Study. In: Pekarek Doehler, Simona/Wagner, Johannes/González-Martínez, Esther (Hrsg.): Longitudinal Studies on the Organization of Social Interaction. London: Palgrave Macmillan, S. 287–328.
Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A Simplest Systematics for the Organization of Turn-Taking for Conversation. In: Language 50, H. 4, S. 696.
Scheffer, Thomas (2013): Die trans-sequentielle Analyse – und ihre formativen Objekte. In: Hörster, Reinhard/Köngeter, Stefan/Müller, Burkhard (Hrsg.): Grenzobjekte: Soziale Welten und ihre Übergänge. Wiesbaden: Springer Fachmedien, S. 89–114.
Schmidt, Axel (2018): Prefiguring the future: Projections and preparations within theatrical rehearsal. In: Deppermann, Arnulf/Streeck, Jürgen (Hrsg.): Time in embodied interaction. Synchronicity and sequentiality of multimodal resources. Amsterdam, Philadelphia: John Benjamins Publishing Company, S. 231–260.
Schütz, Alfred (Hrsg.) (1971): Gesammelte Aufsätze. Band 1: Das Problem der sozialen Wirklichkeit. Den Haag: Nijhoff.
Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion 10, S. 353–402.
Strauss, Anselm L./Fagerhaugh, Shizuko/Suczek, Barbara/Wiener, Carolyn (1985): Social organization of medical work. Chicago: University of Chicago Press.
Strübing, Jörg/Hirschauer, Stefan/Ayaß, Ruth/Krähnke, Uwe/Scheffer, Thomas (2018): Gütekriterien qualitativer Sozialforschung. Ein Diskussionsanstoß. In: Zeitschrift für Soziologie 47, H. 2, S. 83–100.
Wagner, Johannes/Pekarek Doehler, Simona/González-Martínez, Esther (2018): Longitudinal Research on the Organization of Social Interaction: Current Developments and Methodological Challenges. In: Pekarek Doehler, Simona/Wagner, Johannes/González-Martínez, Esther (Hrsg.): Longitudinal Studies on the Organization of Social Interaction. London: Palgrave Macmillan, S. 3–35.
Wootton, A. J. (1997): Interaction and the Development of Mind. Cambridge: Cambridge University Press.

Klangorientiertes kommunikatives Handeln

Herausforderungen und Potenziale der Videoanalyse gemeinsamen Musizierens

Theresa Vollmer

Zusammenfassung: (Gemeinsames) Musizieren ist eine alte Kommunikationsart, die sich überall findet. Sie interpretativ zu untersuchen, konfrontiert die Methoden der Sozialwissenschaften mit Herausforderungen, wenn sie in der Datenerhebung oder -analyse vor allem mit Sprache arbeiten. Videodaten bieten einen besonderen Zugang für die Analyse von Sozialphänomenen, die sich durch die Abwesenheit des Sprechens auszeichnen, ohne still zu sein. Die Interpretation klangorientierter Kommunikation bedarf hierzu Untersuchungsinstrumente, die die flüchtigen akustischen Ereignisse in ihrer Prozesshaftigkeit der Analyse verfügbar machen und die Verbindung mit klangerzeugenden Handlungen aufzeigen. An Alfred Schütz anschließend lege ich die Erkenntnispotenziale der Videographie gemeinsamen Musizierens dar. Illustriert durch einen Fall aus einer Ethnographie zu Streichensembles werden die Herausforderungen der Videoanalyse nachgezeichnet. Die Methodologie der Lösung, das »Transkript der rhythmischen Impulse«, wird fokussiert, um am Beispiel des Chorsingens die Transferierbarkeit des Analyseinstrumentes zu diskutieren.

Schlüsselwörter: Transkription, Synchronisation, Noten, Streichensembles, Chöre

1. Gemeinsames Musizieren: Schütz' Hoffnung auf ein besseres Verständnis von Interaktionen

Wenn Menschen miteinander musizieren, synchronisieren sie ihre Handlungen so präzise, dass sie ein hörbares Kollektivgebilde konstruieren: gemeinsame Musik. Das ist eine herausragende Leistung, die ein Grundinteresse adressiert, das nicht nur die Musiksoziologie, sondern die Soziologie und Sozialtheorie gleichermaßen betrifft: Wie gelingt es Handelnden, gemeinsam Soziales zu erschaffen (Vollmer 2023)? Dass die Untersuchung insbesondere musikalischer Kommunikation zur Beantwortung dieser Frage beitragen könnte, ist eine Hoffnung, die Alfred Schütz in seinem Aufsatz »Gemeinsam Musizieren« weckt (Schütz 2016 [1951]).

Obwohl sich das gemeinsame Musizieren im Fall des Gelingens durch die Abwesenheit des Sprechens auszeichnet, sei es nicht weniger sinn-voll (ebd.,

S. 149). Der musikalische Sinn sei aber charakteristisch anders geartet als der der Sprache, sodass »die Begriffe ›Sinn‹ und ›sinnhafter Zusammenhang‹, ›Verstehen‹ und ›Interpretation‹ […], in einer spezifischen Weise gebraucht [werden], die sich von der Anwendung in anderen Sinnsystemen, wie etwa Sprachen unterscheidet« (Schütz 2016 [1944], S. 83). Der besondere Reiz des musikalischen Sinns liegt für Schütz in dessen genuin zeitlichem Charakter: »Die spezifische Existenz des idealen Gegenstandes ›musikalisches Werk‹ ist seine Ausdehnung in der Zeit« (ebd., S. 89).

Diese Zeitlichkeit des Sinns stellt miteinander musizierende Menschen vor die Aufgabe, ihre inneren Bewusstseinsströme zu synchronisieren (vgl. ebd.), sodass sie eine gemeinsame Musik erzeugen können. Die Untersuchung dieser Kommunikationsleistung fasziniert Schütz, da sie »Einsichten erbringen wird, die für viele andere Formen des sozialen Umgangs gelten. Sie vermag möglicherweise sogar Licht auf einen bestimmten Aspekt der Struktur sozialer Interaktion als solcher zu werfen, der seitens der Sozialwissenschaftler[:innen] bislang noch nicht die Aufmerksamkeit auf sich gezogen hat, die er verdient« (Schütz 2016 [1951], S. 149 f.).

Dabei ist die Beforschung des Musizierens eine Herausforderung für die konventionellen Forschungsprogramme der Sozialwissenschaften, wenn und insofern sie in der Datenerhebung oder -analyse vor allem auf Sprache aufbauen – wie z. B. bei Interviewverfahren oder Feldnotizen. Videodaten schaffen hier ein neuartiges Potenzial für die interpretativ, explorative Untersuchung gemeinsamen Musizierens, die Schütz 1951 in dieser Form noch nicht zur Verfügung standen. Zugleich bleibt die Analyse dieser Daten eine Herausforderung, weil die musikalischen Klänge flüchtig sind und an die Permanenz ihrer Erzeugung und ihres ›Erhörens‹ gebunden sind.

Um diese Videoanalyseprobleme näher zu beleuchten und zu ihrer Lösung beizutragen, wird im Folgenden zuerst eine Erhebungssituation des gemeinsamen Musizierens von Streichensembles vorgestellt, das ich im Rahmen meiner Promotionsforschung untersuchen konnte. Ein für diese Untersuchung integrierendes ethnographisches Forschungsdesign, hatte Videodaten als eine der Hauptdatensorten. Anhand der Darstellung einer aufgezeichneten, exemplarischen Spielsequenz, in der sich ein Streichquartett mit einem Rhythmusproblem konfrontiert sah, werde ich die Herausforderungen der Videoanalyse dieses Musizierens aufzeigen. Das zur Lösung dieser Probleme entwickelte Verfahren »Transkript der rhythmischen Impulse« erlaubt die integrierende Interaktionsanalyse verschiedener sozialer Elemente, die die musikalische Kommunikation des Quartetts bedingten und ausmachten: Die den Musiker:innen vorliegenden Noten, ihre leibkörperliche Kommunikation während des Spiels und die individuell erzeugten musikalischen Klänge. Um die Potenziale und Herausforderungen der Transferierbarkeit dieses Analyseverfahrens auf andere Formen der (musikalischen) Kommunikation diskutieren zu können, wird seine Methodologie

in Verschränkung mit der Sozialtheorie des kommunikativen Handelns herausgestellt. Ob und wie sich das Verfahren auf andere Formen gemeinsamen Musizierens übertragen lässt, wird anhand des Beispiels der Interaktionsanalyse des Chorsingens reflektiert, das ich derzeit mit Michaela Pfadenhauer an der Universität Wien untersuche, um mit einem kurzen Ausblick auf die Chancen der interpretativen Videoanalyse klangorientierter Kommunikation zu schließen.

2. Die Erhebungssituation: Mehrstimmiges Musizieren mit Noten, das Proben und die soziologische Partitur gemeinsamen Musizierens

In meiner ethnographischen Promotionsforschung zum Musizieren von *Streichensembles* (Vollmer 2023) habe ich an die oben angeführten Überlegungen Schütz' angeschlossen. Wenn Menschen in Streichensembles miteinander spielen, erzeugen die einzelnen Akteure klassischerweise differente musikalische Stimmen (ohne die Hilfe von Dirigierenden), deren Zusammenklang einen Sinn sui generis erzeugt, der mehr ist als die Summe des Sinns der einzelnen Stimmen – polyphone Musik (vgl. Schütz 2016 [1951], S. 164).

Dafür verwenden sie *Noten*, die bereits existierende Musikstücke durch konventionale Zeichen objektivieren. Diese Zeichen unterscheiden sich systematisch vom idealen Gegenstand der Musik, denn sie sind »keine Bilder der Töne«, sondern »Ausdrucksmittel in einer konventionalen Sprache für […] Anweisungen« (Schütz 2016 [1951], S. 153), die Musiker:innen befolgen müssen, wenn sie ein Musikstück reproduzieren wollen. Zu diesem Zweck müssen sie zunächst Fähigkeiten erlernen, die sie in die Lage versetzen, die Zeichen der Noten zu lesen und quasi gleichzeitig als musikalische Handlungen mit ihren Instrumenten umzusetzen.

Die Realisation der notierten Musik stellt eine (soziale) Handlungsherausforderung für die Beteiligten dar, sodass das Spiel mit Noten mit einem mehr oder weniger langen Übungs- und Probenprozess verknüpft ist. – Nur ein verhältnismäßig geringer Teil dieses Musizierens realisiert sich in Aufführungen. Dabei erscheinen Noten in den gemeinsamen Musiziersituationen als »Interaktionsskript« (Vollmer 2023, S. 350 ff.): Sie enthalten zu synchronisierende Spiel- oder Singanweisungen – allgemeiner *Musizieranweisungen* – für die einzelnen Beteiligten, die in situ gleichzeitig realisiert werden sollen. Auf diese Weise beeinflussen Noten die Interaktionsordnung des gemeinsamen Musizierens, ohne sie zu determinieren. Es stellt sich die Frage, wie es den Beteiligten gelingt, die in den Noten objektivierten Musikstücke hörbar zu realisieren und ihre individuellen, musikalischen Handlungen so präzise zu synchronisieren, dass im Fall des Gelingens eine gemeinsame Musik erklingt?

Um dieser Frage explorativ nachzugehen, bieten sich *Proben* als Untersuchungssituationen an, weil das Realisieren der notierten Stücke noch nicht

gelingt. Der Umstand dieses ›vorläufigen Scheiterns‹ eröffnet einen verstehenden Zugang zu den Bedingungen gelingenden Musizierens, insofern die Beteiligten selbst situativ zu erkennen geben, wenn etwas noch nicht funktioniert: Das gemeinsame Spiel der notierten Musik wird versucht, unterbrochen, Probleme und Interpretationen besprochen und das Spiel immer wieder wiederholt, bis es peu á peu ›besser‹ gelingt, das Stück im Sinne der Notation und Interpretation der Beteiligten zum Klingen zu bringen.

Dadurch entstehen sequenzielle Strukturen durchaus typisierbarer Formen und gattungsförmiger Züge (vgl. Knoblauch 2024, S. 279), insofern ein reges Wechselspiel zwischen »Spiel-« bzw. Singphasen – allgemeiner *Musizierphasen* – und »Redephasen« entsteht (Vollmer 2023, S. 368 ff.). Dieser Wechsel der Phasen muss als idealtypisch verstanden werden, insofern sie empirisch nicht immer als ›reine‹ Musizier- oder Redephasen auftreten (z. B. kann es vorkommen, dass in Redephasen gesungen oder in Musizierphasen gerufen wird). Der Übergang zwischen den Phasen ist vielmehr selbst eine zu interpretierende, prozesshafte Leistung der Beteiligten und nimmt unterschiedliche Formen an. Die Darstellung der Videoanalyse dieses Beitrages konzentriert sich insbesondere auf eine solche Musizierphase.[1]

In meiner Forschung bin ich der Frage nachgegangen, wie es klassisch ausgebildeten Musiker:innen gelingt, mit ihren Instrumenten und Noten in Streichensembles gemeinsam zu musizieren und wie sie dies dabei erleben. Diese Frage enthält mehrere analytische Dimensionen des Sozialen, die ich mithilfe einer Metapher verknüpft habe, die ich als *soziologische Partitur gemeinsamen Musizierens* (Vollmer 2023, S. 1–12) bezeichne. Sie umfasst wie eine Streichtriokomposition drei zusammengehörende, autarke Stimmen des Sozialen.

Eine Art Generalbass bildet das *Wissen*: als geteilte Wissensbestände ausgebildeter Musiker:innen (i) und gemeinsames Wissen, das Ensembles durch gemeinsame Proben hinweg entwickeln und das sich z. B. in Hörerwartungen widerspiegelt (ii). Die hier adressierten sozialen Aspekte reichen über individuelle Situationen gemeinsamen Musizierens hinaus und können nicht allein aus der situativen musikalischen Kommunikation abgeleitet werden.

Als eine Art Melodie hebt sich die analytische Stimme, die ich als *Hörwelt* bezeichne, ab. Das gemeinsame Musizieren ist mit den subjektiven Bewusstseinsleistungen (vgl. Schütz 2016 [1944] und [1951]), individuellen Handlungen und dem situativen Wahrnehmen der Beteiligten verknüpft, sodass die Hörwelt diese situative Handlungsperspektive von Musiker:innen adressiert.

1 Auch die Untersuchung von Redephasen stiftet relevante Erkenntniszugänge für die interpretative Untersuchung gemeinsamen Musizierens, weil z. B. durch die Verbalisierung etwaiger Probleme das Sinnverstehen ergänzt wird (Vollmer 2023, S. 322–328 und S. 467–471).

Gemeinsames Musizieren kann weder allein durch das Wissen noch aus der individuellen Handlungsperspektive der Musiker:innen verstanden werden. Vielmehr stellt sich die Frage, was sie im kommunikativen Handeln leisten, um gemeinsam zu musizieren. Zentral sind hierbei die erzeugten Klänge, sodass ich diese Stimme der Partitur als *Klangwelt* bezeichne und die in den folgenden Ausführungen fokussiert wird.

3. Die Ethnographie gemeinsamen Musizierens mit Videodaten

Das Streichensemble-Musizieren habe ich mit einem Forschungsdesign untersucht, das durch die fokussierte Ethnographie gerahmt ist (Knoblauch 2001), Methoden der lebensweltanalytischen Ethnographie (Honer 1989) und Videographie (Tuma/Schnettler/Knoblauch 2013) integriert, sodass u. a. Videodaten eine der Hauptdatensorten waren. Die Methodologie der *fokussierten Ethnographie* ist insbesondere durch theoretische Konzepte Goffmans (2009 [1963], S. 29–45), wie »fokussierte Interaktionen« und »soziale Situationen«, geleitet. In der Untersuchung ging es nicht um rein situative Geschehnisse. Vielmehr waren transsituative Wissensbestände und der sozialkulturelle Kontext (vgl. Knoblauch/Vollmer 2022) des Musizierens Ausgangspunkt und Ergebnis der Forschung, um die Fokussierungen der Feldakteure (re)konstruieren zu können, sodass die fokussierte Ethnographie vor allem der Adressierung der sozialanalytischen Stimme »Wissen« der soziologischen Partitur diente.

Die Methoden der *lebensweltanalytischen Ethnographie* dienten insbesondere der explorativen Untersuchung der Handlungsperspektive der Musiker:innen, um die analytische Stimme der »Hörwelt« zu perspektivieren. Zum Beispiel habe ich als Cellistin in Streichensembles musiziert (deren Spiel ich mitunter mit Video aufzeichnen konnte), während sich meine Erfahrungen als Streichmusikerin im Forschungsprozess zunehmend als Teil der grundlegenden Interpretationsressourcen erwiesen, insofern ich als Cellistin Teil des Forschungsfeldes war.

In den Situationen gemeinsamen Musizierens konnte ich neben Datensorten wie klassischen Feldnotizen und ethnographischen Gesprächen meiner teilnehmenden Beobachtungen und Erlebensdaten (Pfadenhauer 2005) meiner beobachtenden Teilnahmen, Videos erstellen, um insbesondere die sozialanalytische Stimme der »Klangwelt« untersuchen zu können. Die Analyse der erstellten Videos wird mit der Ethnographie verbunden und als Forschungsprozess als *soziologische Videographie* bezeichnet (Knoblauch 2024, S. 275 f.).

Auch wenn die *Datensorte Video* durch andere ergänzt werden (muss), erlaubt sie einen besonderen Zugang zur musikalischen Kommunikation, wenn sie als »Forschungsmedium« dient, um die aufgezeichneten Interaktionen zu analysieren (Tuma/Schnettler/Knoblauch 2013, S. 46–48). Audiovisuelle Daten sind in der Lage, Abschnitte einer Feldsituation zu »registrieren« (Bergmann 1985),

ohne die aufgezeichneten Ereignisse durch ihre Versprachlichung vorzuinterpretieren. Sie konservieren das aufgezeichnete Geschehen im Zusammenspiel seiner sozialen Elemente (synchron) als auch in seiner zeitlichen Abfolge (diachron), wobei die zeitliche Struktur der aufgezeichneten Ereignisse dieselben objektiven Zeitmaße aufweist wie in ihrem situativen Vollzug (Tuma/Schnettler/Knoblauch 2013, S. 31; vgl. Knoblauch/Wilke i. d. B.).

Da die Videos das Konstrukt von Kamerahandlungen (Reichertz/Englert 2011) sind, wird die Darstellbarkeit der registrierten Kommunikation beeinflusst: Die Aufnahmegeräte werden ein- und ausgeschaltet, das Musizieren wird aus einer spezifischen Kameraperspektive aufgezeichnet, die Räumlichkeit der Situation wird auf die Zweidimensionalität eines Bildschirms gebannt und es fehlen sinnliche Eindrücke, wie Gerüche oder Temperatur. Auch die akustische Qualität der musikalischen Klänge wird durch die Aufzeichnungen verändert, da sie mit dem wortwörtlichen Raum verbunden sind, in dem sie erklingen – etwa ein Zimmer mit Teppich oder eine Kirche. Wegen des konstruierten Charakters des Forschungsmediums ist die Rekonstruktion der Bedingungen und Möglichkeiten ihrer Erstellung zentral. Zum Beispiel gewinnen die leibkörperlich zueinander ausgerichteten Musiker:innen, die in den ›mit-klingenden Räumen‹ positioniert sind, jeweils einen unterschiedlichen »positionalisierten Höreindruck« (Vollmer 2023, S. 329–341), der allein durch die Videoaufnahmen nicht rekonstruiert werden kann. Und: Die gefilmten Geschehnisse verlieren ihre natürliche, situative Flüchtigkeit, die für die Beobachteten charakteristisch ist (Bergmann 1985), weil die Videos erlauben, die Ereignisse beliebig oft abzuspielen, zu verlangsamen usw. Das macht sie gleichwohl zu einer fabelhaften Datensorte für die interpretative Beforschung minutiöser musikalischer Kommunikation.

4. Der exemplarische Fall: Das Rhythmusproblem eines Streichquartetts

Um die intersubjektiv geteilte »Klangwelt« des Streichensemblespiels untersuchen zu können, war u. a. die Videographie der Proben eines Streichquartetts von Berufsmusiker:innen ein relevanter Erkenntnisschlüssel. Als klassisches Streichquartett setzt es sich aus zwei Geigen, einer Bratsche und einem Cello zusammen. Der im Folgenden beschriebene Fall stammt aus der zweiten von vier Proben für ein Konzert, in dem die Gruppe das »Streichquartett Nr. 4 in C-Dur« von Franz Schubert spielen wird.

Im Laufe der Videoanalyse fokussierte ich einen probentypischen Wechsel von einer Musizier- zu einer Redephase (s. o.) genauer, bei dem es sich um einen *Abbruch* handelt: Die Gruppe spielt einen Abschnitt der Komposition und ohne, dass das Stück ein Ende erreicht hätte, wird das Spiel durch die Musiker:innen unterbrochen. Dabei lässt sich in den letzten elf Sekunden des Musizierens durch

die Mimik der Musiker:innen erkennen, dass ›etwas Außerordentliches‹ vor sich geht. Irgendetwas passiert – oder nicht – im Spiel der letzten sieben Takte des Kompositionsabschnitts (siehe Abb. 1), das zu diesem Abbruch führt.

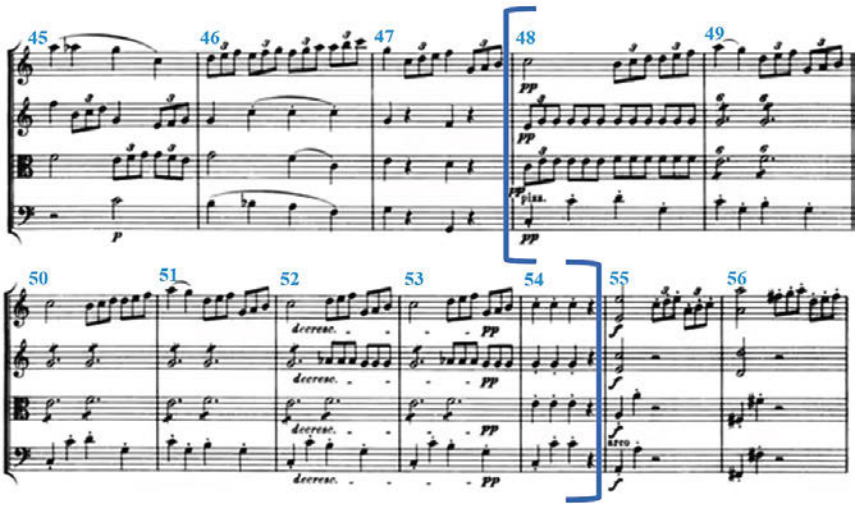


Abbildung 1: Partitur des Kompositionsabschnitts (Schubert 1890 [1812], eigene, blaue Markierungen)

Der Abbruch wird durch den Bratschisten an einer musikalisch günstigen Stelle initiiert – ein musikalischer Sinnabschnitt (vgl. Schütz 2016 [1944], S. 128) wurde gerade beendet, bevor nach einer Viertelpause das Spiel des nächsten Abschnittes fortgeführt werden würde. Das Abbrechen geschieht dabei so fließend und schnell, sodass es wirkt, als ob die musikalische Kommunikation einfach als sprachliche fortgeführt werden würde. Es entsteht folgender Gesprächsauszug:

Bratschist: »Wer hat Recht?«
(Die erste Violinistin und Cellistin lachen.)
Cellistin: »Das habe ich mich auch gerade gefragt. Weiß' nicht.«
Bratschist: »Das Cello hat immer Recht!«
Cellistin: »Wahrscheinlich hatte ich Recht. Ich war aber auch …
kein guter General.«
(Erste und zweite Violinistin und Cellistin lachen.)

Dieser kurze Austausch ist soziologisch betrachtet besonders hübsch, weil verbal nicht expliziert wird, was eigentlich passiert ist. Dennoch reagieren die Musikerinnen u. a. mit Lachen und stellen keine Nachfragen, was gemeint ist. Den Beteiligten scheint klar zu sein, worum es geht, ohne dass das aus dem Inhalt des gesprochenen Wortes hervorgeht.

Im Folgenden werde ich nicht den Gesprächsauszug interpretieren (vgl. dazu Vollmer 2023, S. 467–474), sondern zeigen, wie ich die *Musizierphase* analysiert habe. Als Cellistin konnte ich mithilfe der Videoaufnahme hören, dass die Synchronisierung des Rhythmus nicht gelingt (insofern Streichensembles keine Dirigierenden oder andere ›Richter:innen‹ [z. B. Metronome] einsetzen, müssen die Beteiligten ihr Spiel selbst wechselseitig koordinieren). In Interpretationssitzungen zeigte sich, dass musikalische Laien das Problem der Spielenden mitunter überhaupt nicht hören konnten – auch wenn sie anhand der Mimik der Musiker:innen erkannten, dass ›etwas Außerordentliches‹ vor sich ging und das Spiel abgebrochen wurde. Musikalisch erfahrene Betrachter:innen wiederum hörten (manchmal nach mehrmaligen Schauen), dass ein Rhythmusproblem vorhanden war, ohne aber en Detail Auskunft darüber geben zu können, wie genau es sequenziell erzeugt wurde.

Tatsächlich erwies sich die Analyse dieser Musizierphase als überraschend anspruchsvoll. Überraschend deshalb, weil bereits ein Zeichensystem existiert, um Musik zu notieren: Die Noten, die die Musiker:innen beim gemeinsamen Spiel verwenden. Allerdings zeigte sich in der Untersuchung, dass das situative, individuelle Musizieren ungleich der notierten Musik ist und dass diese Differenz zwischen gespielter und notierter Musik selbst eine Motivation für das Abbrechen sein kann.

5. Das Auswertungsverfahren I: Das »Transkript der rhythmischen Impulse«

Zur Videoanalyse des Musizierens bedarf es eines Interpretationsverfahrens, das nicht nur das Sprechen in den ›Blick‹ nehmen kann. In die Videographie ist zu diesem Zweck die »interpretative Videointeraktionsanalyse« integriert (Tuma/Schnettler/Knoblauch 2013, S. 85–104), die u. a. Verfahren der Workplace Studies (Luff/Hindmarsh/Heath 2000) und Konversationsanalyse (Sacks 1992) aufgreift und adaptiert. Damit werden im Prozess der Videoanalyse Anleihen an den theoretischen Grundlagen der Ethnomethodologie Garfinkels (1984) genommen, die an Schütz' Soziologie anschließt. »Ziel der Analyse ist […] das sequenziell aufeinander bezogene Handeln der Akteure, das die gemeinsam erarbeitete Wirklichkeit beständig in Situationen hervorbringt, also situativ und zugleich interaktiv erzeugt« (Tuma/Schnettler/Knoblauch 2013, S. 15).

Auch wenn diese Methodologie einen Zugang zu musikalischer Kommunikation eröffnete, stellte ich beim Analysieren fest, dass ich ein neues Transkriptionsverfahren brauchte, das die Sequenzialität der hörbar erzeugten Kommunikationsereignisse der Interpretation zugänglich macht. Dabei ging es nicht um die physikalischen Eigenschaften der Akustik, wie z. B. Lautstärke oder Frequenz, die sich mithilfe der Audiospur messen lassen. Ich war in der Analyse vielmehr mit

der Gleichzeitigkeit und Folge vieler einzelner Klänge konfrontiert, die in Verbindung mit den Zeichen der Noten standen und durch die einzelnen Akteure erzeugt wurden. Durch das prozesshafte Zusammenspiel dieser Klänge entfaltete sich das Rhythmusproblem, das u. a. durch die Mimik der Beteiligten während des Spiels sequenziell kommentiert wurde.

Nach mehreren gescheiterten Transkriptionsversuchen bin ich dazu übergegangen, nicht die musikalischen Klänge selbst, sondern die *Zeitlichkeit ihrer Erzeugung* zu erfassen. Die Noten, die die Musiker:innen in dieser Situation vor sich haben, war hierfür ein hilfreicher Schlüssel, da sie die Spielanweisungen für die einzelnen Beteiligten abbilden. In der Partitur lässt sich erkennen, welche von ihnen zeitgleich durch die verschiedenen Akteure realisiert werden sollen (siehe in Abb. 2 exemplarische Markierungen zeitgleich zu realisierender Noten).

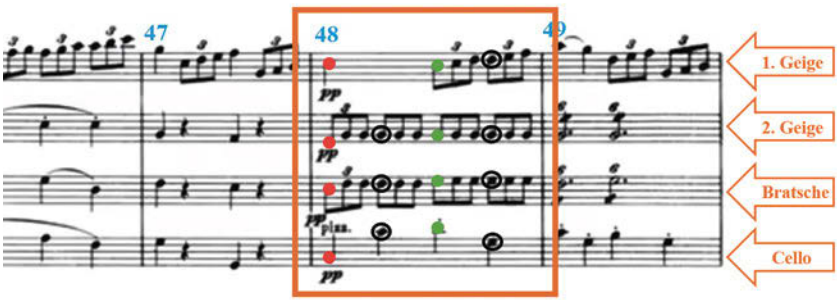


Abbildung 2: Takt 48 (Schubert 1890 [1812], eigene Markierungen)

Die Umsetzung dieser Anweisungen kann beobachtet werden: Die Beteiligten machen von mir sogenannte *Musizierbewegungen* an ihren Instrumenten, die das notierte Stück hörbar machen. Deswegen sind sie hilfreich, um sichtbar zu machen, was situativ hörbar ist. Dazu habe ich die Videosequenz der Musizierphase in Standbilder zerlegt (aufgerundet 17 Bilder pro Sekunde), die wie eine Art Daumenkino funktionieren, bei dem die einzelnen Bilder sozusagen im objektivgleichen Zeitabstand erscheinen. Für die Musiksequenz sind so um die 200 Bilder entstanden, die ich chronologisch nummeriert habe.

Für die Rekonstruktion des erzeugten Rhythmus sind diejenigen Aspekte der Musizierbewegungen relevant, die die Zeitlichkeit der Klänge bedingen. Bei klassischer Streichmusik sind das die Bogen- und Zupfbewegungen der Instrumentalist:innen, wobei ich mich auf das Erzeugen der Töne konzentriert habe, die auf den vier Zählzeiten der Takte liegen (siehe Abb. 2), weil das Quartett ein Rhythmusproblem erzeugt. Um die Momente des individuellen Erzeugens dieser Töne zu ermitteln habe ich die nummerierten Standbilder am Computerbildschirm chronologisch vor- und rückwärts betrachtet. So konnten die sehr schnellen und

kleinen Musizierbewegungen ›in Ruhe‹ rekonstruiert werden, wobei mein Wissen als Cellistin relevant wurde, um erkennen zu können, ›bei welchen Standbildern‹ die notierten Anweisungen umgesetzt wurden. In der Abbildung 3 ist die Darstellung einer exemplarischen Bogenbewegung zu sehen, die die erste Violinistin macht, um ihre erste Note aus Takt 48 zu realisieren (die Nummern unter den anonymisierten Bildern entsprechen dabei der Chronologie: ›In Bild 14 beginnt die Violinistin ihre Bogenbewegung, die sie in Bild 25 beendet hat‹).

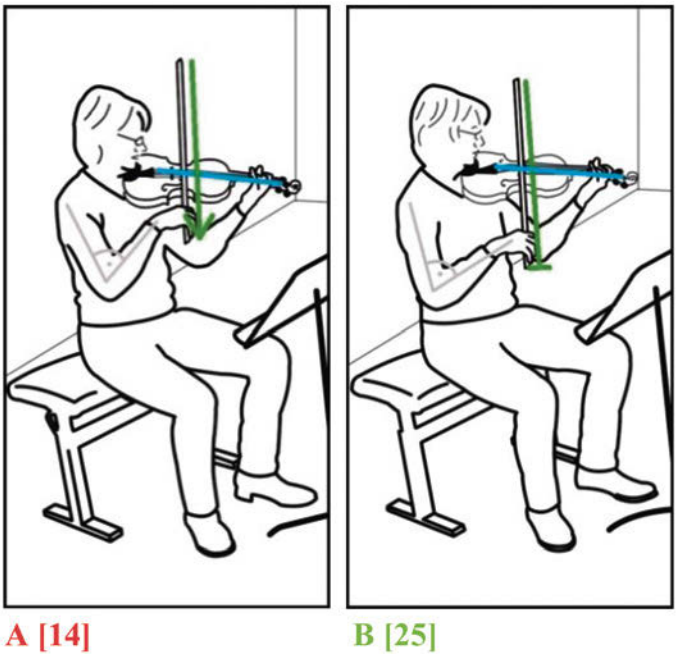


Abbildung 3: Bogenbewegung der ersten Violinistin zur Realisierung ihrer ersten Note aus Takt 48 (eigene Darstellung)

Vereinfacht gilt die Faustregel: Solange eine Saite der Instrumente in Schwingung versetzt wird, ist ein musikalischer Klang zu hören, sodass die Klänge mehr oder weniger direkt an die Zeitlichkeit ihrer Erzeugung gebunden sind. Mithilfe der nummerierten Standbilder war es möglich, diese Zeitlichkeit der individuellen Bewegungen in ein Verhältnis zueinander zu setzen: Ich ermittelte alle individuellen Standbildnummern, ›in denen‹ die Musiker:innen ihre Töne der Zählzeiten produzierten. Diese Nummern übertrug ich in ein Transkript, dass ich das *Transkript der rhythmischen Impulse* nenne (siehe Abb. 4). Die fortlaufenden Nummern des Transkriptes entsprechen den nummerierten Standbildern, die zu Päckchen angeordnet sind, die die gemeinsame Erzeugung der einzelnen Takte des Kompositionsabschnitts abbilden.

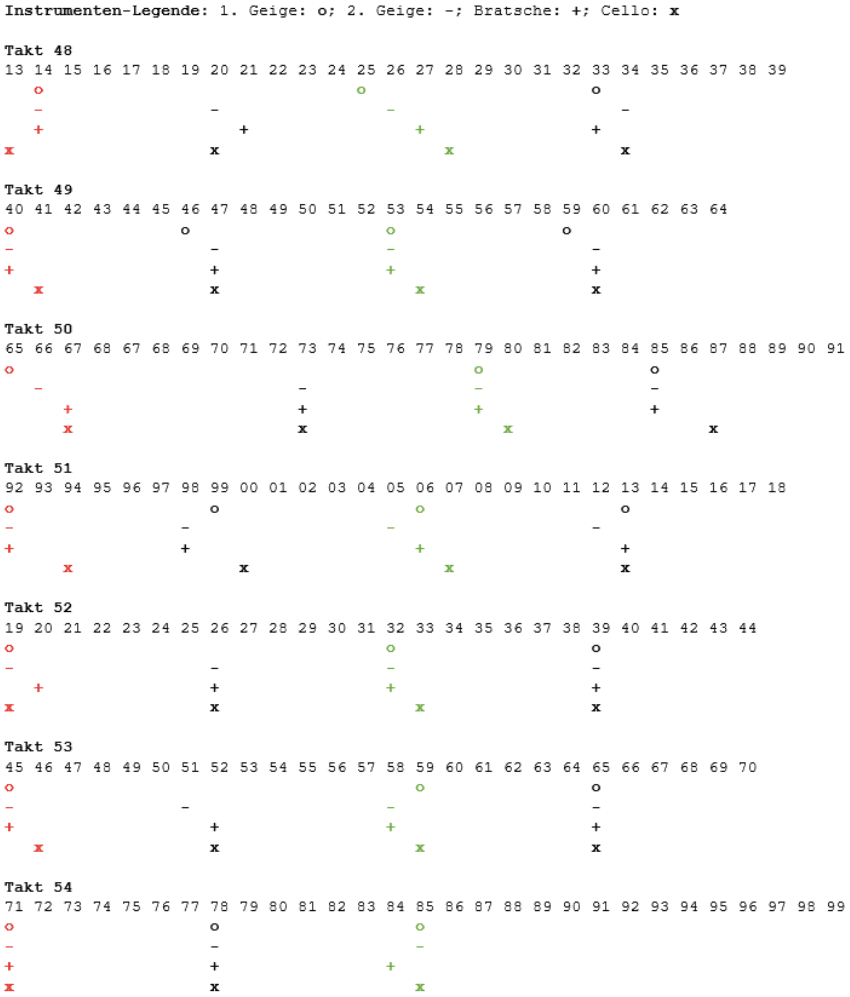


Abbildung 4: »Transkript der rhythmischen Impulse« (eigene Darstellung)

Durch dieses Transkript lässt sich erkennen, dass die erste Violinistin ihre Töne der Zählzeiten in den meisten Fällen etwas schneller erzeugt als die Cellistin bzw. vice versa. Genau diese beiden Stimmen waren in diesem Kompositionsabschnitt besonders gut herauszuhören: Die erste Violinistin spielt die hohe Melodiestimme. Die Cellistin zupft in diesen Takten jeweils vier Töne des 4/4-Taktes, sodass auch ihre Stimme durch den Zupfklang gut herauszuhören ist. Besonders an den Stellen, wo die Symbole für die Zählzeiten-Töne der ersten Violinistin und Cellistin weit(er) auseinander liegen, ist das Auseinanderfallen des Rhythmus deutlicher zu hören (z. B. bei der 3. Zählzeit von Takt 48, 1. und 4. Zählzeit von Takt 50, 1. Zählzeit von Takt 51). Zudem lässt sich ablesen, wie die zweite Violinistin und

der Bratschist ihren Rhythmus realisieren. Sie spielen eine Art Begleitung, die in diesem Kompositionsabschnitt nicht stark hervorsticht. Durch das Transkript lässt sich feststellen, dass beide immer wieder zwischen dem Rhythmus der ersten Violinistin und der Cellistin hin und her schwanken. Es scheint so, als ob sie nicht wüssten, an wem sie sich orientieren sollen. Das war nicht nur eine neue, sondern auch ›interpretativ hübsche‹ Feststellung, wenn man sich in Erinnerung ruft, dass der Bratschist, das Musizieren mit den Worten »Wer hat Recht?« abbricht. Dieses Transkriptionsverfahren ermöglichte es darüber hinaus, z. B. den Zusammenhang zwischen der Mimik der Beteiligten im Verhältnis zum entstehenden Rhythmusproblem zu untersuchen, ohne den Bezug zu den Noten zu verlieren.

6. Das Auswertungsverfahren II: Die Methodologie des Transkriptionsverfahrens kommunikativer Klangobjektivierungen

Die Methodologie dieses Verfahrens lässt sich durch die Begriffe der Sozialtheorie kommunikativen Handelns (Knoblauch 2017, S. 75–188) explizieren. Da empirische Wirklichkeit grundsätzlich darauf angewiesen sei, dass »soziales Handeln kommuniziert […] und dadurch erst beobachtbar oder berichtbar wird«, wird sie »durch Objektivierungen erfahrbar« (ebd., S. 85). »Wie immer der Sinn geartet ist und welche Handlung auch immer intendiert sein mag – Sinn und Handeln müssen den Bereich des subjektiven Sinns überschreiten, um zum Gegenstand der Erfahrung Anderer und damit sozial werden zu können. […] Dank dieser Objektivierungen ist das soziale Handeln auch immer ein kommunikatives Handeln« (ebd., S. 86).

Die »Modalität der Kommunikation« (vgl. Kress 2010 und Tuma/Schnettler/Knoblauch 2013, S. 15) in Musizierphasen zeichnet sich dadurch aus, dass die Beteiligten insbesondere durch *Klänge* miteinander wirken, wenn sie hörbare Musik erschaffen. Sie sind »Ereignisse in der äußeren Welt« (Schütz 2016 [1951], S. 155) und sind – analytisch betrachtet – mehrseitig, insofern sie zum einen durch die individuellen Handlungen der Beteiligten *produziert* werden, während sie zugleich (inter-)subjektiv *hörbar* (und teilweise spürbar) sind. Diese sozialtheoretischen Facetten der empirischen Klänge lassen sich durch die Konzepte der »Sinnlichkeit« (Knoblauch 2017, S. 124–131) und »Objektivierungen« (ebd., S. 155–169) perspektivieren. Dabei ist das Konzept der Sinnlichkeit ebenso ›vielseitig‹, weil es mit mehreren analytischen Dimensionen verknüpft ist: »Was immer ›Materie‹ sein mag – es muss *cum grano salis* sinnlich erfahrbar sein, und zwar so, dass die Art der Materialität durch die Art der Sinnlichkeit bestimmt wird. Die Sinnlichkeit ist somit die Form der Materialität« (ebd., S. 128).

Klänge als Objektivierungen kommunikativen Handelns gehen dabei mit einer Art der »Materialität« einher, die sie besonders flüchtig macht, insofern ihre »Objektpermanenz« (ebd., S. 166 f.) die einzelnen Situationen des Musizierens nicht überdauert (wenn sie nicht durch Aufzeichnungen zu »Objektivationen« [ebd., S. 161–169] gemacht werden): Objektivierungen werden als prozessuale Phänomene gefasst, die mit den leibkörperlichen Handlungsvollzügen, die sie hervorbringen, verbunden sind (z. B. das ›Zum-Schwingen-Bringen‹ der Saiten eines Streichinstrumentes). Sie stellen die »Wirkseite« des kommunikativen Handelns dar und sind »das dritte Moment, das die Relation zwischen zwei [oder mehreren] Subjekten« ausmacht (ebd., S. 161). Dabei korrespondiert die Zeitlichkeit der objektivierten Klänge durch ihre Sinnlichkeit sozusagen mit der Permanenz des Hörens derselben.

Dabei darf nicht vergessen werden, dass die Beteiligten bereits musikalisches Vorwissen haben (s. o.). Das gilt für ihre Fähigkeiten des Notenlesens und des Klangproduzierens ebenso wie für das Hören beim Musizieren. Dieser sozialisierte Aspekt des Wahrnehmens ist von Relevanz: »Was wir wahrnehmen, ist immer auch durch kommunikativ vermittelten Sinn sozialisiert. Freilich haben wir innere Wahrnehmungen, die sich auf unseren Körper beziehen können (der Bauch drückt), auf psychischen Vorstellungen beruhen (wie stellen uns vor, wie es jetzt wäre, ein kühles Bier zu trinken) oder in der Reflexion sich auch auf Gedanken beziehen können (die wir uns wiederum überlegen können). All dies sind aber vollständig sozialisierte Wahrnehmungen. Ob das Bauchdrücken überhaupt der Rede wert ist, was an einem kühlen Bier so attraktiv und was an den Gedanken so eigen ist, hängt von sozialen Konventionen ab« (ebd., S. 126).

Da die Klänge durch die (sozialisierte) Wahrnehmung – hier besser: Sinnlichkeit – der Musiker:innen erfahren werden können, können sie durch die Permanenz ihres Hörens den Sinn der aufeinanderfolgenden Klänge im »Bewusstseinsstrom der inneren Zeit« (vgl. Schütz 2016 [1944], S. 101) als Musik erleben. Zugleich sind diese Klänge ein Ereignis in der »äußeren Welt« und die Voraussetzung der Kommunikation musikalischer Gedanken (Schütz 2016 [1951], S. 155). Dabei produzieren die Musiker:innen die einzelnen Klänge nicht willkürlich oder als individuelle akustische Ereignisse. Sie stehen vielmehr mit den Zeichen der notierten Musik in Verbindung, sodass sie als *Töne* produziert werden. Töne haben dadurch selbst zeichenhaften Charakter, weil sie auf die entsprechenden Anweisungen der Noten verweisen. Als Töne werden die Klänge auch von den Mitmusiker:innen gehört, sodass sich exemplarisch im »Transkript der rhythmischen Impulse« zeigen ließ, dass das Rhythmusproblem des Ensembles damit verbunden ist, dass Töne, die synchron produziert und gehört werden sollen, zeitlich auseinanderfielen. Und zwar derart, dass sich bereits während des Spiels eine mimische Kommunikation beobachten lässt, die zeigt, dass die Beteiligten dieses Problem gleich beim ersten Mal, noch während es sich in der Ausdehnung der Zeit entwickelte, hören konnten (während sich in diversen

Interpretationssitzungen zeigte, dass außenstehende Beobachter:innen das Problem mitunter überhaupt nicht akustisch erkennen konnten; s. o.). Insofern handelt es sich auch um eine beeindruckende Leistung des Ensembles, das Spiel abzubrechen.

7. Ein Nutzungsszenario: Die Videoanalyse des Chorsingens

Es stellt sich die Frage, ob und inwiefern sich das vorgestellte Untersuchungsinstrument auf andere Formen der Erzeugung von Kollektivgebilden wie gemeinsamer Musik übertragen lässt. Zur Auseinandersetzung mit dieser Frage berichte ich von den Herausforderungen und Chancen der Transferierbarkeit der interpretativen Videoanalyse auf Chorgesang, den ich derzeit gemeinsam mit Michaela Pfadenhauer untersuche. Dabei kann ich exemplarisch auf die Analyse des Singens eines Kirchenchores Bezug nehmen, den wir u. a. videographisch untersuchen. Die Musiziergattungen der Streichensembles und Chöre teilen einige typische Gemeinsamkeiten: Es werden bereits existierende, mehrstimmige Musikstücke realisiert, die (in den überwiegenden Fällen) als Noten objektiviert sind, während das gemeinsame Musizieren geprobt wird, bevor es zur Aufführung kommt. Zugleich finden sich charakteristische Unterschiede zum Streichensemblespiel, die die Anpassung der Arbeit am Videomaterial fordern – etwa: die musikalische Klangerzeugung beim (Chor-)Singen (1), die Verteilung der musikalischen Stimmen der Kompositionen auf die Sänger:innen (2), die Anzahl der Beteiligten (3) und die damit verbundene Heterogenität der Arbeitsteilung des Chorsingens (4).

1. *Die musikalische Klangerzeugung beim (Chor-)Singen:* Während (Streich-)Instrumentalist:innen durch Bewegungen mit und an ihren Instrumenten Töne hervorbringen, verwenden Chorsänger:innen ihre eigene menschliche Stimme. Sie ist keine von den Singenden losgelöste Objektivation – kein Instrument, das als Objekt ›hinzugezogen‹ wird. Der physische Stimmapparat wird quasi ›im‹ Körper der Singenden zum Schwingen gebracht. Vor diesem Hintergrund könnte die interpretative Interaktionsanalyse des (Chor-)Singens erschwert sein. Allerdings zeigt sich, dass die Musizierbewegungen des Singens mit dem Forschungsmedium Video relativ gut beobachtbar sind, wobei ihre Formen sich sichtbar z. B. vom Summen, Schreien, Rappen, Growling oder Sprechen unterscheidet. Da das Öffnen des Mundes für das Singen relevant ist, können die Bewegungen des Kiefers, der Lippen und damit die Veränderungen der Mundöffnungen, die die gesanglichen Klänge formen, relativ gut in ihrem kommunikativen Zusammenspiel analysiert werden. Dabei ist auch das für das (Chor-)Singen relevante Atmen gut erkennbar, das mitunter den gesamten Oberkörper der Beteiligten bewegt.

2. *Die Verteilung der musikalischen Stimmen der Kompositionen auf die Beteiligten:* Während in klassischen Streichensembles jede:r Instrumentalist:in jeweils eine Stimme zur Mehrstimmigkeit der Komposition beiträgt, ist die Verteilung der Stimmen im Chor eine charakteristisch andere – sie werden (Stimm-)Gruppen zugeordnet (siehe eine beispielhafte Rekonstruktion der Verteilung der musikalischen Stimmen des exemplarischen Kirchenchores in Abb. 5). Dabei ›sollen‹ die individuellen Stimmen der Beteiligten nicht mehr heraushörbar sein. Der Höreindruck des chorischen Klanges ist vielmehr mit der herausragenden Leistung verbunden, die einzelnen Stimmen zu einem mehrstimmigen Gesamtklang zu verbinden, in dem sie quasi zu einem kollektiven Klangkörper verschmelzen. Eine Herausforderung des Chorsingens besteht damit in der synchronen Realisation derselben musikalischen Stimme durch die Mitglieder einer Stimmgruppe, eine weitere in der harmonischen Koordination und Synchronisation dieser Gruppen zueinander. Vor diesem Hintergrund fokussieren wir uns in der Videoanalyse auch auf die beobachtbare Performanz der Stimmgruppen. Die audiovisuellen Daten erlauben dabei der Frage nachzugehen, wie sich im Probenprozess der »Partizipationsstatus der Beteiligten« (Knoblauch 2024, S. 280) verändert und wie die Gleichförmigkeit des Singens der Vielen erzeugt und koordiniert wird.

Abbildung 5: Anordnung der Singenden im Verhältnis zur Mehrstimmigkeit der gesungenen Komposition (eigene Darstellung)

3. *Die Anzahl der Beteiligten:* Als größere Herausforderung für die interpretative Videoanalyse erweist sich die Anzahl der Mitglieder eines Chores. Während in Streichensembles ›nur‹ etwa eine Handvoll Musiker:innen zusammenkommen, singen in Chören mitunter mehrere Dutzend Beteiligte. Das fordert bereits das Erstellen der Videos im ethnographischen Forschungsprozess heraus: Gelingt es, die kommunikativen Handlungen Aller registrieren zu können? Für die Untersuchung des Kirchenchores haben wir zu diesem

Zweck mit drei Kameras gearbeitet, die mehr oder weniger in der Raumtotale, verschiedene Blickachsen auf das kommunikative Geschehen erlauben (während für das Streichensemblespiel jeweils eine Kameraperspektive ausreichend war). Zudem ist es eine Frage der technischen Eigenschaften der Kameras, wie hoch die Auflösung der Aufnahmen ist: Kann ich derart in die Aufnahmen ›hineinzoomen‹, dass ich als Analytikerin die relativ kleinen Singbewegungen der Vielen rekonstruieren kann, um das geprobte Gelingen der Synchronizität des Singens auf einer hohen »Granularitätsstufe« (Knoblauch/Vollmer 2022, S. 360) – quasi besonders fein – vergleichen zu können? Zugleich rückt insbesondere die Leistung der Synchronisierung als Zusammenspiel Vieler in die Reichweite einer feingliedrigen Analyse.

4. *Die Heterogenität der Arbeitsteilung in Bezug auf das gemeinsame Musizieren:* Während sich Streichensembles durch die Abwesenheit eines Dirigats auszeichnen, ist eine musikalische Konstellation, die derart viele Menschen umfasst, wie ein Chor, zumeist auf eine Leitung angewiesen (vgl. Schütz 2016 [1951], S. 167). Sie wird im von uns untersuchten Kirchenchor durch eine Chorleiterin geleistet. Wenn wir Schütz folgen, müssen ihre dirigierenden Bewegungen alle »musikalischen Ereignisse der inneren Zeit« der vielen Singenden als Synthese »ersetzen« (ebd.). Dabei deutet sich eine doppelte Zeitlichkeit des Dirigierens an: Die Leiterin reagiert durch ihre Gesten in Musizierphasen auf die Klangobjektivierungen des Chores und führt ihn mit diesen Gesten gleichzeitig an. Dabei nimmt die Chorleiterin eine im doppelten Wortsinn herausgestellte, kommunikative Position im Chor ein: Zum einen hat sie ein kommunikatives Privileg bei der »Erfassung des musikalischen Gedankens des Komponisten« (vgl. ebd., S. 159), insofern sie durch ihre Leitung die musikalische Interpretation maßgeblich – etwa durch Ansprachen an den Chor – anleitet. Zum anderen nimmt sie eine besondere räumliche Position ein, in der sie den Singenden gegenüber auf einem Podest steht, die zu ihr mehr oder weniger frontal zu Stimmgruppen angeordnet ausgerichtet sind. Diese räumliche Anordnung[2] des Chorsingens (siehe exemplarisch Abb. 5) kann durch die Videodaten in die Reichweite einer feingliedrigen Analyse rücken (vgl. Knoblauch 2024, S. 284 ff.). – Eine Aufgabe, zu der Videodaten detailliertere Einblicke liefern können, wie sie selbst den daran Beteiligten nicht mehr rekonstruierbar sind (vgl. Knoblauch 2024, S. 275 f.).

Die in diesem Beitrag vorgestellten weiterentwickelten Methoden der Videointeraktionsanalyse werden für das Chorsingen weiter erprobt und angepasst. Sie sollen zur tiefergehenden Exploration der Frage beitragen, wie (klingende) Kollektivphänomene durch das Zusammenspiel individueller Handlungen geleistet

2 Zur Relevanz der räumlichen Anordnung von Streichensembles siehe Vollmer (2023, S. 303–341 et passim).

werden, sodass sie kollektive Formen annehmen (vgl. ebd., S. 285 f.), wie z. B. der Chorklang als Leistung organisierten, gemeinsamen Singens.

Die drei analytischen Stimmen der vorgestellten soziologischen Partitur »Wissen«, »Hör-« und »Klangwelt« korrespondieren dabei in der empirischen Wirklichkeit gemeinsamen Musizierens (Vollmer 2023, S. 475–490). Insbesondere die soziale Stimme des »Wissens« harrt weiterführender Untersuchungen und Differenzierung. Dennoch: Insbesondere die Verbindung zwischen der Hör- und Klangwelt bzw. die Verbindung der »inneren« und »äußeren Zeit« nach Schütz können durch die weiterentwickelten Begriffe z. B. der Sozialtheorie des kommunikativen Handelns nuancierter aufeinander bezogen werden. Dabei soll es *nicht* darum gehen, die empirischen Erkenntnisse der Untersuchung des Streichensemblespiels oder des Chorsingens anderen Formen der (z. B. nicht-westlich-abendländischen, improvisierten oder programmierten) Musik ›überzustülpen‹. Die Ausführungen sollen vielmehr dazu beitragen, die analytischen Elemente einer Sozialtheorie gemeinsamen Musizierens näher zu bestimmen und in Verbindung zu setzen, um zur Grundlage eines vertiefenden Dialogs verschiedenartiger (sozial-)wissenschaftlicher Erkenntnisse beitragen zu können. Die Weiterentwicklung der interpretativen Videoanalyse bietet dabei die Chance, Formen der Kommunikation in den Blick zu nehmen, die sich durch die Abwesenheit des Sprechens auszeichnen können, ohne dabei auch nur im Entferntesten (kommunikativ) still zu sein.

Literatur

Bergmann, Jörg (1985): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. In: Entzauberte Wissenschaft, Sonderband 3. Schwartz, S. 299–320.

Garfinkel, Harold (1984): Studies in ethnomethodology. Polity Press.

Goffman, Erving (2009 [1963]): Interaktion im öffentlichen Raum. Frankfurt am Main: Campus.

Honer, Anne (1989): Einige Probleme lebensweltlicher Ethnographie – Zur Methodologie und Methodik einer interpretativen Sozialforschung. Zeitschrift für Soziologie, 18(4), S. 297–312.

Knoblauch, Hubert (2001): Fokussierte Ethnographie. Sozialer Sinn. Neuere Entwicklungen in der Ethnographie, S. 123–42.

Knoblauch, Hubert (2017): Die kommunikative Konstruktion der Wirklichkeit. Wiesbaden: Springer VS.

Knoblauch, Hubert (2024): Efferveszenz: Videografische Analyse kommunikativer Formen kollektiver Emotionen. In: Diefenbach, Aletta/Zink, Veronika (Hrsg.): Emotions- und Affektsoziologie. Eine Einführung. Berlin/Boston: De Gruyter Oldenbourg, S. 273–288. https://doi.org/10.1515/9783110589214-024.

Knoblauch, Hubert/Vollmer, Theresa (2022): Fokussierte Ethnographie. In: Poferl, Angelika/Schröer, Norbert (Hrsg.): Handbuch Soziologische Ethnographie. Wiesbaden: Springer VS, S. 353–366.

Kress, Gunther (2010): Multimodality: A Social-Semiotic Approach to Contemporary Communication. Routledge.

Luff, Paul/Hindmarsh, Jon/Heath, Christian (Hrsg.) (2000): Workplace Studies. Recovering Work Practice and Informing System Design, Cambridge: Cambridge University Press.

Pfadenhauer, Michaela (2005): Ethnography of Scenes. Towards a Sociological Life-world Analysis of (Post-traditional) Comunity-building. FQS, 6(3), Art. 43. http://nbn-resolving.de/urn:nbn:de:0114-fqs0503430

Reichertz, Jo/Englert, Carina (2011): Einführung in die qualitative Videoanalyse: Eine hermeneutisch-wissenssoziologische Fallanalyse. Wiesbaden: Springer VS.
Sacks, Harvey (1992): Lectures on Conversation. Oxford: Blackwell.
Schubert, Franz (1890 [1812]). »Quartett (C Dur) für 2 Violinen, Viola und Vionloncell«. In Franz Schubert's Werke. Kritische durchgesehene Gesamtausgabe. Serie 5. Quartette für Streichinstrumente. Breitkopf und Härtel.
Schütz, Alfred (2016 [1944]): Fragmente zur Phänomenologie der Musik. In: Sebald, Gerd/Stascheit, Andreas (Hrsg.): Werkausgabe: ASW. Schriften zur Musik. Konstanz/München: UVK, S. 79–146.
Schütz, Alfred (2016 [1951]): Gemeinsam Musizieren. Eine Studie sozialer Beziehungen. In: Sebald, Gerd/Stascheit, Andreas (Hrsg.): Werkausgabe: ASW. Schriften zur Musik. Konstanz/München: UVK, S. 147–76.
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie: Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.
Vollmer, Theresa (2023): Gemeinsam Musizieren. Zur Soziologie und Ethnographie der Streichensembles. Wiesbaden: Springer VS. https://link.springer.com/book/10.1007/978-3-658-42162-5.

Erhebung – Analyse – Nachnutzung

Audio-/Videodaten im Kontext sprachlicher Begabungsförderung

Caterina Mempel und Jenny Winterscheid

Zusammenfassung: Am Beispiel von drei Teilprojekten der Bund-Länder-Initiative *Leistung macht Schule* werden eine begabungsgerechte Ausgestaltung des Fachunterrichts Deutsch sowie ein iteratives, partizipatives und dialogisches Vorgehen im Rahmen der Erhebung, Analyse und Nachnutzung von Audio- und/oder Videodaten in den Blick genommen. Neben Erkenntnissen zur Ausdifferenzierung sprachlicher Begabung konnten Einblicke in die Unterrichtspraxis und in schulische Bedarfe gewonnen werden, die Grundlage für die Entwicklung von Konzepten und Materialien sowie passgenauen Professionalisierungsmaßnahmen für Lehrpersonen waren/sind. Die Arbeit mit aufbereiteten Unterrichtsdaten in Form von Transkripten oder Text- und Videovignetten unterstützt den Dialog zwischen Forschung und pädagogischer Praxis.

Schlüsselwörter: sprachliche Begabungsförderung, partizipative Forschung, Gesprächsanalyse, Transkription, Vignettenarbeit

1. Einleitung

Optimale Lern- und Bildungschancen für Kinder und Jugendliche zu schaffen, ist ein zentrales Ziel der Bildungspolitik von Bund und Ländern. Je individueller die schulische Förderung auf die Bedürfnisse der einzelnen Schüler:innen abgestimmt ist, desto besser können Kinder und Jugendliche ihre Potenziale entfalten. In der 2018 gestarteten Bund-Länder-Initiative *Leistung macht Schule* (*LemaS*) geht es um eine Optimierung von Lern- und Bildungschancen und eine möglichst voraussetzungsarme Förderung leistungsstarker und potenziell leistungsfähiger Schüler:innen in Deutschland »– und zwar unabhängig von ihrer Herkunft, ihrem Geschlecht oder ihrem sozialen Status« (vgl. z. B. Leistung macht Schule o. J.[a]). Der Forschungsverbund *LemaS* zielt mit seinem Verbundprojekt auf eine Haltungsänderung der Akteur:innen und die theorie- und evidenzbasierte Optimierung begabungs- und leistungsfördernder Schul- und Unterrichtsgestaltung durch wissenschaftliche Beratung, Weiterqualifizierungen und differenzierend einsetzbares Material. Eng mit diesen Zielen verknüpft ist eine partizipative Ausrichtung, wobei Kooperation und Dialog zwischen Forschung und pädagogischer

Praxis einen besonderen Stellenwert einnehmen. Die Transferphase (2023–2027) strebt die Vertiefung und Verbreitung der Ergebnisse aus der ersten LemaS-phase an (LemaS Forschungsverbund o. J.[a]). Ähnlich wie in der ersten Phase begleitet der Forschungsverbund *LemaS-Transfer* die Schulen mit vielfältigen Professionalisierungs- und Weiterbildungsangeboten und beforscht zugleich die Transfer- und Implementationsprozesse.

Im Beitrag wird am Beispiel dreier in der Deutschdidaktik angesiedelten Teilprojekte die Arbeit mit Audio-/Videomaterial im Rahmen der Erhebung, Analyse und Nachnutzung in der ersten Phase sowie der Transferphase näher beleuchtet und eine begabungsgerechte Ausgestaltung des Fachunterrichts in den Blick genommen.

2. Verortung in *Leistung macht Schule*

Die drei Teilprojekte des Deutsch-Clusters »literarisches Schreiben« (vgl. z. B. Laudenberg/Neuweiler 2020), »LemaS-GRiP – Förderung des sprachlich-ästhetischen Gesprächs im Regelunterricht im Fach Deutsch in der Primarstufe« (vgl. z. B. Mayer/Mempel 2020) und »sprachlich-rhetorische Kommunikation« (vgl. z. B. Spiegel/Winterscheid 2020) entstanden aus dem Bedarf an konkreten konzeptionellen Orientierungen, adäquaten Materialien und empirischen Evidenzen im Bereich der sprachlich-rhetorischen und literarischen Förderung von leistungsstarken und potenziell leistungsfähigen Schüler:innen im Fach Deutsch. Dies ist besonders relevant, wenn neben traditionell segregierenden Maßnahmen auch die binnendifferenzierte Unterstützung möglichst aller Schüler:innen im gemeinsamen Unterricht berücksichtigt werden soll (vgl. z. B. Preckel/Vock 2013, S. 171 ff.; Mayer 2018).

Erhebungsinstrument in allen drei Teilprojekten waren Audio- und/oder Videomaterial, welches in der Schule erhoben wurde, um sprachliche Begabungen sowie Erfahrungs- und Lernprozesse zu erkennen, fachinhaltlich zu beurteilen und daraus didaktische Modellierungen angemessen zu planen und umzusetzen. Die erhobenen, aufbereiteten und analysierten Daten wurden – neben den wissenschaftlichen Ergebnissen der Schul- und Unterrichtsforschung – für die Gestaltung von Fort- und Weiterbildungen für Lehrpersonen sowie für die Entwicklung von Unterrichts- und Professionalisierungsmaterialien herangezogen.

Gemeinsam mit Projektschulen wurden in der ersten *LemaS*-Phase *P³rodukte*[1] entwickelt, erprobt und verfeinert, um Lehrpersonen dabei zu unterstützen,

1 *LemaS-P³rodukte* sind Strategien, Konzepte, Maßnahmen und Materialien zur Förderung (potenziell) leistungsstarker Schüler:innen und dienen der begabungs- und leistungsfördernden Schul- und Unterrichtsentwicklung (z. B. LemaS Forschungsverbund o. J.[b]).

ihren Unterricht so zu gestalten, dass sie die sprachlichen und literarischen Potenziale möglichst aller Schüler:innen erkennen und gezielt im Unterricht fördern können. Diese wurden im intensiven Austausch mit den Lehrpersonen implementiert, auf ihre Einsetzbarkeit und Wirksamkeit hin überprüft, immer wieder auch angepasst und sollen in der Transferphase nun in neue Schulen transferiert werden.

3. Erhebungskontext

Praxeologische Ansätze in der Unterrichtsforschung erfordern die detaillierte Beobachtung und Rekonstruktion von alltäglichen Unterrichtspraktiken. Routinen und Praktiken sind oft so tief im Alltagswissen verwurzelt, dass sie von den Beteiligten nicht einfach in Interviews oder Metagesprächen reflektiert/thematisiert werden können. Gleichzeitig werden immer noch viele Gesprächsmythen (vgl. z. B. Spiegel 2013, S. 303–307) weitergegeben und »in den Lehrwerken [lässt sich weiterhin] ein sogenannter *written-bias*, d. h. eine starke Orientierung an der Schriftlichkeit« (Grundler 2022, S. 182; Herv. i. O.) ausmachen. Zur Bewertung der fachdidaktischen Qualität des Unterrichts und der Vermittlung ist es demnach entscheidend, direkt in der Unterrichtsinteraktion zu beobachten, welche Indikatoren für Verstehens- oder Verständigungsprozesse vorhanden sind und wie und welche Interaktionen während des Unterrichts stattfinden. Um Ansatzpunkte für die Förderung zu erkennen und die *LemaS-P³rodukte* auf die Erfordernisse und Wünsche von Lehrpersonen und Schüler:innen anzupassen, wurden in den Teilprojekten des Deutsch-Clusters daher u. a. unterschiedliche Audio-/Videodaten, wie z. B. Interviews mit Lehrpersonen und Schüler:innen oder Unterrichtsinteraktionen, erhoben.

3.1 Feldzugang

Nach der Beantragung einer offiziellen Genehmigung für die Forschungsvorhaben bei den zuständigen Ministerien wurden die Schulen, welche nach einem Matchingverfahren den Projekten zugeordnet wurden, informiert. Neben schriftlichen Informationen für Eltern und Schüler:innen, Lehrpersonen und Schulleitungen wurde auch der persönliche Austausch mit allen Interessensgruppen gesucht und das Einverständnis für unterschiedliche Datenerhebungs- und Aufbereitungsformen und vereinzelt auch die Möglichkeiten der (Nach-)Nutzung eingeholt. Der organisatorische Aufwand ist beträchtlich, jedoch aufgrund der sensiblen personenbezogenen Daten, die erhoben wurden, im Sinne von Art. 4 Nr. 1 Datenschutz-Grundverordnung (DSGVO) gerechtfertigt.

3.2 Technische und organisatorische Aspekte der Datenaufzeichnung

Die Erhebungen fanden an den Teilprojekten zugeordneten Schulen durch die Projekt-Mitarbeitenden oder vereinzelt auch durch Lehrpersonen selbst statt.[2] Die Einzel- oder Gruppeninterviews mit Schüler:innen oder Lehrpersonen wurden mit Audioaufnahmegeräten aufgezeichnet, welche zwischen den Interagierenden positioniert worden waren. Für die Datenerhebung des Unterrichts kamen üblicherweise zwei statische Kameras zum Einsatz, die so platziert wurden, dass möglichst viele Beteiligte erfasst werden konnten. Eine diagonale Aufstellung ermöglichte es, sowohl die Lehrperson als auch die Schüler:innen von vorne zu filmen. Um flexibel auf unerwartete Unterrichtsverläufe reagieren zu können, wurde teilweise eine mobile Kamera bereitgehalten. Zusätzlich zeichnete ein zentral aufgestelltes Audioaufnahmegerät die Unterrichtskommunikation auf, um die Audioaufnahme gegebenenfalls unter die Videoaufnahmen zu legen, da die Videokameras räumlich etwas entfernt vom Unterrichtsgeschehen standen. In einem Feldtagebuch bzw. Beobachtungsprotokoll konnten Sitzpläne, der Verlauf der Unterrichtsstunden und relevante Beobachtungen dokumentiert werden, was die Auswertung und das Auffinden bestimmter Sequenzen erleichtern sollte. Um das Geschehen im Unterricht zu rekonstruieren und bei der Analyse den Einbezug von Erarbeitetem oder Unterrichtsmaterialien berücksichtigen zu können, wurden außerdem Unterrichtsplanungen sowie Tafelbilder, Handreichungen, Schüler:innenprodukte etc. erfasst, die im Unterricht verwendet oder erstellt worden waren.

3.3 Datenaufbereitung

Nach der Erhebung sichteten die Mitarbeitenden die Daten und bearbeiteten diese technisch. Aufgrund der Fülle an gewonnen Datenmaterial wurden für die Analyse nur relevante Ausschnitte verschriftet. Die Transkription der identifizierten Sequenzen erfolgte nach cGAT-Konventionen (Schmidt et al. 2023), einer an die Arbeit mit Transkriptionseditoren angepassten Version der GAT 2-Konventionen (Selting et al. 2009). Mit diesem Vorgang ging auch die Pseudonymisierung der Daten einher.

Im Teilprojekt *LemaS-GRiP* wurden ferner ausgewählte Fotogramme in das Transkript aufgenommen, um die Bedeutung der körperlichen und materiellen Ebene für die Interaktion detailliert zu veranschaulichen und die im Unterricht als komplexe Simultanstruktur (vgl. Wagner-Willi 2004) vorzufindende Vielschichtigkeit zu beschreiben (siehe Abb. 1, vgl. auch Mempel 2023, Mempel/Mayer 2021).

2 Im Teilprojekt *LemaS-GRiP* wurde zusätzlich den Projekt-Lehrpersonen eine technische Grundausstattung für den Projektzeitraum bereitgestellt, um sie als Co-Forschende in den Prozess der Datenerhebung einzubeziehen (vgl. Mayer/Mempel 2022).

SU01 <<all> oder er MACHT da jetzt einfach so>
((zu SU02 sprechend, streicht mit rechter Hand
über linke Schulter und beißt die Zähne
zusammen))

Abbildung 1: Screenshot Transkriptauszug (Mempel 2024)

4. Analysekontext

In diesem Abschnitt werden die methodologischen Ansätze und die konkrete Arbeit der ersten *LemaS*-Phase näher beschrieben.

4.1 Wissenschaftliche Analysen

4.1.1 LemaS-GRiP

Im Teilprojekt *LemaS-GRiP* orientieren sich die Analysen an den Prinzipien der sequenziellen Interaktionsanalyse im Sinne der dokumentarischen Methode (vgl. Asbrand/Martens 2018), die bei allen Interpretationsschritten die verbale und nonverbale bzw. visuelle Ebene des videografierten Datenmaterials berücksichtigte (ausführlich dazu Mempel 2023). Interviews und Protokolle wurden triangulierend miteinbezogen. Der Einzelfall diente hier der Problemerschließung und Analyse von Strukturen, Deutungsmustern oder Praktiken. Die Rekonstruktion der videografierten Unterrichtsstunden konnte zeigen, dass eine gesprächsförmige Begegnung mit einem anspruchsvollen literarischen Text wesentlich zur Kompetenzentwicklung beiträgt, indem die Mehrdeutigkeit des Textes interpretierende Sinnentwürfe von den Rezipierenden fordert, die es kollektiv auszuhandeln gilt (vgl. z. B. Mayer/Mempel 2022; Mempel 2022, 2023; Mempel/Mayer 2021). Dazu gehört auch, Erfahrungen mit der Vielstimmigkeit literarischer Texte zu machen und zu lernen, Momente des Nicht-Verstehens, der Irritation und Ambivalenz auszuhalten und als Ausgangspunkt für eine erneute Deutungssuche zu nutzen. Ebenso lässt sich eine wachsende Verantwortung aller Mitwirkenden für ein gemeinsam getragenes Gespräch – das durch Achtsamkeit für den Text und das Gegenüber geprägt ist – über den Projektzeitraum der ersten Phase beobachten, wenn literarische Gespräche regelmäßig im Unterricht umgesetzt wurden. Die Diversität der Gruppe kann hier positiv als »Chance« (Bräu/Schwerdt 2005, S. 9) für individuelle und gemeinsame Lernprozesse gewertet und als Ressource anerkannt werden, wobei die individuellen Fähigkeiten, Relevanzen und Interessen gleichberechtigt nebeneinander positioniert sind (vgl. »Pädagogik der Vielfalt« nach Prengel 1993).

4.1.2 Literarisches Schreiben und sprachlich-rhetorische Kommunikation

Zusätzlich zu den erhobenen Unterrichtsinteraktionen an den Kooperationsschulen und/oder den aufgezeichneten Interviews wurden bei den Teilprojekten zum *literarischen Schreiben* und zur *sprachlich-rhetorischen Kommunikation* auch Identifikatoraufgaben, Schüler:innenprodukte, Fragebögen sowie Daten eines Vorgängerprojekts zu Hochbegabten (Korpus Spiegel) einbezogen (vgl. z. B. Laudenberg et al. 2022).

Hintergrund der Arbeit im Teilprojekt zur *sprachlich-rhetorischen Kommunikation* bildeten gesprächsanalytische Erkenntnisse zur Schul- und Unterrichtsforschung und die gesprächsanalytische Auswertung des Audio- und Videomaterials (vgl. z. B. Deppermann 2008). Die Gesprächsanalyse beschäftigt sich mit der Frage, »nach welchen Prinzipien und mit welchem sprachlichen und anderen kommunikativen Ressourcen Menschen ihren Austausch gestalten und dabei die Wirklichkeit, in der sie leben, herstellen« (ebd., S. 9). Zu den Grundprinzipien gehört u. a. die Arbeit mit Audio- und/oder Videomaterial und deren Verschriftlichung (Transkription) nach bestimmten Konventionen (s. o.). Das Vorgehen ist dabei deskriptiv, induktiv, d. h. »from the data themselves« (Schegloff/Sacks 1973, S. 291) entwickelt, und erfolgt sequenziell (Sacks/Schegloff/Jefferson 1974). Bei der gesprächsanalytischen Auswertung der Daten wurde zum einen berücksichtigt, wie Schüler:innen die jeweilige Aufgabe zum Argumentieren, Präsentieren oder Feedbackgeben sprachlich, parasprachlich und körpersprachlich bearbeiteten, weil diese »bildungssprachlichen Praktiken« (vgl. Morek/Heller 2012) nicht nur im schulischen, sondern auch im Alltag und später im Beruf von entscheidender Bedeutung sind. Zum anderen wurde untersucht, was die Performance von Sprachbegabten im Besonderen auszeichnet: Dies geschah zunächst anhand von Einzelfallanalysen (vgl. z. B. Spiegel 2019; Winterscheid 2019; Spiegel 2018), die Ergebnisse wurden anschließend jedoch auch über die verschiedenen Daten hinweg – fallübergreifend – betrachtet (Deppermann 2008, S. 49–103).

Beim Vergleich der Ergebnisse der Untersuchung der schulischen Performances in den beiden Teilprojekten zum *literarischen Schreiben* und zur *sprachlich-rhetorischen Kommunikation* fanden sich im Hinblick auf die Ausdifferenzierung von sprachlicher Begabung, obgleich sich die Anforderungen an sprachlich-rhetorische Kommunikationssituationen und das literarische Schreiben in wesentlichen Punkten unterscheiden, dennoch einige Aspekte, die in beiden Bereichen gleichermaßen herausgestellt wurden (vgl. z. B. Winterscheid/Sellinger 2023, Laudenberg et al. 2022; Laudenberg/Spiegel 2020; LemaS Forschungsverbund o. J.[b]). Die daraus abgeleiteten Kriterienkataloge bildeten die Grundlage der Ausgestaltung der *LemaS-P³rodukte* und stellten Anknüpfungspunkte für die Förderung an den Schulen dar. Darüber hinaus bieten sie eine Orientierung bei den Rückmeldungen an die Schüler:innen und der Auswahl der weiteren Angebote.

So kann beispielsweise eine Schülerin, die kreativ mit Worten spielen kann, aber sich mit Handlungssträngen schwertut, entsprechende Aufgaben aus dem Bausteinkasten zum *literarischen Schreiben* bearbeiten und in der Auseinandersetzung mit einem literarischen Vorbild des entsprechenden Bausteins ihre Fähigkeiten ausbauen. Viele spannende Anregungen finden die Schüler:innen zudem in den dort vorhandenen Videoclips namhafter Autor:innen, die sich zu bestimmten Aspekten des Schreibens und des Schreibprozesses äußern und Tipps geben (in diesem Fall etwa einen Videoclip von Katharina Herzog zum Spannungsaufbau). Eine Lehrerin reflektierte nach dem Einsatz eines solchen Videos etwa: »Wir haben kürzlich das Video von Andrea Karimé zum lautmalerischen Schreiben eingesetzt, und der Funke ist sofort übergesprungen« (Leistung macht Schule o. J.[b]).

Ein Schüler, der sehr pointiert Sachverhalte darstellen kann, aber dessen Performance noch nicht so ausgeprägt ist, kann dafür das Spiel mit der Stimme und den Einsatz von körpersprachlichen Mitteln konkret mithilfe von Übungen aus der Aufgabensammlung zur *sprachlich-rhetorischen Kommunikation* (in den Heften oder auch in den WebQuests; zu den in der ersten LemaS-Phase entstandenen WebQuests s. z. B. Winterscheid 2024) trainieren, indem er aufgabengeleitet an authentischen Beispielen herausarbeitet, welche sprachlichen Mittel begabte Redner:innen wie und wann einsetzen (epistemisches Wissen), in praktischen Übungen Gelerntes ausprobiert (handlungspraktisches Wissen), dann mit anderen im Austausch, die sich ebenfalls mit diesem Aspekt beschäftigt haben, Rückmeldungen erhält und selbst Feedback gibt (Reflexion) (vgl. z. B. Mayer et al. 2024, S. 413 ff.).

4.2 Analytisches Arbeiten in der Praxis am Beispiel literarisches Schreiben und sprachlich-rhetorische Kommunikation

Die Analysen in den beiden Teilbereichen zum *literarischen Schreiben* und der *sprachlich-rhetorischen Kommunikation* bildeten die Grundlage für verschiedene *P³rodukte*; die Lehrenden und Lernenden sollten aber im Rahmen des Projekts zugleich dazu befähigt werden, selbst analytisch vorzugehen: Die kooperierenden Lehrpersonen im Teilprojekt *sprachlich-rhetorische Kommunikation* erhielten in den Fortbildungen theoretischen Input, die Möglichkeit, sich vorab mit den Aufgabenformaten auseinanderzusetzen, sowie Tipps für den Einsatz im Unterricht. Gleichzeitig wurden neben den Erkenntnissen der gesprächsanalytischen Forschung auch deren Herangehensweise und Prinzipien, wie etwa eine deskriptive Haltung bei der Betrachtung von Kommunikationssituationen einzunehmen (vgl. z. B. Deppermann 2008, S. 106), vermittelt. Für eine adäquate Förderung der Schüler:innen gehört zudem die Schulung des Blicks auf Kommunikation.[3] Audio- und Videodaten[4] sowie die Arbeit mit Verschriftungen/Transkripten und

3 Warum das wichtig ist, siehe u. a. auch Grundler (2022); Becker-Mrotzek (2008) etc.

4 Es handelt sich jedoch nicht um erhobene Aufnahmen im Projektkontext, sondern ausgewählte frei verfügbare Videodateien. Geeignete Videobeispiele authentischer Interaktionen

bisweilen der eigenen Verschriftung/Transkription (siehe bspw. Abb. 2), – analog zur Arbeitsweise innerhalb der Gesprächsanalyse – bilden überdies einen festen Bestandteil der Aufgaben(-Formate) im Teilprojekt und gehen zumeist mit konkreten Beobachtungsaufträgen und auch dem Einsatz von auf den Kriterienkatalogen basierenden Beobachtungsbögen (vgl. hierzu auch Winterscheid 2024) einher.

Abbildung 2: Screenshots aus den WebQuests zum Argumentieren (Niveaustufen 1 und 2)

Man kann diese Annäherung an ein gesprächsanalytisches Arbeiten im Unterricht entsprechend zu Schwarzes Einordnung als propädeutisches Vorgehen (vgl. Schwarze 2023) umschreiben. Gerade vor der Arbeit mit Transkripten und dem Transkribieren schreckten anfangs noch einige Lehrpersonen zurück. Von den Kooperationspartner:innen, die die neuen Aufgabenformate mit dieser Aufgabe des Verschriftens dennoch implementierten, erfolgten unterschiedliche Rückmeldungen: Manche Schüler:innen ächzten wohl wegen der doch recht aufwändigen Arbeit, bei den anderen überwog die Neugierde, sich auf eine bisher unbekannte

finden sich bspw. in für Lehre und Wissenschaft zusammengestellten und öffentlich gemachten Korpora, in Mediatheken, auf YouTube etc.

Art und Weise Interaktionen zu nähern, sowie die Freude über den Erkenntnisgewinn, der sich mit der Arbeit einstellte. Insgesamt konnten die Lehrpersonen beobachten, dass z. B. bei der Bearbeitung der WebQuest zum Argumentieren, bei der sowohl mit Transkripten gearbeitet wird, aber auch selbst verschriftet/transkribiert (abhängig vom Niveau der WebQuest) werden sollte, »die Mehrheit der Klasse begeistert mitgemacht« und eine Bearbeitung stattgefunden hatte, welche »mit konventionellem Material oder mit traditionellen Unterrichtsmethoden so nicht passiert [wäre]« (Leistung macht Schule o. J.[c]).

Auch beim *literarischen Schreiben* setzen sich die Schüler:innen mit einem kurzen Ausschnitt eines Textes, einem Gedicht oder einem Autorenvideo auseinander und produzieren danach, basierend auf dem Vorbild/Input und angeregt durch eine sehr weit gefasste Aufgabe, ein eigenes Produkt (vgl. z. B. Laudenberg 2024; Laudenberg/Neuweiler 2020). Dabei ist die Analyse des literarischen Ausschnitts dem Prozess der Adaption quasi inhärent, insofern die Begabung in der Ausarbeitung »sichtbar [wird] und durch Vermittlung handwerklicher Fähigkeiten« (Abraham 2017, S. 82) gefördert werden kann.

Die Schüler:innen werden in beiden Teilprojekten dazu angeregt, eigenständig analytisch tätig zu werden, sich mit literarischen Texten, verschiedenen Textsorten, Audio- und Videodaten sowie deren Verschriftlichungen/Transkripten, bildungssprachlichen Praktiken, verschiedenen (rhetorischen) Mitteln und dem Wesen von Lyrik, Literatur und Kommunikation auseinanderzusetzen, und schließlich das neu erworbene Wissen in der Performance zu entfalten. Wenn dieser Prozess in Auseinandersetzung mit der literarischen Vorlage und/oder in Kleingruppen geschieht, entzieht er sich zunächst einmal der Kontrolle der Lehrperson, denn »[i]m Klassenverband oder im Einzel-Feedback lässt sich überprüfen, welche formalen und inhaltlichen Aspekte der Textvorlage – womöglich auf originelle Weise – berücksichtig sind und inwiefern dies zu einem gelungenen Gesamterlebnis beiträgt bzw. an welchen Stellen noch Überarbeitungsbedarf besteht« (Laudenberg 2024, S. 37). In solchen Lehr-Lern-Kontexten ändert sich die Rolle der Lehrperson, die während der Erarbeitung »im Hintergrund« ist und den Jugendlichen begleitend zur Seite [steht]« (vgl. auch z. B. Leistung macht Schule o. J.[c]).

Die entwickelten, erprobten und zum Teil mehrfach überarbeiteten Aufgaben(-Formate) finden sich in den Bausteinkästen/Arbeitsheften und können in dieser Form – oder nur in Teilen – an den *LemaS*-Schulen eingesetzt bzw. an die Schüler:innen herausgegeben werden. Solche Aufgaben können allerdings auch analog zu den in dem Teilprojekt kreierten Konzepten und Methoden adaptiert und ganz individuell eingesetzt werden. Dazu gehört, dass »Lehrpersonen sich in der ›Kreativität des Findens‹ (Roussel 2012) üben und bereits bei der Lektüre auf Schlüsselstellen achten, die in sprachlicher und stilistischer respektive literar-ästhetischer Hinsicht als Muster dienen können« (Mayer et al. 2024, S. 411) – dasselbe gilt für spannende Reden, Argumentationen etc.

Gerade in Bezug auf gesprochensprachliche Kompetenzen sei darüber hinaus empfohlen, »außerhalb« der Aufgabenformate den Schüler:innen immer wieder Aufträge zu übertragen, wie Erklärаufgaben, »kriterienbasierte Zuhör- und Beobachtungsaufgaben« und viele mehr, mit denen ebenfalls Fähigkeiten aller Schüler:innen kontinuierlich geschult werden können (vgl. ebd., S. 414 ff.).

5. Nutzungsszenarien

Während die erste Phase der Entwicklung, Implementation und Evaluation der *P³rodukte* diente, sollen diese in der zweiten Phase der Bund-Länder Initiative »LemaS-Transfer in die Schullandschaft« (LemaS Forschungsverbund o. J.[a]) weiter Verbreitung finden. Lehrpersonen, die bereits in der ersten Phase mit dem Forschungsverbund zusammengearbeitet haben, werden in der Transferphase zu Multiplikator:innen fortgebildet. In Schulnetzwerken treffen sie auf neue Schulvertreter:innen, denen sie die Anwendung der *LemaS-P³rodukte* näherbringen und mit denen sie ihre Erfahrungen zur Förderung von Potenzialen und Stärken der Schüler:innen teilen. Der Forschungsverbund *LemaS-Transfer* begleitet die Multiplikator:innen und Transferschulen bei dieser Aufgabe mit Professionalisierungs- und Qualifizierungsangeboten.

5.1 Regelhaftigkeit und Komplexität unterrichtlicher Wirklichkeit

Professionelle Unterrichtswahrnehmung bezieht sich auf die situationsspezifische Fähigkeit von Lehrpersonen, die komplexen Interaktionen und Dynamiken, die im Unterricht stattfinden, zu bemerken (*noticing*), diese wissensgestützt zu beschreiben, zu erklären, davon wissensbasierte Schlussfolgerungen abzuleiten (*knowledge-based reasoning*) (vgl. van Es/Sherin 2008) und den Unterricht entsprechend der Bedürfnisse ihrer Schüler:innen zu gestalten. In der Lehrer:innen(weiter)bildung werden diese Konzepte genutzt, um (angehende) Lehrpersonen auf die reflexiven Praktiken vorzubereiten, die notwendig sind, um im Unterricht wirksam zu sein. Durch gezieltes Training in *Noticing* und *Reasoning* können Lehrpersonen lernen, wie sie ihre Beobachtungen strukturiert analysieren und dabei evidenzbasierte Entscheidungen treffen können. Der Ansatz des »Lernens am Fall« (Crapa et al. 2021) bietet dabei die Möglichkeit, sich mit dieser komplexen Realität auseinanderzusetzen, die eigene Praxis zu hinterfragen, neue Perspektiven einzunehmen und Handlungsmöglichkeiten zu diskutieren. Der professionstheoretisch fundierte Ansatz unterstützt (auch angehende) Lehrer:innen darin, domänenspezifische Anforderungen zu verstehen und ihre Erfahrungen theoretisch zu reflektieren, wodurch Lernmomente sichtbar werden.

Ein Fall entsteht nach diesem Ansatz erst dann, wenn etwas beobachtet, dokumentiert oder ausgewählt und damit zum Fall gemacht wird (vgl. Steiner 2004,

S. 14). Der Fall ist also ein – authentischer oder konstruierter – Ausschnitt der – eigenen oder fremden – unterrichtlichen Praxis, der mündlich, schriftlich oder filmisch vorgelegt werden kann: z. B. als beschriebene Beobachtung, Textprotokoll, dichte Beschreibungen, als Interaktionssequenz in Form eines Transkripts, als Text- oder Videovignette. Bei der Analyse wird nur ein bestimmter Moment unterrichtlicher Wirklichkeit herausgegriffen, wodurch »eine zweckgerichtete, handlungsentlastende und verlangsamte Beschäftigung (Crapa et al. 2021, S. 23) mit unterrichtlicher Praxis einhergeht, d. h. ein Innehalten, was im Alltag so nur selten möglich ist.

Am Beispiel des Teilprojekts *LemaS-GRiP* zeigen wir im Folgenden, wie die Videodaten aus der ersten Phase als schriftliche und filmische Fallvignetten in der Transferphase für die Professionalisierung der Multiplikator:innen in der zweiten LemaS-Phase genutzt werden. Die aufbereiteten Vignetten veranschaulichen zum einen Schlüsselstellen eines gesprächsförmigen Literaturunterrichts und dienen zum anderen dazu, Herausforderungen bei der Umsetzung der literarischen Gesprächsformate im Unterricht transparent und wahrnehmbar zu machen.

5.2 Kriterien für die Auswahl in *LemaS-GRiP*: Problemhaltige Unterrichtssituationen

Die Text- und Videovignetten für die Transferphase stammen aus einem Fundus von Unterrichtsanalysen im Teilprojekt aus der ersten Phase von *LemaS* (vgl. Abschnitt 4.1.1), die auf der Basis des »Rahmenmodells zur Systematisierung typischer Problemsituationen« nach Heins (2018) geskriptet wurden. Nach Heins besteht die Relevanz des Modells für hochschuldidaktische Zusammenhänge darin, »eine Basis für die gezielte Auswahl von […] Unterrichtssituationen zu sein« (ebd., S. 29). Diese Situationen ermöglichen es, Strukturen zu rekonstruieren, die charakteristisch für den Literaturunterricht sind, und fördern die Entwicklung professioneller Wahrnehmung, Interpretation und Entscheidungsfindung im Handlungskontext (ebd.). Nach diesem Ansatz wird eine Problemsituation als eine solche definiert, die aus literaturdidaktischer Sicht hinsichtlich der normativen Angemessenheit des Handelns von Lehrpersonen je nach Fokus der Betrachtung besonders umstritten interpretiert und bewertet werden kann (ebd.).

Das von Heins entwickelte Rahmenmodell umfasst sechs Problemfelder, zu denen sich die problematischen Unterrichtssituationen bündeln lassen (ebd.):

- Problemfeld 1 – Probleme der Phasierung
- Problemfeld 2 – Probleme der Gegenstandsauswahl und -bewertung
- Problemfeld 3 – Probleme der Verstehensziele
- Problemfeld 4 – Probleme im Interpretationsprozess
- Problemfeld 5 – Probleme der Gegenstandsangemessenheit
- Problemfeld 6 – Probleme der Methoden – Funktionalität.

In einem ergänzenden Modell werden die Problemfelder auf zugrundeliegende Spannungsverhältnisse zwischen mindestens zwei der stets fachspezifisch ausgeprägten Unterrichtskomponenten (a) Lehr-Lernziel, (b) Gegenstand, (c) Disposition der Lernenden und (d) Lehr-Lernmethoden zurückgeführt (vgl. ebd., S. 37 f.).

Mithilfe des Modells wurde die Fülle an videografierten Situationen der ersten Phase daraufhin befragt, welchem Problemfeld sie zugeordnet werden können und welches Spannungsverhältnis sich in ihnen ausdrückt, und Video- und Textvignetten für die Arbeit mit den Multiplikator:innen erstellt.

5.3 Erstellung von Text- und Videovignetten in *LemaS-GRiP*

Die Vignetten sollen ein möglichst geringes Vorwissen und allenfalls theoretisch abgesicherte Vorkenntnisse erfordern und mehr oder weniger stark explizierte bzw. eher leicht oder schwer auffindbare Hinweise auf die jeweils relevanten Dimensionen eines literaturdidaktisch fundierten Problems beinhalten. Die für die Qualifizierung der Lehrpersonen entwickelten Vignetten bestehen aus einer kurzen Kontextbeschreibung (u. a. zu Klassenstufe, Gruppengröße, Unterrichtsgegenstand, Art der Textbegegnung), konkretisiert durch einen literarischen Textauszug oder eine kurze Inhaltszusammenfassung sowie einem angepassten Transkriptauszug bzw. eine Videosequenz (siehe Abb. 3).

Abbildung 3: Screenshot Videovignette eingebunden in LernBar-Kurs »Vorlesegespräche planen« (Mempel 2024)

Während die Betrachtung der Videovignetten durch einen Beobachtungsauftrag unterstützt wird (siehe Abb. 3), werden die Textvignetten durch einen konkreten Bearbeitungsimpuls gerahmt, welcher gezielt ein bestimmtes Problemfeld in den Blick nimmt und zentrale Analyseschritte direkt in den Bearbeitungsimpuls aufführt (siehe Abb. 4).

M1 Arbeitsauftrag/Impulse

- Welche der Schüler:innenäußerungen stufe ich als textlich fundierte Deutungen ein und welche als textferne Spekulationen?
- Welchen Einfluss haben meines Erachtens verschiedene Unterrichtsfaktoren (Schüler:innen/Klasse, Gegenstand, Lehr-Lern-Ziele, Methoden/Gesprächsformat und Lehrperson) auf die Äußerungen der Kinder?
- Mit welchen Impulsen hätte die Lehrperson intervenieren können, damit die Kinder textferne Spekulationen selbst erkennen?
- Wie beurteile ich das Verhältnis von textnahen Deutungen und textfernen Spekulationen vor dem Hintergrund meiner bisherigen Erfahrungen zum Gesprächsformat Vorlesegespräch?
- Welche Konsequenzen ziehe ich hieraus für meine zukünftigen Vorlesegespräche?

Abbildung 4: Beispiel Arbeitsauftrag zum Problemfeld 4 – Probleme im Interpretationsprozess (Mempel 2024)

Eingebunden sind die Vignetten in interaktive Selbstlernmodule, in sogenannte LernBar-Kurse,[5] innerhalb eines Moodle-Kurses auf der Angebotsplattform von LemaS.

5.4 Auseinandersetzung mit Vignetten im Rahmen der Qualifizierung in *LemaS-Transfer*

In der Auseinandersetzung mit Vignetten können Problemfelder und das zugrundeliegende Spannungsverhältnis (vgl. Abschnitt 5.2) situationsspezifisch ausgeleuchtet werden. Aus fachdidaktischer Perspektive lässt sich das in den Vignetten abgebildete Lehrpersonenhandeln auf der Grundlage des in den LernBar-Kursen (bereitgestellten) theoretischen Wissens in den Blick nehmen und interpretieren. Allein oder in der Fallbesprechung im Team können Handlungsalternativen entwickelt und begründet werden. Die theoretische Einordnung und Diskussion der Problemsituation als Ausdruck eines bestimmten Spannungsverhältnisses ermöglicht es, andere vergleichbare Problemsituationen oder auch

5 Das LernBar-Autorensystem ist ein Programm zur Erstellung interaktiver Lerneinheiten und wird von studiumdigitale an der Goethe Universität entwickelt (vgl. https://lehre-virtuell.uni-frankfurt.de/knowhow/lernbar/).

inspirierende Unterrichtssequenzen zu betrachten und miteinander zu vergleichen (vgl. Heins 2018). Die Fallarbeit in der Transferphase zielt darauf ab, theoretisches Wissen zu vermitteln bzw. zu festigen, dessen praktische Anwendung zu fördern und eine reflexive Auseinandersetzung mit der Praxis zu ermöglichen (vgl. Heinzel 2021). In individuellen und gemeinsamen Vignettenlektüren (vgl. Agostini/Anderegg 2023, S. 27) in asynchronen Selbstlernkursen oder synchronen Blended Learning-Kursen wird an konkreten Momenten angesetzt. Dies ermöglicht zum einen ein intuitives Verständnis von Erfahrungssituationen, das oft in der eigenen Unterrichtspraxis notwendig ist. Zum anderen kann dieser durchaus emotionale Zugang durch die Einbeziehung theoretischer Konzepte erweitert werden, wodurch die jeweilige Beispielsituation aus verschiedenen Blickwinkeln betrachtet und variiert werden kann. Vignettenlektüren werfen Fragen auf, die Reflexionen und (Handlungs-)Möglichkeiten freisetzen können. Es gibt nicht einen richtigen Umgang, sondern verschiedene mögliche Alternativen, die im Nachhinein besprochen werden können, aber in der jeweiligen Situation ermittelt werden müssen (vgl. ebd., S. 28). Sie »fordern ein Sichpositionieren heraus und werfen häufig mehr Fragen als Antworten auf« (ebd.). Gleichzeitig geht es nicht darum, das Handeln von Lehrenden oder Lernenden normativ zu bewerten, sondern vielmehr darum, die unterschiedlichen Erfahrungen, Sichtweisen und Interpretationsmuster zu illustrieren (vgl. ebd.).

6. Ausblick

Mit dem Beitrag haben wir die Arbeit der drei *LemaS*-Teilprojekte des Deutsch-Clusters – *literarisches Schreiben*, *LemaS-GRiP* und *sprachlich-rhetorische Kommunikation* – in ihrer Zusammenarbeit mit den Schulvertreter:innen an den Kooperationsschulen beleuchtet und dabei aufgezeigt, wie Praxis und Forschung sich im Hinblick auf die Unterrichtsforschung und die Professionalisierung von Lehrpersonen gegenseitig befruchten können und wie entscheidend der Einsatz von Audio- und Videodaten auf verschiedenen Ebenen dabei war bzw. jetzt auch in der Transferphase – z. B. mittels Vignetten – noch ist. Anhand der Analyse von Audio-/Videomaterial konnten Erkenntnisse zu sprachlicher Begabung sowie wertvolle Einblicke in die Unterrichtspraxis und schulische Bedarfe gewonnen werden, die Grundlage für die Entwicklung von Konzepten und Materialien sowie passgenauen Professionalisierungsmaßnahmen für Lehrpersonen waren/sind. Die Arbeit mit den aufbereiteten Unterrichtsdaten in Form von Transkripten oder Text- und Videovignetten stellt zudem einen wesentlichen Anteil der Kooperation und des Dialogs zwischen Forschung und pädagogischer Praxis dar. Gerade der Einsatz von Vignetten zur Reflexion und Analyse von Unterrichtssituationen kann als praxisorientierter Ansatz verstanden werden, um komplexe unterrichtliche Herausforderungen zu bewältigen, die Fachlichkeit des

Unterrichts zu verbessern und wertvolle Impulse für die Weiterentwicklung der pädagogischen Praxis zu gewinnen.

Die durch die Bund-Länder-Initiative *Leistung macht Schule* angestoßene Weiterentwicklung und Verbreitung von evidenzbasierten Methoden dient der Förderung von Potenzialen und Fähigkeiten aller Schüler:innen. Mit der zweiten Phase rückt die Verbreitung und Vertiefung der in der ersten Phase gewonnenen Erkenntnisse und Erfahrungen in den Fokus. Der Forschungsverbund wird in dieser Phase die Implementierung und den Transfer der entwickelten *LemaS-P³rodukte* in neue Schulen unterstützen und begleiten. Besonders hervorzuheben ist dabei das Engagement der Lehrpersonen, die als Multiplikator:innen fungieren und ihre Erfahrungen sowie Inspiring-Practice-Beispiele teilen. Der erfolgreiche Transfer dieser Erkenntnisse und Materialien in neue Bildungskontexte wird entscheidend für die weitere Optimierung von Lern- und Bildungschancen und einer möglichst voraussetzungsarmen sprachlichen Förderung leistungsstarker und potenziell leistungsfähiger Schüler:innen sein.

Literatur

Abraham, Ulf (2017): Ist literarisches Schreiben »Produktionsorientierung«? Schreiben im Literatur- und Literatur im Schreibunterricht: spannende Verhältnisse. In: Wrobel, Dieter/Brand, Tilman von/Engelns, Markus (Hrsg.): Gestaltungsraum Deutschunterricht: Literatur – Kultur – Sprache. Baltmannsweiler: Schneider Verlag Hohengehren, S. 77–84.

Agostini, Evi/Anderegg, Niels (2021): »Den Zauber von Unterricht erfassen«. Die Arbeit mit Vignetten als Beitrag zur Professionalisierung und Schulentwicklung. In: Lernende Schule – Für die Praxis pädagogischer Schulentwicklung 94, H. 24, S. 26–29.

Asbrand, Barbara/Martens, Matthias (2018): Dokumentarische Unterrichtsforschung. Wiesbaden: Springer VS.

Becker-Mrotzek, Michael (2008): Gesprächskompetenz vermitteln und ermitteln: Gute Aufgaben im Bereich Sprechen und Zuhören. In: Bremerich-Vos, Albert/Granzer, Dietlinde/Köller, Olaf (Hrsg.): Lernstandsbestimmung im Fach Deutsch. Gute Aufgaben für den Unterricht. Weinheim/Basel: Beltz, S. 50–77.

Bergmann, Jörg R. (1980): Ethnomethodologische Konversationsanalyse. In: Schröder, Peter/Steger, Hugo (Hrsg.): Dialogforschung. Jahrbuch 1980 des Instituts für deutsche Sprache. Düsseldorf: Schwann, S. 9–51.

Bräu, Karin/Schwerdt, Ulrich (2005): Einleitung. In: Bräu, Karin/Schwerdt, Ulrich (Hrsg.): Heterogenität als Chance: Vom produktiven Umgang mit Gleichheit und Differenz in der Schule. Münster: Lit, S. 9–15.

Crapa, Gianmarco/Heinzel, Friederike/Hoffmann, Julia/Koch, Barbara (2021): Lernen am Fall. Reflexionen in der Schulentwicklung. In: Lernende Schule – Für die Praxis pädagogischer Schulentwicklung 94, H. 24, S. 23–25.

Deppermann, Arnulf (2008): Gespräche analysieren. Eine Einführung. 4. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

Grundler, Elke (2022): Miteinander sprechen. Gespräche als Lerngegenstand im Unterricht. In: Brand, Tilman von/Kilian, Jörg/Sosna, Anette/Riecke-Baulecke, Thomas (Hrsg.): Basiswissen Lehrerbildung: Deutsch unterrichten. Hannover: Klett/Kallmeyer, S. 176–189.

Heins, Jochen (2018): Was sind typische Problemsituationen im Literaturunterricht? Ein Rahmenmodell zur Systematisierung von Unterrichtssituationen für die Entwicklung von Vignetten. In: Didaktik Deutsch 44, S. 27–48.

Heinzel, Friederike (2021): Der Fall aus der Perspektive von Schulpädagogik und Lehrer*innenbildung. Ein Ordnungsversuch. In: Wittek, Doris/Rabe, Thorid/Ritter, Michael (Hrsg.): Kasuistik in Forschung und Lehre. Erziehungswissenschaftliche und fachdidaktische Ordnungsversuche. Bad Heilbrunn: Klinkhardt, S. 41–64. https://doi.org/10.25656/01:21561.

Kramer, Charlotte/König, Johannes/Kaiser, Gabriele/Ligtvoet, Rudy/Blömeke, Sigrid (2017): Der Einsatz von Unterrichtsvideos in der universitären Ausbildung. Zur Wirksamkeit video- und transkriptgestützter Seminare zur Klassenführung auf pädagogisches Wissen und situationsspezifische Fähigkeiten angehender Lehrkräfte. In: Zeitschrift für Erziehungswissenschaft, H. 1, S. 137–164.

Laudenberg, Beate (2024): Imitatives Schreiben. Die (Fan-)Rolle fiktiver Schreibvorbilder nutzen. In: Deutsch. Unterrichtspraxis für die Klassen 5–10, H. 79, S. 36–37.

Laudenberg, Beate/Neuweiler, Simone (2020): Teilprojekt 15: Förderung des literarischen Schreibens im Deutschunterricht der Primar- und Sekundarstufe I. In: Weigand, Gabriele/Fischer, Christian/Käpnick, Friedhelm/Perleth, Christoph/Preckel, Franzis/Vock, Miriam/Wollersheim, Heinz-Werner (Hrsg.): Leistung macht Schule. Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler. Weinheim und Basel: Beltz, S. 158–166.

Laudenberg, Beate/Sellinger, Lisa/Spiegel, Carmen/Winterscheid, Jenny (2022): Erkennen von sprachlich Leistungsstarken im Deutschunterricht – zwei Zugänge. In: Weigand, Gabriele/Fischer, Christian/Käpnick, Friedhelm/Perleth, Christoph/Preckel, Franzis/Vock, Miriam/Wollersheim, Heinz-Werner (Hrsg.): Dimensionen der Begabungs- und Begabtenförderung in der Schule. Band 2: Leistung macht Schule. Bielefeld: wbv, S. 357–369.

Laudenberg, Beate/Spiegel, Carmen (2020): Jenseits des Normalen? – zur Förderung sprachlicher und sprachästhetischer Begabungen. In: Fischer, Christian/Fischer-Ontrup, Christiane/Käpnick, Friedhelm/Neuber, Nils/Solzbacher, Claudia/Zwitserlood, Pienie (Hrsg.): Begabungsförderung, Leistungsentwicklung, Bildungsgerechtigkeit – für alle! Beiträge aus der Begabungsförderung. Münster/New York: Waxmann, S. 57–71 (Band 10: Begabungsförderung. Individuelle Förderung und inklusive Bildung hrsg. v. Christian Fischer).

Leistung macht Schule (o. J.[a]): Das ist die gemeinsame Bund-Länder-Initiative. https://www.leistung-macht-schule.de/index.html (04.10.2024).

Leistung macht Schule (o. J.[b]): Mit Freude literarische schreiben. https://www.leistung-macht-schule.de/de/Mit-Freude-literarisch-schreiben-1908.html (04.10.2024).

Leistung macht Schule (o. J.[c]): Mit einer WebQuest Jugendliche für Rhetorik begeistern. https://www.leistung-macht-schule.de/de/Mit-einer-WebQuest-Jugendliche-fur-Rhetorik-begeistern-1916.html (04.10.2024).

LemaS Forschungsverbund (o. J.[a]): LemaS-Transfer in die Schullandschaft. Die 2. Phase der Initiative Leistung macht Schule. https://phase1.lemas-forschung.de/themen/lemas-transfer (04.10.2024).

LemaS Forschungsverbund (o. J.[b]): Glossar. Zentrale Begriffe in LemaS. https://phase1.lemas-forschung.de/glossar (04.10.2024).

Mayer, Johannes (2018): Diversitätssensible Förderung von Begabten und (potenziell) Leistungsstarken im inklusiven Literaturunterricht. In: Laudenberg, Beate/Spiegel, Carmen (Hrsg.): Begabte und Leistungsstarke im Deutschunterricht. Band 1: Primarstufe. Baltmannsweiler: Schneider Verlag Hohengehren, S. 49–59.

Mayer, Johannes/Laudenberg, Beate/Mempel, Caterina/Winterscheid, Jenny/Zuber, Jasmin (2024): Sprachlich-literarisches Lernen begabungs- und begabtenförderlich gestalten (Symposium). In: Rogl, Silke/Resch, Claudia/Bögl, Elisabeth/Gürtler, Barbara/Hinterplattner, Sara/Klug, Julia (Hrsg.): Begabung verändert – förderliche Lernwelten erforschen, gestalten, implementieren. Münster/New York: Waxmann, S. 394–419.

Mayer, Johannes/Mempel, Caterina (2020): Teilprojekt 16: Förderung des sprachlich-ästhetischen Gesprächs im Regelunterricht im Fach Deutsch in der Primarstufe (LemaS-GRiP). In: Weigand, Gabriele/Fischer, Christian/Käpnick, Friedhelm/Perleth, Christoph/Preckel, Franzis/Vock, Miriam/Wollersheim, Heinz-Werner (Hrsg.): Leistung macht Schule. Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler. Weinheim/Basel: Beltz, S. 167–177.

Mayer, Johannes/Mempel, Caterina (2022): Potenziale von Grundschulkindern in literarischen Gesprächen erkennen und entfalten. In: Weigand, Gabriele/Fischer, Christian/Käpnick, Friedhelm/Perleth, Christoph/Preckel, Franzis/Vock, Miriam/Wollersheim, Heinz-Werner (Hrsg.):

Dimensionen der Begabungs- und Begabtenförderung in der Schule. Band 2: Leistung macht Schule. Bielefeld: wbv, S. 327–339.
Mempel, Caterina (2022): »Sie glänzte im blassen Schein der Sterne« Literarische und sprachliche Potenziale in Vorlesegesprächen entdecken und entfalten. In: Farkas, Katarina/Laudenberg, Beate/Mayer, Johannes/Rott, David (Hrsg.): Begabte Figuren in Literatur und Unterricht. Münster/New York: Waxmann, S. 113–130.
Mempel, Caterina (2023): Das weiß ich von jener Hexe. Peerinteraktive Bedeutungsaushandlung im Vorlesegespräch. In: Lingnau, Beate/Preußer, Ulrike (Hrsg.): Anschluss- und Begleitkommunikation zu literarischen Texten, S. 94–117. SLLD Sprachlich-literarisches Lernen und Deutschdidaktik. https://doi.org/10.46586/SLLD.263.
Mempel, Caterina/Mayer, Johannes (2021): Diversität als Ressource und Zielperspektive für die Gestaltung offener Gespräche im inklusiven Literaturunterricht. In: k:ON – Kölner Online Journal für Lehrer*innenbildung 2, H. 1, S. 83–108. https://doi.org/10.18716/ojs/kON/2021.1.5
Morek, Miriam/Heller, Vivien (2012): Bildungssprache – Kommunikative, epistemische, soziale und interaktive Aspekte ihres Gebrauchs. In: Zeitschrift für angewandte Linguistik 57, H. 1, S. 67–101.
Preckel, Franzis/Vock, Miriam (2013): Hochbegabung. Ein Lehrbuch zu Grundlagen, Diagnose und Fördermöglichkeiten. Göttingen: Hogrefe.
Prengel, Annedore (1993): Pädagogik der Vielfalt. Verschiedenheit und Gleichberechtigung in Interkultureller, Feministischer und Integrativer Pädagogik. Wiesbaden: VS Verlag für Sozialwissenschaften.
Roussel, Martin (Hrsg.) (2012): Kreativität des Findens, Figurationen des Zitats. München: Fink. https://doi.org/10.30965/9783846753057.
Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A Simplest Systematics for the Organization of Turn-Taking for Conversation. In: Language 50, H. 4, S. 696–735.
Schegloff, Emanuel A./Jefferson, Gail (1973): Opening up Closings. In: Semiotica 8, H. 4, S. 289–327.
Schmidt, Thomas/Schütte, Wilfried/Winterscheid, Jenny/Schürmann, Mia/Reineke, Silke/Schedl, Evi (2023): cGAT. Konventionen für das computergestützte Transkribieren in Anlehnung an das Gesprächsanalytische Transkriptionssystem 2 (GAT2). Version 1.2. Arbeitspapier. Mannheim: Leibniz-Institut für Deutsche Sprache. https://doi.org/10.14618/chrz-zy56.
Schwarze, Cordula (2023): Gespräche im Seminar über mündliche Rhetorik: Gesprächsanalytische Ergebnisse und hochschuldidaktische Potenziale. In: k:ON – Kölner Online Journal für Lehrer*innenbildung, Sonderausgabe 3, S. 66–93. https://doi.org/10.18716/ojs/kON/2023.s4.1.
Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar/Bergmann, Jörg/Bergmann, Pia/Birkner, Karin/Couper-Kuhlen, Elizabeth/Deppermann, Arnulf/Gilles, Peter/Günthner, Susanne/Hartung, Martin/Kern, Friederike/Mertzlufft, Christine/Meyer, Christian/Morek, Miriam/Oberzaucher, Frank/Peters, Jörg/Quasthoff, Uta/Schütte, Wilfried/Stukenbrock, Anja/Uhmann, Susanne (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion, Ausgabe 10, S. 353–402. www.gespraechsforschung-ozs.de (02.12.2024).
Spiegel, Carmen (2013): Die ausbildungsvorbereitende Förderung von Gesprächskompetenzen an Realschulen. In: Efing, Christian (Hrsg.): Ausbildungsvorbereitung im Deutschunterricht der Sekundarstufe I. Frankfurt am Main: Lang, S. 299–318.
Spiegel, Carmen (2018): Wenn sprachbegabte Kinder kommunizieren … In: Laudenberg, Beate/Spiegel, Carmen (Hrsg.): Begabte und Leistungsstarke im Deutschunterricht. Band 1: Primarstufe. Baltmannsweiler: Schneider Verlag Hohengehren, S. 26–35.
Spiegel, Carmen (2019): Gesprächskompetenzen Leistungsstarker und Begabter fördern. In: Spiegel, Carmen/Laudenberg, Beate (Hrsg.): Leistungsstarke und Begabte im Deutschunterricht. Band 2: Sekundarstufe. Baltmannsweiler: Schneider Verlag Hohengehren, S. 36–47.
Spiegel, Carmen/Laudenberg, Beate (2018): Einführung zur fachdidaktischen Begabungsforschung und Leistungsförderung. In: Laudenberg, Beate/Spiegel, Carmen (Hrsg.): Leistungsstarke und Begabte im Deutschunterricht. Baltmannsweiler: Schneider Verlag Hohengehren, S. 3–11.
Spiegel, Carmen/Winterscheid, Jenny (2020): Teilprojekt 17: Sprachlich-rhetorische Kompetenzen im Deutschunterricht fördern. In: Weigand, Gabriele/Fischer, Christian/Käpnick, Friedhelm/Perleth, Christoph/Preckel, Franzis/Vock, Miriam/Wollersheim, Heinz-Werner (Hrsg.): Leistung macht Schule. Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler. Weinheim/Basel: Beltz, S. 176–182.

Steiner, Edmund (2004): Erkenntnisentwicklung durch Arbeiten am Fall. Ein Beitrag zur Theorie fallbezogenen Lehrens und Lernens in Professionsausbildungen mit besonderer Berücksichtigung des Semiotischen Pragmatismus von Charles Sanders Peirce. Dissertation. Zürich: Philosophische Fakultät. http://www.ewi.tu-berlin.de/fileadmin/i49/dokumente/1143711480_diss_steiner.pdf (21.11.2024).

Es, Elizabeth A. van/Sherin, Miriam G. (2008): Mathematics teachers' »learning to notice« in the context of a video club. In: Teaching and Teacher Education 24, H. 2, S. 244–276.

Wagner-Willi, Monika (2004): Videointerpretation als mehrdimensionale Mikroanalyse am Beispiel schulischer Alltagsszenen. Zeitschrift für qualitative Bildungs- Beratungs- und Sozialforschung 5, H. 1, S. 59–66.

Winterscheid, Jenny (2019): Wenn Begabte und Leistungsstarke präsentieren … In: Spiegel, Carmen/Laudenberg, Beate (Hrsg.): Leistungsstarke und Begabte im Deutschunterricht. Band 2: Sekundarstufe. Baltmannsweiler: Schneider Verlag Hohengehren, S. 48–58.

Winterscheid, Jenny (2024): Förderung rhetorisch-sprachlicher Kompetenzen im Unterricht – exemplarisch gezeigt an zwei Produkten. In: Weigand, Gabriele/Fischer, Christian/Käpnick, Friedhelm/Perleth, Christoph/Preckel, Franzis/Vock, Miriam/Wollersheim, Heinz-Werner (Hrsg.): Wege der Begabungsförderung in Schule und Unterricht. Transformative Impulse aus Wissenschaft und Praxis. LemaS-Sammelband 3. Bielefeld: wbv, S. 271–280.

Winterscheid, Jenny/Sellinger, Lisa (2023): Wie äußert sich sprachliche Begabung und wie lässt sie sich fördern? In: Fischer, Christian/Fischer-Ontrup, Christina/Käpnick, Friedhelm/Neuber, Nils/Reintjes, Christian (Hrsg.): Potenziale erkennen – Talente entwickeln – Bildung nachhaltig gestalten. Beiträge aus der Begabungsförderung. Münster/New York: Waxmann, S. 137–145.

Verdeckte Videographie

Datenerhebung und -analyse im Spannungsverhältnis von Erkenntnisinteresse und Forschungsethik

Ajit Singh

Zusammenfassung: Mein Beitrag reflektiert methodologische und forschungsethische Implikationen verdeckter Erhebungen von Videodaten und setzt diese in Relation zum Erkenntnisinteresse videographischer Forschung. Während ethnographische Studien seit langem verdeckte Beobachtungen in ihr Methodenrepertoire integrieren und durchaus fruchtbare Erkenntnisse zu Tage fördern, ist videographische Forschung mit einem doppelten »Problem« konfrontiert: Erstens ist ihr Aufzeichnungsgerät, die Videokamera, vom Körper entkoppelt und damit i. d. R. sichtbar. Zweitens werden die untersuchten Realitätsausschnitte (Orte, Räume, Akteure) immer technisch registriert und als Datum audiovisuell abgebildet. Auf der Basis eigener videographischer Aufzeichnungen während der Covid-19-Pandemie werden exemplarische Forschungserfahrungen reflexiv zum Gegenstand gemacht, um insbesondere die »Ethnomethoden« verdeckter Videoforschung ins Auge zu fassen. Neben der spezifischen Auswirkung auf die Positionalität der Forscher*innen, wird auch berücksichtigt, inwieweit sich die Praxis verdeckter Videoforschung in das Datenmaterial einschreibt.

Schlüsselwörter: verdeckte Videographie, Ethnographie, Forschungsethik, öffentlicher Raum, Reaktanz

1. Einleitung: Ethnographische Videoforschung und Forschungsethik

In der qualitativen Sozialforschung lässt sich anhand verschiedener methodischer Verfahren – von qualitativen Interviews bis zur teilnehmenden Beobachtung in der Ethnographie – aufzeigen, dass mit spezifischen Techniken der Verdeckung geforscht wird, die mehr oder weniger als legitim gelten. In meinem Beitrag möchte ich die methodologische Frage behandeln, ob und unter welchen Bedingungen auch die *verdeckte Videographie* als ein möglicher Weg der Erkenntnisproduktion angesehen werden kann. Meine Ausführungen beschränken sich dabei vor allem auf *verdeckte Erhebungen* von Videodaten *im öffentlichen*

Raum.[1] Diese Frage lässt sich auf den ersten Blick insofern klar beantworten, dass verdecktes Videographieren die rechtlichen Anforderungen hinsichtlich der Informationspflicht (DSGVO Art. 13) und der Selbstbestimmung der Beforschten (DSGVO Art. 7) ignoriert und dadurch forschungsethische Grundsätze unterläuft. Festmachen lässt sich der methodologische ›Stellenwert‹ verdeckter ethnographischer Videoforschung vor allem durch seine Nicht-Thematisierung in Beiträgen, die sich zwar explizit mit Videoforschung und Ethik befassen, dabei verdeckte Videoforschung aber nicht mal erwähnen (Everri et al. 2020). In der videographischen Forschung werden forschungsethische Fragen üblicherweise im Vorfeld der Erhebung formal durch die »informierte Einwilligung« geklärt. Eine weitere Form des Schutzes der Beforschten findet sich auf der Ebene der Analyse und Ergebnisdarstellung durch die Anonymisierung der Videodaten durch Filter. Für Publikationen werden Standbilder verwendet und durch Nachzeichnungen, Modifikationen, Extraktionen verfremdet oder dekontextualisiert (u. a. Goodwin 1995; Laurier 2014; Tuma/Knoblauch/Schnettler 2013; Wilke 2018). Womöglich auch deshalb werden Probleme und Möglichkeiten verdeckter Videographie kaum mit dem Prozess der Datenerhebung und der Datenanalyse in Verbindung gebracht.

Forschungsethik wird in diesem Beitrag nicht wissenschaftsphilosophisch reflektiert, sondern hinsichtlich ihrer Realisierung in der qualitativen Erforschung sozialer Phänomene. Die Sensibilisierung für forschungsethische Standards existiert nicht erst seit dem Aufkommen von Videotechnologien oder der späten Digitalisierung des Internets. Im Jahr 1969 hat die American Sociological Association (ASA) einen Ethikkodex für Sozialforschung entwickelt. Ein wesentliches Ziel dieses, wie später auch anderer Kodexe, war es, einen Orientierungsrahmen für Forschungshandeln und den Umgang mit den Forschungsteilnehmer*innen zu liefern, aus dem auch für Nicht-Forschende der rechtliche und normative Rahmen von Forschung ersichtlich wird. Idealiter wird sich an den primären Grundsätzen orientiert, dass Beforschte *erstens* nicht geschädigt werden dürfen, *zweitens* dass ihre Persönlichkeitsrechte gewahrt bleiben und *drittens*, dass sie informiert werden müssen, wenn sie Teil einer Erhebung werden oder es auch »passiv« werden können (vgl. Hopf 2018). Es dürfte kein Zufall sein, dass die British Sociological Association (BSA), die für verdeckte Forschung durchaus Raum lässt, in ihrem »statement of ethical practice« verdecktes Forschungshandeln unter dem Thema »Relations to the research participants« behandelt (BSA 2017, S. 4 f.). Im besten Fall sind dies nicht nur formale Standards, sondern Richtlinien, die an die praktische

1 Mein Beitrag basiert auf einem Vortrag, den ich im August 2022 auf der Midterm Conference des ESA Research Networks Qualitative Methods (RN20) in Kopenhagen gehalten habe. Für weitere Hinweise und Diskussionen bedanke ich mich bei Paul Eisewicht, Bernt Schnettler sowie bei René Wilke.

und reflexive Selbststeuerung der Forschenden und mitunter auch der Teilnehmenden in der situativen Forschungspraxis angepasst werden (vgl. Mondada 2014).

In meinem Beitrag konzentriere ich mich auf ethnographisches Forschungshandeln, das ich als Referenz für Videographie nehme. Dafür befasse ich mich zunächst mit dem Verhältnis von offener und verdeckter Ethnographie. Im Anschluss erläutere ich einige Implikationen zur Videographie, ihrem Gegenstand und ihrer Datenerhebungstechnik. Danach beschreibe ich die Forschungserfahrungen meines verdeckten videographischen Selbstversuchs während der Covid-19-Pandemie. Im Schlussteil reflektiere und diskutiere ich Nutzen, Grenzen und die Relevanz verdeckter Videographie im Lichte aktueller gesellschaftlicher Entwicklungen und neuer Nutzungsszenarien von Videodaten.

2. Verdeckt und offen Forschen in der (soziologischen) Ethnographie

Die Ethnographie ist weder eine Methode noch eine monomethodische Forschungspraxis. Vielmehr verbirgt sich hinter dem Label der Ethnographie ein »integrierter Forschungsansatz« (Breidenstein et al. 2013, S. 34) und ein multiprogrammatischer Forschungsstil, der sich im Laufe der letzten 60 Jahre in verschiedene Stile ausdifferenziert hat (vgl. Eisewicht 2016; Poferl/Schröer 2022). Kennzeichnend für Ethnographien ist, dass sie sich bei der Auswahl ihrer Erhebungsmethoden und ihrer Forschungsgegenstände nicht allzu starke Selbstbeschränkungen auferlegt, wenngleich über die »teilnehmende Beobachtung« (Spradley 1980) oder auch über die »beobachtende Teilnahme« (u. a. Honer 1993; Hitzler/Eisewicht 2016) ein zentraler Zugang zur Erforschung sozialer Gemeinschaften und Milieus sowie für die Rekonstruktion feldspezifischer Praktiken und kommunikativer Handlungen gelegt wird.

Etwas diffiziler wird es, wenn man nach den rechtlichen und ethischen Standards *offener und verdeckter Beobachtungen* fragt. Für die ethnographische Stadtforschung in der Tradition der Chicagoer Schule war neben der offenen auch die verdeckte Beobachtung immer Bestandteil ethnographischer Praxis. Wie mit dem Problem der Verdeckung umgegangen wurde, lässt sich am Beispiel dreier Klassiker der Ethnographie umreißen: *Howard Becker* untersuchte in den 1940er Jahren deviantes Verhalten am Beispiel von Dance Musicians. Da er selbst in der Chicagoer Musikszene als Klavierspieler ›unterwegs‹ war, musste er seine Rolle als Forscher überhaupt nicht sichtbar machen. Statt formelle Interviews zu führen, hörte er stattdessen den Musikerinnen bei ihren Gesprächen zu und legte weder Rolle, Interesse noch, dass er Daten erhebt, offen (Becker 1966, S. 84). *Erving Goffman* wiederum hat seine Forscherrolle meistens insofern kaschiert, dass er in Rollen geschlüpft ist. Für seine bahnbrechende Studie zu totalen Institutionen (»Asylums«, Goffman 1961) am St. Elizabeth Hospital in Washington D. C.

war Goffman »an assistant to the athletic director, when pressed avowing to be a student of recreation and community life« (Smith 2022, S. 22). Feldrollen zu übernehmen hatte bei ihm Methode, im Casino, in Shetland und nicht zuletzt bei seinen Beobachtungen von Verhalten in öffentlichen Räumen. Vermutlich würde heute diese Art ethnographischer Forschung in Ethik-Ausschüssen kritischer beleuchtet. Seine Erträge sind jedoch immer noch von unschätzbarem Wert. *Willliam Foot Whyte* (1979) hat wiederum darauf hingewiesen, dass verdeckte Forschung nicht heißt, »undercover« zu sein. Sie zielt auch nicht darauf, die Untersuchten in ihrem Fehlverhalten zu entlarven oder wie ein*e Spion*in eine Gegenseite auszuspielen, sondern Wissen zu produzieren, ohne den Beforschten zu schaden.[2] Für Whyte legitimiert sich verdeckte Forschung im Kontext der eigenen Gesellschaft mitunter als einziger Weg oder als eine kombinierte Praxis beider Forschungsstile (Whyte 1979, S. 57). Er negiert vor allem nicht den hohen Stellenwert, den der Aufbau von Feldbeziehungen hat und betont, dass dies in seinem Verständnis der Königsweg der Ethnographie ist.

Dabei ist zu klären, was eigentlich genau verdeckt wird: Neben der eigenen (Feld-)Rolle, ist dies auch das Forschungsinteresse. So ist es durchaus zum Allgemeinplatz geworden, dass das eigene Erkenntnisinteresse im Feld zwar ausgewiesen wird, aber eben nur so weit wie nötig, um das Feld – auch im Verständnis einer Reaktivität des Feldes – nicht zu stark zu beeinflussen. Ebenso wichtig scheint mir jedoch zu sein, genauer auf die *Erhebungs- und Analysekontexte* zu schauen und welchen Einfluss die verdeckte Forschungspraxis auf die Aufzeichnung und analytische Verwendung von Videodaten hat.

Richtet man den Blick auf die *Erhebungskontexte*, kommt der Ethnographie ein besonderer Aspekt »zugute«: Ein wesentlicher forschungspraktischer Vorteil der klassischerweise als teilnehmende Beobachtung durchgeführten Ethnographie liegt darin, dass sie prinzipiell ohne großes Equipment auskommt. Ein zentrales Beobachtungsinstrument ist allen voran der eigene Leib/Körper mit seinen multisensorischen Wahrnehmungsqualitäten (u. a. Goffman 1989, S. 125; Poferl et al. 2023). In physischer Kopräsenz können Forschende gar nicht anders, als die von ihnen bewohnten Situationen auch leibkörperlich wahrzunehmen, d. h. Orte zu riechen, Dinge zu spüren oder die Handlungen anderer zu hören und zu sehen. Ethnographisches Forschen in diesem leibkörperlichen Sinne ist zumindest je nach Grad der Involviertheit *immer ein wenig verdeckt*. Oder um es praktisch zu formulieren: kein*e Ethnograph*in würde seine*ihre ethnographische Forschung dadurch ausweisen, indem sie*er einen Aushang macht, dass sie*er gerade mit ihrem*seinem Körper Daten erhebt. Die Anwendung von Erhebungsmethoden sieht man auch im Falle von ethnographischen

2 »The researcher should be committed to the gathering of information to increase scientific knowledge. In reporting the findings, the researcher is obligated to try to avoid damaging the organization or its individual members« (Whyte 1979, S. 58).

Gesprächen (sofern nicht, wie etwa bei qualitativen Interviews üblich, Daten mit einem Aufzeichnungsgerät erhoben werden) gemeinhin nicht direkt. Die Fähigkeit zu verdecken oder so ins Feld einzutauchen, dass man irgendwann sozial inventarisiert wird, ist vielmehr eine Kernkompetenz teilnehmender ethnographischer Beobachtung. Je nach methodologischem Paradigma, wird diese Art der vermeintlichen Distanzlosigkeit oder das hohe Maß an Subjektivität zwar auch kritisch betrachtet. Neben diesen kritischen Tönen und Reflexionen – auch gegenüber einer ethischen Überregulierung durch Ethik-Komitees (vgl. Hammersley 2009) – weist die Literatur zur Ethnographie aber auch Positionen auf, die eine methodologische Legitimierung verdeckter Forschungspraktiken adressieren.

So reflektiert Homan (1980) anhand seiner ethnographischen Studie zu Pfingstbewegungen, dass verdeckte Forschung (in seinem Fall in der Rolle eines potenziellen Konvertiten) unter pragmatischen Aspekten sinnvoll erscheinen kann, weil sie Zugänge zu sozialen Welten ermöglicht, die eher im Geheimen existierten. Unter ethischen Gesichtspunkten sieht er darin auch einen Schutz der »Ungestörtheit« der Probanden*innen gewahrt. Darüber hinaus betont Homan, dass sowohl die individuelle Moral als auch die Berufsethik der Forschenden davon berührt werden.[3] Würde man die unterschiedlichen Studien, die sich auf verdeckte Methoden berufen, ohne den Anspruch auf Vollständigkeit auf einen Nenner bringen, dann würde dieser sicherlich in den Gegenständen begründet liegen, die untersucht werden. Anvisiert werden i. d. R. solche Felder, die schwer zugänglich sind, die am Rande gesellschaftlicher Normen angesiedelt sind, prekäre Arbeitsbedingungen (Loyd 2012), Szenen oder deviante Subkulturen ins Auge fassen, wie die Welt von Türsteher*innen (Hobbs/O'Brien/Westmarland 2003; Calvey 2019) oder Hooligans (Pearson 2009).

Die Legitimation solcher verdeckten Studien vollzieht sich daher auch in der Zuschreibung eines gesellschaftlichen Mehrwertes, weil soziale Probleme sichtbar gemacht werden, die sonst verborgen blieben (vgl. den aufschlussreichen Überblick von Roulet et al. 2017). Dahinter verbergen sich spezifische Forschungstraditionen und Erkenntnisinteressen, die auch entlang der Frage zu diskutieren wären, inwieweit entweder ein ethischer Common Sense qualitativer Sozialforschung unterlaufen wird oder hier gar eine »kreative« Form der Gegenstandsangemessenheit ethnographischer Methoden proklamierbar ist (vgl. Calvey 2019).

3 Diese Trennung der Forscher*innenrolle als individuelles und professionelles Subjekt wird von Bulmer kritisiert: »The pretence of role- playing and the individual moral consequences of doing so are bound up together. If the researcher employs ›a dishonest representation of his interests‹ and indeed of his identity, he does so as a sociologist in a professional role as well as an individual in his role as citizen« (Bulmer 1980, S. 60).

3. Vom Leibkörper als Erkenntnismittel zur Kamera als Aufzeichnungsapparat

Wie eingangs erwähnt, orientieren sich meine Ausführungen an einer bestimmten Verfahrensform der soziologischen Ethnographie, die als »Fokussierte Ethnographie« (Knoblauch 2001) und eng damit verbunden auch als »Videographie« (Tuma/Knoblauch/Schnettler 2013) bezeichnet wird. Der Fokus liegt auf Prozessdaten, also auf der interpretativen Analyse real-ablaufender Interaktionen in ihren jeweiligen sozialen Situationen, die i. d. R., aber nicht ausschließlich, von den Forschenden audiovisuell und videographisch aufgezeichnet werden. Die Traditionen der fokussierten Ethnographie liegen in den Forschungslinien der Ethnomethodologie (Garfinkel 1967), der Konversationsanalyse (Sacks 1992) und in der ethnographischen Untersuchung sozialer Interaktionsordnungen (Goffman 1983). Hierbei wird die Reflexivität des forschenden Subjekts als kommunikativ handelnder Akteur stets mitgedacht (Knoblauch/Tuma 2015). Forschende sind damit immer auch Teil nicht nur der Gesellschaft, sondern der konkret von ihnen untersuchten Situationen. Und damit ist auch die Entscheidung, offen oder verdeckt zu forschen, eine Entscheidung der Forschenden, die in welcher Form auch immer Folgen für die Datenerhebung, die Beschaffenheit von Daten und für ihre Analyse hat.

Vom Standpunkt verdeckter Datenerhebung aus ist videographische Forschung mit einem doppelten »Problem« konfrontiert: Erstens ist ihr Aufzeichnungsgerät, die Videokamera, vom Körper entkoppelt und damit i. d. R. sichtbar. Die »Unauffälligkeit« teilnehmender Beobachtung wird dahingehend gebrochen, dass das Instrument der Beobachtung nicht mehr nur der forschende Körper ist, sondern ein ausgelagertes Aufzeichnungsgerät. Das markiert eine grundlegende Veränderung des ethnographischen Modus Operandi, weil Forschende mit der Kamera in der Hand kaum mehr teilnehmen, sondern größtenteils die Position des/der Aufzeichnenden einnehmen (zur Kombination videographierender und beobachtend teilnehmender Ethnographie vgl. Vollmer 2024). Zweitens werden die untersuchten Realitätsausschnitte (Orte, Räume, Akteure) immer technisch »registriert« (Bergmann 1985) und als Datum audiovisuell und technisch abgebildet.[4] Die Anonymisierung von Videodaten spielt also per se eine wichtige Rolle, weil sie im Gegensatz zur rekonstruktiven, schriftsprachlichen Beschreibung soziale Situationen und ihre Akteure erkennbar werden lassen.

Weil soziale Interaktionen im Fokus der Videographie stehen, orientiert sich ihre videobasierte Analyse stark an ethnomethodologischen,

4 Den Begriff der »Abbildung« verwende ich hier ausschließlich in einem technischen Sinne und nicht im theoretischen und methodologischen Verständnis einer objektiven Abbildung von Realität. In diese Kerbe schlug ja bereits die Writing Culture Debatte in der Ethnographie, die sich auf die Bildlichkeit visueller Ethnographien übertragen lässt.

konversationsanalytischen und hermeneutischen Prämissen (Knoblauch 2004; Knoblauch/Schnettler 2012; Heath/Hindmarsh/Luff 2010). Videographische Forschung folgt damit einem Erkenntnisinteresse, Interaktionen nicht als Selbstdeutungen des Feldes, sondern in einem ethnomethodologischen Sinne als »naturally organized ordinary activities« (Garfinkel 1991, S. 14) zu untersuchen und soziale Ordnung und die Herstellung sozialer Wirklichkeit aus ihrem aktualen Vollzug heraus zu analysieren und zu verstehen (vgl. Knoblauch/Wilke i. d. B.). Neben ihrer methodischen Fokussierung auf das situative Geschehen richten sich die Fokussierte Ethnographie wie auch die Videographie auf »Handlungs- und Kommunikationszusammenhänge« (Knoblauch 2001, S. 137) und beziehen analytisch deshalb sehr zentral auch solche Wissensbestände ein, die für die Konstitution feldspezifischer Ordnungen relevant sind. Was methodologisch gemeinhin als »Kontextwissen« (Tuma/Knoblauch/Schnettler 2013, S. 88 f.) bezeichnet wird, meint gleichsam ein Verstehen der emischen Perspektive, nicht nur einzelner Subjekte, sondern auch der kontextuellen Strukturen eines spezifischen Feldes und seiner orientierungsstiftenden Wissensvorräte.

Die Videographie, wie auch andere verwandte Ansätze der videobasierten Erhebung sozialer Interaktionen, vertreten weitestgehend eine eindeutige Haltung zu verdeckter Forschung. Sie orientieren sich am ethischen Kanon qualitativer Sozialforschung und setzen den »informed consent« als maßgeblich voraus (u. a. Heath/Hindmarsh/Luff 2010, S. 18; Tuma/Knoblauch/Schnettler 2013, S. 67; vom Lehn 2018, S. 129 f.). In der Regel werden Aushänge gemacht oder Informationsblätter verteilt. Menschen erhalten damit auch die Gelegenheit, ihr Gefilmtwerden abzulehnen. Für videographisch Forschende kann dies ein Problem darstellen, sofern die Verwendung einer Videoaufnahme im Nachhinein nicht gestattet wird. Mit der verdeckten Videographie lässt sich einem weiteren Problem begegnen, das als methodologisches Kernargument für verdecktes Videographieren schlechthin gelten kann, nämlich die »Reaktanz«. Oftmals wird angemahnt, dass der Einfluss von Video(forschung) und im Speziellen des Kamerahandelns auf die Erhebungssituation und damit die Aufzeichnung »natürlicher Daten« (vgl. hierzu auch Lynch 2002) zu groß sei und Menschen ihr Handeln in Orientierung an der Kamera verändern. Gegen die Invasivitätsannahme der Kamera wird eingewendet, dass sich Gewöhnungseffekte einstellen, die aus der wechselseitigen Herstellung von Vertrautheit resultieren (Tuma/Knoblauch/Schnettler 2013, S 13 f.). Jüngere Studien betonen hingegen, dass zwischen Gewöhnung und Reaktanz ein Unterschied besteht und dass Reaktanz nicht immer eindeutig im Material zu identifizieren ist (Goerigk et al. 2021).[5] Darüber hinaus muss Reaktanz nicht nur mit einer Veränderung des Handelns einhergehen.

5 Trotz neuerer Erkenntnisse ist festzuhalten, dass Untersuchungen zur Reaktanz in der Videoforschung (ähnlich wie Untersuchungen zur verdeckten Videographie) bis auf wenige Ausnahmen (vgl. Lahn/Klette 2023, S. 14) kaum systematisch durchgeführt werden.

Gleichermaßen lässt sich das Agieren mit und das Schauen in die Kamera als Hinweis deuten, dass die Beforschten sich über die Aufzeichnung bewusst sind – womit eine »Exit-Strategie« zumindest potenziell möglich wäre.

4. Videographische Forschung während der Covid-19-Pandemie

Meine Ausführungen resultieren aus dem Forschungskontext des deutschlandweit durchgeführten Datenerhebungsprojektes »Social Distancing und neue Raumformen der Interaktion«.[6] Zwischen März 2020 bis etwa November 2020 führten verschiedene Forscher*innen in öffentlich wie auch in nicht-öffentlich zugänglichen Räumen Erhebungen durch, um die sichtbaren Veränderungen sozialer Erscheinungsformen in der Covid-19-Pandemie an möglichst vielen Orten wie Kaufhäusern (vgl. Sibert i. E.), urbanen Räumen, Parks, Boulderhallen, bei Demonstrationen (vgl. Meier zu Verl/Koch/Meyer 2021) beim Einkaufen, Spazierengehen, beim Schlangestehen vor Geschäften (vgl. Singh/Eisewicht i. E.) zu untersuchen.[7] Das Erkenntnisinteresse des Projektes gründet auf der Beobachtung, dass die Herstellung von »Social Distancing« in vielen sozialen Situationen in Konflikt zu routinisierten und institutionalisierten Handlungsformen stand und die Menschen beispielsweise vor Probleme stellte, wie Abstände kommuniziert und wie die Einhaltung von Abstandsregeln situativ hergestellt werden (Knoblauch 2020). Um diese sozialen und räumlichen Interaktionsformen zu erfassen und feinkörnig zu rekonstruieren, wurde ein videographischer Zugang gewählt.

»Überwältigt« von der sich auf alle Wirklichkeitsbereiche ausweitenden pandemischen Situation, stellte sich für uns die Frage, wie wir diese einmalige Phase – mit Bezug auf die spezifische Situation in Deutschland – methodisch erfassen können, die so offenkundig das soziale und öffentliche Leben veränderte. Begab man sich in Organisationen oder kleinere Geschäfte wie einen Copyshop, lag es völlig nahe, dass man zunächst mit den Betreiber*innen solcher Geschäfte Kontakt aufnimmt (vgl. Singh/Meinert 2023). Das entspricht der Vorgehensweise,

6 Finanziert wurde das Datenerhebungsprojekt durch die Berlin University Alliance (BUA). Ziel der Studie ist es, auf Basis videografisch erhobener Daten das Verständnis komplexer, kulturell und situativ organisierter Dynamiken von öffentlichen Begegnungen im Zusammenhang mit dem Infektionsgeschehen und Hygienemaßnahmen genauer zu verstehen (vgl. Tuma et al. in Vorbereitung). Geleitet wird das Projekt von Hubert Knoblauch, Uwe Flick und René Tuma (zur Projektvorstellung siehe https://www.tu.berlin/as/forschung/laufende-projekte/social-distancing-und-neue-raumformen-der-interaktion wie auch: https://www.youtube.com/watch?v=EDElsiCKSEU, 26.10.2021).

7 Die Daten wurden größtenteils in Zusammenarbeit mit dem Forschungsdatenzentrum aviDa an der TU Berlin archiviert und für die Nachnutzung aufbereitet (s.https://depositonce.tu-berlin.de/items/95703e69-2501-445c-a975-c27b691d646c). Zur Nachnutzung von Forschungsdaten vgl. auch Wilke/Knoblauch (2022).

wie sie für die ethnographische Organisationsforschung (vgl. Bachmann 2009, S. 251) und in jüngerer Zeit auch für die videographische Untersuchung von (Groß-)Veranstaltungen und Events (Knoblauch/Singh 2023) üblich ist. Für die Erhebung von Videodaten in solchen Kontexten ist der formale Weg des Feldzugangs nahezu unumgänglich.

Bei der Forschung in öffentlichen Räumen bewegt sich die Videographie seit jeher in einem Graubereich zwischen der Wahrung von Persönlichkeitsrechten und Wissenschaftsfreiheit (GG, Art. 5 Abs. 3): Ab wann müssen Menschen in welcher Form informiert werden? Unter welche Bereiche der Rechtsprechung fällt dies und welche ist zu priorisieren? Wann läuft man Gefahr, Persönlichkeitsrechte zu verletzen (vgl. Tuma/Knoblauch/Schnettler 2013, S. 67 f.)? Einige dieser Fragen sind zwar im Grundgesetz (GG), im Bundesdatenschutzgesetz (BDSG, § 27) und der Datenschutz-Grundverordnung (DSGVO, Art. 15–18; 21, 77) geregelt. Dennoch treffen diese Gesetze auf eine forscherische Praxis mit sehr spezifischen Kontextbedingungen und Erkenntnisinteressen, sodass die Abwägung für ein Für oder ein Wider etwa im Sinne der Forschungsfreiheit schwierig zu überblicken und eindeutig zu definieren ist.

4.1 Erhebungs- und Analysekontext: Verdeckte Videoaufzeichnung im öffentlichen Raum

Die folgenden Datenauszüge entstanden im Rahmen einer *verdeckten* Aufzeichnung. Im ersten Fall haben wir uns (als zwei Forschende) auf eine Bank gesetzt, die sich an einem öffentlich zugänglichen Vorplatz befand. Auch wenn es Gelegenheiten gibt, sich dort länger aufzuhalten, handelt es sich bei diesem Ort eher um eine Durchgangspassage, die je nach Richtung in einen belebten Stadtpark hinein- oder aus diesem hinausführt. Die Videokamera von etwa 20 cm Länge und 8 cm Breite und Höhe platzierten wir zwischen uns auf der Parkbank und bedeckten sie mit einem Kleidungsstück, sodass die Kameralinse freie Sicht hatte. Das videographische Interesse an diesem Ort bestand darin, Menschen dabei zu beobachten, wie sie aus unterschiedlichen Richtungen kommend diesen Vorplatz überqueren. Von der einen Seite kamen Spaziergänger*innen, Jogger*innen oder Fahrradfahrer*innen von der Uferpromenade eines Flusses, von der anderen Seite gingen Menschen, die zumeist von dem direkt angrenzenden S-Bahnhof kamen, in Richtung des Flusses und des Parks. Die dritte Route liegt frontal zur Kameraperspektive (siehe Abb. 1a und b).[8]

8 Bei Kim Blaik bedanke ich mich für die Nachzeichnung der Videostandbilder.

Abbildung 1a: Paare

Abbildung 1b: wenig Distanz

Im Fokus unserer Ortsbeobachtung standen also *Menschen und ihre Körper*, die sich allein oder in kleineren Gruppen fortbewegten (Abb. 1a). Selbst in solchen Situationen (siehe Abb. 1b), in denen sich Menschen auf engstem Raum begegneten, wurden Kollisionen umstandslos und sehr koordiniert vermieden. In unseren Aufzeichnungen ließ sich tatsächlich keine ›auffällige‹ Situation identifizieren, in denen Passant*innen sich auf eine störende Weise zu nahegekommen sind. Überdies ist festzuhalten, dass der von uns ausgesuchte öffentliche Raum nicht coronaspezifisch »refiguriert« war. Es gab keine Wegmarkierungen oder Hinweisschilder, die ein spezifisches Verhalten instruierten, wie dies in manchen Einkaufsstraßen oder in Geschäften der Fall war. Der zum Erhebungszeitpunkt vorherrschende gesamtgesellschaftliche Kontext der Pandemie schien sich nicht sonderlich in den verkörperten Handlungen der Fußgänger*innen und Passant*innen sichtbar (»accountable«) zu machen.

Mit Phänomenen der öffentlichen Ordnungsbildung, auch solchen im Straßenverkehr, hat sich Goffman bereits in den 1970er Jahren befasst. Analytisch bezeichnete er Fußgänger*innen als »vehicular units« (Goffman 1971), die sich wie Pilot*innen durch den Verkehr navigieren. Fußgänger*innen verfügen über verschiedene Techniken, um sich erfolgreich und schadlos durch den öffentlichen Raum zu manövrieren. Eine besondere Praktik beschreibt Goffman als »Scanning« (ebd., S. 11), was sich sowohl auf die Beobachtung des physischen Raumes als auch auf die »Externalisierungen« der anderen Verkehrsteilnehmer*innen richtet (ebd., S. 11 f.). Auch wir wurden mitunter ›gescannt‹, ob als Ensemble dieser lokalen Szenerie oder doch als mit verdeckter Kamera aufzeichnende Akteure, lässt sich nachträglich nicht auflösen. Das Setting und die Funktion dieser Passage wie auch unser Kameraplatz auf einer Parkbank schien uns aber eine unspezifische Daseinsberechtigung zu geben, auch weil die Passant*innen nicht länger als nötig dort verharrten, sondern größtenteils unter Einhaltung der Verkehrsregeln (»traffic codes«) vom einen Punkt zu einem anderen gelangen wollten (vgl. Goffman 1971, S. 6). Aufgrund der Flüchtigkeit der dort aufkommenden Situationen erschien es daher nicht besonders riskant, Aufzeichnungen verdeckt vorzunehmen.

Für die zweite Aufnahmesituation (siehe Abb. 2) suchten wir einen anderen Ort in unmittelbarer Nähe auf, dessen räumliche Infrastruktur sich vom ersten deutlich unterscheidet. Es handelt sich um eine Promenade, die an einem Fluss

entlangführt. Direkt daran angeschlossen sind verschiedene Bootsanlegestellen, von denen aus für gewöhnlich kommerzielle, touristische Bootstouren angeboten werden. Daran angebunden sind kleinere Fischräucher- und Pommesbuden, Cafés etc., die Spaziergänger*innen zum Kauf und Verzehr von Speisen und Getränken animieren. Wir wählten einen Spot aus, von dem aus wir gut das Treiben vor einem der Cafés beobachten konnten und setzten uns dafür an einen Tisch. Die Kamera positionierten wir so auf dem Tisch, dass sie für die Passant*innen nicht direkt als Aufnahmegerät identifizierbar war und bedeckten sie erneut mit ein paar Kleidungsstücken. Das Objektiv war aber so positioniert, dass die Sicht auf die Aktivitäten vor dem Café freigelegt war. Weil der Verkauf außerhalb des Cafés vor einem Fenster stattfand, bildeten sich vor dem Laden kleinere Warteschlangen. Die Warteschlangen setzten sich teilweise bis auf die Promenade fort, sodass ständig Passant*innen durch die Schlangen hindurch- oder an ihnen vorbeimussten. Infolgedessen wurde der Raum durch die Kund*innen und die Spaziergänger*innen in unterschiedliche Zonen aufgegliedert, die sich vereinfacht gesagt als Wartebereich und als Passage beschreiben lassen (siehe Abb. 2, Bild 1, der Mann mit Hut, der zwischen C und D hindurch geht).

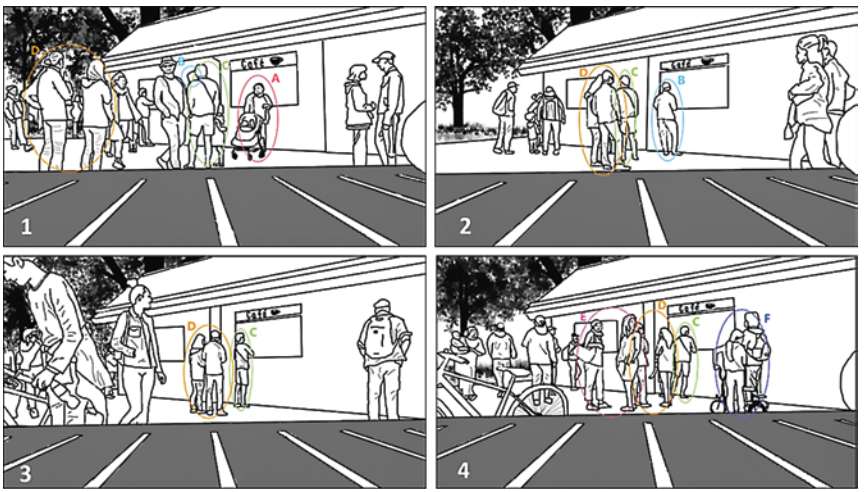


Abbildung 2: Schlangestehen

Im Fokus standen in dieser Erhebungssituation die körperlich-räumlichen Arrangements, die wir als »doing waiting« (Ayaß 2020) oder als »Queuing« (vgl. Singh/Eisewicht i. E.) identifizieren und analysieren. Auch hier sind Körper in Bewegung, allerdings erweist sich die räumliche Situation als weniger flüchtig, weil die die Warteschlange bildenden Akteure deutlich länger an einem Ort verweilen und mehr Aufwand leisten müssen, um die lokale räumliche und physische Ordnung der Warteschlange sichtbar aufrechtzuerhalten. Wie sich an Abbildung 2 (Bilder 1–4) zeigt, wandelt sich die Anordnung der Schlange mit der Zeit

(zunächst frontales Anstehen in den Bildern 1–3, später von der rechten Seite anstehend, s. Bild 4). Die unterschiedlichen Abstände der jeweiligen »Glieder« (in Buchstaben A–F) in der Warteschlange liefern aber Hinweise darauf, dass der Umgang mit der pandemischen Situation zwar unterschiedlich gehandhabt, aber auch deutlicher im »display behavior« der Handelnden angezeigt wird (vgl. Abb. 2, Bild 1 der Abstand von D zu C und Bild 4 der Abstand von F zu E). In einer Warteschlange zeigen Wartende nicht nur an, dass sie sich in einer Schlange befinden. Ebenso wie sich fortbewegende »vehicular units«, markieren sie durch ihre körperliche Anordnung auch ihre lokale Zugehörigkeit zu einem bestimmten Glied innerhalb der Kette, also der Vater mit dem Kinderwagen (A), der Mann mit seinem Kind (C) oder die drei Frauen (D), die sich gemeinsam anstellen und gleichermaßen Abstand zu (C) halten. Durch die spezifische Ausrichtung frontal zum Café und inmitten auf der Promenade ist jedoch für die Wartenden das Problem zu lösen, wie mit jenen umzugehen ist, die sich im Bewegungsstrom auf einer »Route« befinden. Wie sich zeigt, werden in Abbildung 2, Bild 1 über die größere Lücke zwischen D zu B nicht nur spezifische Abstände zueinander innerhalb der Schlange koordiniert. Gleichzeitig öffnet sich zwischen beiden Gruppierungen – ob intendiert oder nicht – eine Art Korridor, der von den Spaziergänger*innen, als Durchgang wahrgenommen und genutzt wird.

4.2 Die ›Ethnomethoden‹ verdeckter Videographie und ihre Daten

Im Unterschied zu einer *offenen Datenerhebung*, die im besten Fall Gewöhnungseffekte nach sich zieht, führt der Erhebungskontext der verdeckten Videographie auf beiden Seiten nicht zur Gewöhnung. Wenn man verdeckt forscht, muss man einiges dafür tun, dass die Datenerhebung auch verdeckt bleibt. Forschende unternehmen dafür Anstrengungen, die Kamera zu verstecken, sich so zu positionieren, dass man nicht direkt wahrgenommen wird. Die Kamera wird dafür – wenn man so will – unsichtbar gemacht. Der konstruierte Zauber dieser Art Erhebungssituation liegt in der Idealisierung der Natürlichkeitsunterstellung des situativen Geschehens, die sich in den *Ethnomethoden verdeckter videographischer Forschung* ausdrücken.[9] Während die Offenlegung der Forschung sogar ein Türöffner sein kann, wird die verdeckte Erhebung von Videodaten von vorneherein wie ein Stigma oder ein Geheimnis gehütet, das es vor »Anderen« zu schützen gilt. Die Kamera ist dabei ein störendes Artefakt, weil sie gleichermaßen verdächtig macht. Das Managen des forschenden Selbst, das permanent ertappt werden könnte, wirkt sich nicht nur auf die Erhebungspraxis aus, sondern ist dabei auch ein Gradmesser für die Situierung von Ethik und Normerwartung.

9 Zu den Ethnomethoden ethnographischer Forschung vgl. Schindler (2012).

Verdeckte Forschung ist jedoch nicht nur folgenreich für den Prozess der Datenerhebung, sondern auch für die *Beschaffenheit der Daten und den Analysekontext*. Die verdeckte Erhebung von Videodaten schreibt sich unweigerlich in das Material ein. Dies zeigt sich in verschiedenen Momenten der Datensichtung: *Erstens* wird dies erkennbar an den Bildausschnitten, die entscheidend dafür sind, in welchem Detailierungsgrad *Realitätsausschnitte* untersucht werden können. Die Wahl der Kameraperspektive durch das Kamerahandeln konstruiert im Wesentlichen das Datum mit. Eine sichtbare Positionierung der Kamera im (öffentlichen) Raum ermöglicht es, den Fokus der Aufzeichnung zu kalibrieren und zu überprüfen. In Abbildung 1 ist wiederum deutlich zu sehen, dass durch die verdeckte Erhebung nur von unten aufgezeichnet wurde, sodass uns die Passant*innen in einem spezifischen Blickwinkel erscheinen. In Abbildung 2 ist wiederum erkennbar, dass der Tisch sehr prominent einen Teil des Bildausschnitts bestimmt. Für die Analyse mag dies nicht prinzipiell problematisch sein. Dennoch zeigt sich im ersten wie im zweiten Beispiel, dass sich eine transparente videographische Praxis durch den Einsatz entsprechenden Equipments (Stativ, Mikrofone etc.) und durch die Herstellung einer geeigneten Bildperspektive kennzeichnet, die in unserer Aufzeichnung nur unzureichend gewährleistet war.

Zweitens lassen sich nicht zwangsläufig alle relevanten Aspekte sozialer Interaktion erfassen. Auf manche Modalitäten erhielten wir durch unser Kamerahandeln nur eingeschränkt visuellen Zugriff oder es fehlte gänzlich an adäquater Akustik, weil wir nicht nah genug an die Orte und Menschen herankamen. Auch die Analyse des performativen Zusammenspiels von gesprochener Sprache (Prosodie, Intonation etc.) und Körperlichkeit (u. a. Mimik, Gestik) ließ sich nicht vollständig ausschöpfen, weil bestimmte Modalitäten des kommunikativen Handelns nicht technisch registriert werden konnten. Diese technischen Hürden des Tons und der Perspektivität sind ja generell in videographischen Forschungsarbeiten zu überwinden. Bei der verdeckten Erhebung werden diese Einschränkungen jedoch besonders virulent, weil die Flexibilität des Kamerahandelns (bspw. durch korrigierende Positionswechsel, mit Stativ oder dynamisch) auf ein Minimum reduziert ist. Die Aufzeichnung audiovisueller Daten und ihre valide Auswertung lebt deshalb auch von ihrer Präzision bei der Auswahl des geeigneten Aufzeichnungsstandortes. Eine gute Kameraposition, die dort hin filmt, wo sich die relevanten Aktivitäten ereignen, ist eine zentrale Voraussetzung guter videographischer Forschung.

Drittens kommt bei der Analyse verdeckt erhobener Daten dem Kontextwissen eine erhebliche Bedeutung zu. Dies lässt sich aus der Zugänglichkeit zu öffentlichen Räumen ableiten, die es den Forschenden ermöglicht, einen Ort erst über einen längeren Zeitraum hinweg zu beobachten und dann zu filmen. Videographische Forschung in öffentlichen Räumen richtet sich oftmals auf die fokussierte Beobachtung solcher Handlungsformen, die eine wiederkehrende Typik aufweisen (bspw. die Praxis des aneinander Vorbeigehens, des Wartens oder des

Schlangestehens) und die über den spezifischen Kontext hinaus auf eine alltägliche Vertrautheit rekurrieren. Wie auch Goffmans Studien (u. a. 1971) zeigten, basieren diese Untersuchungen fundamental auf gesellschaftlich geteilten Alltagserfahrungen (und weniger auf einem geheimen oder professionellen Sonderwissen) und praktisch erworbenen Handlungskompetenzen der Forschenden, die der Analyse als Kontext- und als ein bestehendes Alltagswissen hinzugefügt werden.

So mag für eine verdeckte videographische Erhebung in öffentlichen Räumen durchaus sprechen, dass trotz verminderter Qualität der Videodaten oder der Vermeidung von Reaktanz auf ein gesellschaftlich geteiltes Alltagswissen zurückgegriffen werden kann, um eine weitestgehend valide und belastbare Analyse sozialer Interaktion zu gewährleisten. Gleichzeitig ist zu konstatieren, dass die Stärken einer feingranularen Analyse kommunikativer Aktivitäten eingeschränkt werden. Nicht davon unberührt bleiben zudem die spezifischen Bedingungen des Erhebungskontextes und die einseitige Rahmung einer Forschungssituation, in der Daten in Unkenntlichmachung einer Erhebungsabsicht erzeugt wurden. Dies lenkt den Blick auf die Reflexivität und Positionalität der eigenen Forscher*innenrolle jenseits des Rechtsrahmens. Denn der Einsatz von Videotechnik erzeugt in der Aufnahmesituation immer schon eine technisch vermittelte Asymmetrie, die aus der Anordnung von Aufzeichnenden zu Aufgezeichneten resultiert (vgl. Everri et al. 2020, S. 79). Diese Dysbalance von Machtverhältnissen verschärft sich gerade dann, wenn die Aufzeichnung nicht transparent gemacht wird und im weiteren Verlauf der Forschung Daten nicht nur gespeichert, sondern ggf. auch nachgenutzt werden können.

5. Diskussion: Grenzen, Nutzen und Relevanz verdeckter Videographie

Meine Ausführungen zielten auf die Reflexion verdeckter videographischer Forschung im öffentlichen Raum während der Covid-19-Pandemie. Für die Interaktionsforschung erwies sich die im Alltag auch als ›Coronapandemie‹ bezeichnete Krise als ein historisch einmaliges Phänomen, weil sich der Umgang mit dem Sars-Cov-2 Virus auf alle Bereiche der Gesellschaft (auch der Forschung) ausgeweitet hatte. Eine schwierige soziale Nebenfolge für die Videographie resultierte seinerzeit daraus, in den Interaktionspartner*innen (und vice versa) immer auch ein potenzielles Risiko zu sehen. Das erforderte eine Reflexion unserer eigenen Positionalität als Forscher*innen, d. h. wie viel Nähe innerhalb der Erhebungen angemessen ist, was man ›riskieren‹ kann oder in jedem Fall nicht tun sollte. Aus dem Geist dieser historischen Lage sowie der Ungewissheit um Dauer und Verlauf der Pandemie, erschien es naheliegend, über Verfahren der verdeckten Datenerhebung nachzudenken.

Verdecktes Videographieren kann als eine Technik betrachtet werden, die in ihrer besten Absicht einen möglichst unverstellten Blick auf Sozialität gewährt. Für ethnographische Studien hat sich die verdeckte Beobachtung vor allem dort als nützlich erwiesen, wo entweder dezidiert abweichendes Verhalten und Handeln am Rande der gesellschaftlichen Norm untersucht wurde und wo Feldzugänge entsprechend problematisch waren. Überdies wird in der Forschungsliteratur betont, dass verdeckte Beobachtung Forscher*innen noch authentischer in die Lage eines »being a ›pure‹ insider« versetzt, was »also allows the researcher to experience firsthand the phenomena under study in the same way that the participants experience it« (Roulet et al. 2017, S. 492).

Für die Videographie (Tuma/Knoblauch/Schnettler 2013) wie auch für ihre »verwandten« Verfahren der fokussierten Ethnographie (Knoblauch 2001) und der Konversationsanalyse (Sacks 1992) ist zwar die Feldzugangsproblematik und die Schwierigkeiten der situativen Erhebung von (mitunter sensiblen) Prozessdaten ein bekanntes Phänomen. Von größerer Relevanz war für unsere Datenerhebung die durch die Videokamera evozierte »Reaktanz«, die als Kernargument für die verdeckte Erhebung von Videodaten angesehen werden kann. Aus dieser Perspektive stellt sich zuvorderst die Frage nach der Qualität und der analytischen Verwendbarkeit von Videodaten, die durch Reaktanzen gekennzeichnet oder gar beeinflusst sind (vgl. Goerigk et al. 2021). Die verdeckte Videographie versucht dieses Problem dezidiert zu umgehen, worin sicherlich ein Potenzial dieser Vorgehensweise liegt. Sie handelt sich dabei aber einige eklatante forschungsethische und methodologische Schwierigkeiten ein, die hier nochmal knapp zusammengefasst werden: Erstens lässt sich verdecktes Videographieren, sowohl in öffentlichen wie privaten Räumen, rechtlich und forschungsethisch kaum legitimieren. Der forschungsethische Regelbruch tangiert in zweifacher Hinsicht den professionellen Status der Forschenden, die am Ende nicht nur sich, sondern auch ihre Forschung legitimieren müssen. Gleichzeitig markiert die Tatsache, dass Forscher*innen in der Situation als Forschende zurechenbar sind und bleiben, ihre spezifische Positionalität im Feld wie auch innerhalb der Gesellschaft, die sie untersuchen. Zweitens schreibt sich die eigentümliche Erhebungssituation in die Daten ein und kann dahingehend nicht nur die Qualität des audiovisuellen Materials, sondern auch deren Analyse limitieren. Drittens kommt dem Kontextwissen gerade dann eine wichtige Funktion zu, wenn die untersuchten Phänomene aus dem Alltagswissen heraus rekonstruiert werden können.

Darüber hinaus stellt sich aber die Frage, inwieweit es nicht nur ethische und methodologische, sondern auch empirische und strukturelle Grenzen videographischer Forschung gibt, und zwar insbesondere dort, wo die Erhebung eigener Videodaten kaum oder nicht mal mehr verdeckt möglich ist. Dazu zählt in jüngerer Zeit die videobasierte Untersuchung von situierten

Gewalthandlungen mittels fremdproduzierter Videos (vgl. Weenink/Tuma/van Bruchem 2022; vgl. auch Kissmann; Egbert/Janssen; Meyer/Grauert/Oberzaucher i. d. B.). Die in der videographischen Literatur oft aufgeführte Beobachtung der Zunahme der alltäglichen Verfügbarkeit von Aufzeichnungstechnologien (von Videokameras bis zu Smartphones), öffentliche Videoüberwachung und die digitale Verbreitung von Videomaterial perpetuieren offenkundig neue Nutzungsszenarien (vgl. Knoblauch/Wilke i. d. B.), sodass eigene, auch verdeckte audiovisuelle Aufzeichnungen kaum noch erforderlich sind, um soziale Situationen videogestützt zu analysieren. Gemeint ist eine Form der registrierenden Selbstbeobachtung der Gesellschaft, die uns in gewisser Weise einen Teil unserer soziologischen Forschung ›abzunehmen‹ scheint. Selbsterzeugte Videos, die auf YouTube, Instagram, Tiktok oder auf der Kurznachrichtenplattform »X« hochgeladen werden und auch Webcam-Aufzeichnungen öffentlicher Orte sind vielfach frei und digital im Internet verfügbar und analysierbar (vgl. Nassauer 2018). Videoaufnahmen sind damit auch ohne großen Aufwand der Forscher*innen potenziell ›erhebbar‹, mehr oder weniger als ›Daten‹ verwendbar und deshalb mittlerweile auch Gegenstand forschungsethischer Fragen und methodologisch reflektierter Lösungsansätze (vgl. Legewie/Nassauer 2018). Im Verhältnis zu verdeckter Videoforschung lässt sich konstatieren, dass diese Videodaten nicht genuin für Forschungszwecke erzeugt wurden und i. d. R. auch die Kontexte ihrer Aufzeichnung für Forschende nur unzureichend bekannt sind. Aus Sicht der Aufgezeichneten bleibt wiederum nicht nur die Aufzeichnung selbst verdeckt, sondern auch die spezifische Rekontextualisierung und Weiternutzung solcher Daten, selbst wenn sie sich um die Einhaltung forschungsethischer Prämissen bemüht.

In den jeweiligen Ethikkodexen der ASA (2008) oder der BSA (2017) wird dem Thema »covert research« ein eigener Abschnitt (ebd., S. 5) gewidmet. Es besteht weiterhin ein loser Konsens darüber, dass verdeckte Forschung unter bestimmten Bedingungen möglich und in Teilen sogar erlaubt ist. Eine eindeutige Lösung für die Legitimierbarkeit verdeckter Videographie lässt sich meines Erachtens nicht gänzlich darlegen. Vielleicht wird sich die Frage der Notwendigkeit verdeckter Videographie künftig auch deshalb nicht in dem Maße stellen, wenn Videodaten vermehrt als fremdproduzierte Onlinedaten zugänglich sind. Dennoch dürfte ein entscheidendes Argument für eine mehr oder weniger offene oder verdeckte videobasierte Forschung von öffentlichen Räumen im Spannungsfeld der Wissenschaftsfreiheit liegen und der Tatsache, dass videographische Forschende ihre Daten nicht für kommerzielle Zwecke erheben und verarbeiten. Die Diskussion um das Für und Wider von verdeckter Videographie verdeutlicht damit vor allem, dass in diesem Fall Methodologie, Forschungsethik und Rechtsprechung nicht gänzlich losgelöst voneinander zu betrachten sind.

Literatur

Atkinson, Paul (2009): Ethics and ethnography. In: Twenty-First Century Society, 4(1), S. 17–30.

Ayaß, Ruth (2020): Doing Waiting: An Ethnomethodological Analysis. In: Journal of Contemporary Ethnography, 49(4), S. 419–455.

Becker, Howard S. (1966): Outsiders. Studies in the sociology of deviance. New York: The Free Press.

Breidenstein, Georg/Hirschauer, Stefan/Kalthoff, Herbert/Nieswand, Boris (2013): Ethnographie. Konstanz: UVK.

British Sociological Association (2017): Statement of Ethical Practice. BSA Publications.

Bulmer, Martin (1980): Comment on ›The Ethics of Covert Methods‹. In: British Journal of Sociology, 31(1), S. 59–65.

Calvey, David (2019): The everyday world of bouncers: a rehabilitated role for covert ethnography. In: Qualitative Research, 19(3), S. 247–262.

Eisewicht, Paul (2016): Zur Kartographierung ethnographischer Forschungsprogramme. In: Hitzler, Ronald/Kreher, Simone/Poferl, Angelika/Schröer, Norbert (Hrsg.): Old School – New School? Zur Frage der Optimierung ethnographischer Datengenerierungen. Essen: Oldib, S. 29–39.

Everri, Marina/Heitmayer, Maxi/Paulius, Yamin-Slotkus/Saadi, Lahlou (2020): Ethical challenges of using video for qualitative research and ethnography. State of the art and guidelines. In: Lähdesmäki, Tuuli/Koskinen-Koivisto, Eerika/Čeginskas, Viktorija L. A./Koistinen, Aino-Kaisa (Hrsg.): Challenges and solutions in ethnographic research. Ethnography with a twist. London: Routledge, S. 68–83.

Garfinkel, Harold (1967): Studies in Ethnomethodology. Englewood Cliffs, NJ: Prentice Hall.

Garfinkel, Harold (1991): Respecification: Evidence for Locally Produced, Naturally Accountable Phenomena of Order, Logic, Reason, Meaning, Method, etc. in and as of the Essential Haecceity of Immortal Ordinary Society (I) – an Announcement of Studies. In: Button, Graham (Hrsg.): Ethnomethodology and the Human Sciences. Cambridge: Cambridge University Press, S. 10–19.

Goerigk, Paul/Weiser-Zurmühlen, Kristin/Erichsen, Göntje/Wild, Kathrin (2021): Zur Einschätzung von Reaktanz und Invasivität in videografischen Daten. Ein kontextsensibler Ansatz zur Beurteilung der Kamerarolle in Aufnahmesituationen [63 Absätze]. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 22(2), Art. 3. http://dx.doi.org/10.17169/fqs-22.2.3547.

Goffman, Erving (1961): Asylums. Essays on the Social Situation of Mental Patients and Other Inmates. New York: Doubleday.

Goffman, Erving (1971): Relations in Public. Microstudies of social order. New York: Basic Books.

Goffman, Erving (1983): Presidential Address: The Interaction Order«. In: American Sociological Review, 48, S. 1–17.

Goffman, Erving (1989): On Fieldwork. In: Journal of Contemporary Ethnography, 18(2), S. 123–132.

Hammersley, Martin (2009): Against the ethicists: on the evils of ethical regulation. In: International Journal of Social Research Methodology, 12(3), S. 211–225.

Heath, Christian/Hindmarsh, Jon/Luff, Paul (2010): Video in Qualitative Research. London: Sage.

Hitzler, Ronald/Eisewicht, Paul (2016): Lebensweltliche Ethnographie. Weinheim/Basel: Beltz Juventa.

Hobbs, Dick/O'Brien, Kate/Westmarland Louise (2007): Connecting the gendered door: women, violence and door work. In: British Journal of Sociology, 58(1), S. 21–38.

Honer, Anne (1993): Lebensweltliche Ethnographie. Wiesbaden: Deutscher Universitäts Verlag.

Hopf, Christel (2016): Forschungsethik und qualitative Forschung. In: Hopf, Wulf/Kuckartz, Udo (Hrsg.): Schriften zu Methodologie und Methoden qualitativer Sozialforschung. Wiesbaden: Springer VS, S. 195–205.

Knoblauch, Hubert (2001): Fokussierte Ethnographie. Soziologie, Ethnologie und die neue Welle der Ethnographie. In: Sozialer Sinn, 2(1), S. 123–142.

Knoblauch, Hubert (2004): Die Video-Interaktions-Analyse. In: Sozialer Sinn, 1, S. 123–138.

Knoblauch, Hubert (2020): Learning to Dance: Social Distancing und die Refiguration der Interaktionsordnung; Beitrag im Blog des SFB 1265. https://sfb1265.de/blog/learning-to-dance-social-distancing-und-die-refiguration-der-interaktionsordnung (12.09.2024).

Knoblauch, Hubert/Schnettler, Bernt (2012): Videography: Analysing Video Data as ›Focused‹ Ethnographic and Hermeneutical Exercise. In: Qualitative Research, 1(3), S. 334–356.

Knoblauch, Hubert/Singh, Ajit (Hrsg.) (2023): Kommunikative Gattungen und Events. Wiesbaden: VS Verlag für Sozialwissenschaften.
Knoblauch, Hubert/Tuma, René (2015): Praxis, kommunikatives Handeln und die Videoanalyse der Videoanalyse In: Deppermann, Arnulf/Feilke, Helmuth/Linke, Angelika (Hrsg.): Sprachliche und kommunikative Praktiken. Berlin/Boston: De Gruyter, S. 229–250.
Knoblauch, Hubert/Vollmer, Theresa (2022): Fokussierte Ethnographie. In: Poferl, Angelika/Schröer, Norbert (Hrsg.): Handbuch Soziologische Ethnographie. Wiesbaden: Springer VS, S. 353–366.
Lahn, Leif C./Kirsti Klette (2022): Reactivity beyond Contamination. An Integrative Literature Review of Video Studies in Educational Research. In: International Journal of Research & Method in Education, 46(1), S. 1–18. https://doi.org/10.1080/1743727X.2022.2094356.
Laurier, Eric (2014): The Graphic Transcript: Poaching Comic Book Grammar for Inscribing the Visual, Spatial and Temporal Aspects of Action. In: Geography Compass, 8(4), S. 235–248.
Legewie, Nicolas/Nassauer, Anne (2018): YouTube, Google, Facebook: 21st Century Online Video Research and Research Ethics [51 paragraphs]. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 19(3), Art. 32.
Lloyd, Anthony (2012): Working to live, not living to work: Work, leisure and youth identity among call centre workers in North East England. In: Current Sociology, 60(5), S. 619–635.
Lynch, Michael (2002): From naturally occurring data to naturally organized ordinary activities: comment on Speer. In: Discourse Studies, 4(4), S. 531–537.
Meier zu Verl, Christian/Koch, Sebastian/Meyer, Christian (2021): Streit um den Gemeinsinn. Interaktionen zwischen Protestierenden und anderen Anwesenden bei den Corona-Protesten in Konstanz. In: Reichardt, Sven (Hrsg.): Die Misstrauensgemeinschaft der »Querdenker«. Frankfurt am Main: Campus, S. 255–291.
Mondada, Lorenza (2014): Ethics in Action: Anonymization as a Participant's Concern and a Participant's Practice. In: Human Studies, 37, S. 179–209.
Nassauer, Anne (2018): How Robberies Succeed or Fail: Analyzing Crime Caught on CCTV. In: Journal of Research in Crime and Delinquency, 55(1), S. 125–54.
Pearson, Geoff (2009): The researcher as hooligan: where ›participant‹ observation means breaking the law. In: International Journal of Social Research Methodology, 12(3), S. 243–255.
Poferl, Angelika/Schöer, Norbert (Hrsg.) (2022): Handbuch Soziologische Ethnographie. Wiesbaden: VS Verlag für Sozialwissenschaften.
Poferl, Angelika/Schröer, Norbert/Hitzler, Ronald/Kreher, Simone (Hrsg.) (2023): Leib-Körper-Ethnographie: Erkundungen zum Leib-Sein und Körper-Haben. Essen: Oldib.
Roulet, Thomas J./Gill, Michael. J./Stenger, Sebastian/Gill, David. James (2017): Reconsidering the Value of Covert Research: The Role of Ambiguous Consent in Participant Observation. In: Organizational Research Methods, 20(3), S. 487–517.
Sacks, Harvey (1992): Lectures on Conversation, Vol. I. Oxford: Blackwell.
Schindler, Larissa (2012): Visuelle Kommunikation und die Ethnomethoden der Ethnographie. In: Österreichische Zeitschrift für Soziologie, 37, S. 165–183.
Singh, Ajit/Eisewicht, Paul (i. E.): Queueing interaction bodies: indexicality and (mis)interpretations of bodily-spatial arrangements in spaces of consumption. In: Singh, Ajit/Meier zu Verl, Christian/Tuma, René (Hrsg.): Knowledge on rewind: how video analysis contributes to social theory. London: Routledge.
Singh, Ajit/Meinert, Leopold (2023): Doing Videography – Mit der Kamera im Feld während der Corona-Pandemie. In: Donlic, Jasmin (Hrsg.): Qualitative Methoden in der Forschungspraxis: Perspektiven, Erfahrungen und Anwendungsfelder. Leverkusen: Budrich, S. 256–277.
Smith, Gregory W. H. (2022): Goffman's Empirical Work. In: Lenz, Karl/Hettlage, Robert (Hrsg.): Goffman-Handbuch. Leben – Werk – Wirkung, Stuttgart: J. B. Metzler, S. 19–23.
Spradley, James P. (1980): Participant Observation. New York: Holt, Rinehart, and Winston.
Tuma, René (2022): Videobasierte Ethnographie. In: Poferl, Angelika/Schröer, Norbert (Hrsg.): Handbuch Soziologische Ethnographie. Wiesbaden: Springer VS, S. 563–574.
Tuma, René/Knoblauch, Hubert/Schnettler, Bernt (2013): Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: VS Verlag für Sozialwissenschaften.
vom Lehn, Dirk (2018): Ethnomethodologische Interaktionsanalyse: Videodaten analysieren und die Organisation von Handlungen darstellen. Weinheim/Basel: Beltz Juventa.

Weenink, Don/Tuma, René/van Bruchem, Marly (2022): How to Start a Fight: A Qualitative Video Analysis of the Trajectories Toward Violence Based on Phone-Camera Recorded Fights. In: Human Studies, 45(3), S. 577–605.
Whyte, William F. (1979): On Making the Most of Participant Observation. In: The American Sociologist, 14(1), S. 56–6.
Wilke, René (2018): Das Interpretations-Bild. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 485–500.
Wilke, René/Knoblauch, Hubert (2021): Faire Prinzipien – für die Nachnutzung von Forschungsdaten und die Primärforschung. In: Soziologie, 50(4), S. 431–438.

III. Interaktionsanalysen mit Ethnographischen Videos

2. Interaktion und Interaktivität mit Gegenständen, Technik und digitalen Technologien

Analyse atypischer Kommunikation

Transkription von Videosequenzen als analytischer Ansatzpunkt von Mehrpersonengesprächen im Kontext Unterstützter Kommunikation

Imke Niediek

Zusammenfassung: Der Beitrag möchte die Bedeutung der Transkription für die Analyse von Videodaten (natürlicher) Interaktionssituationen in atypischen Gesprächssituationen am Beispiel der Nutzung von elektronischen Kommunikationshilfsmitteln (sogenannten Talkern) diskutieren. Durch die Nutzung von ›Talker‹ werden erwartbare Interaktionsvollzüge irritiert, sodass die Gesprächsorganisation eine besondere Herausforderung für die Interaktionsbeteiligten darstellt. Am Beispiel von zwei Videosequenzen wird herausgearbeitet, wie diese Herausforderungen sich auch auf die Transkription und Analyse von Mehrpersonengesprächen mit Talker auswirken. Neben methodischen Aspekten, wie ergänzenden Transkriptionskonventionen, wird deutlich, dass für die Analyse von Mehrpersonengesprächen mit Talker bereits der Transkriptionsprozess selbst einen erheblichen Anteil an der Datenanalyse hat und auch die Klärung methodologischer Fragen, wie dem Status Talkers für das Interaktionsgeschehen, beinhaltet. Letztlich wird dafür plädiert, die mehrfache Transkription von Sequenzen als Bestandteil der Analyse zu nutzen.

Schlüsselwörter: atypische Gesprächssituation, Transkription, Unterstützte Kommunikation, Gesprächsorganisation, Interaktionsanalyse

1. Einleitung

Der Beitrag möchte die Bedeutung der Transkription für die Analyse von Videodaten (natürlicher) Interaktionssituationen für das Forschungsfeld der atypischen Gesprächssituationen diskutieren. Dies erfolgt am Beispiel der Kommunikation mit elektronischen Kommunikationshilfsmitteln von Menschen mit kommunikativen Beeinträchtigungen.

Alternative und ergänzende Kommunikationsformen wie Gebärden/Gesten, Bildkarten, Kommunikationstafeln, Sprachcomputer (im weiteren Verlauf ›Talker‹ genannt) sollen die Interaktion mit anderen, die Einflussnahme auf die persönliche Umgebung und schließlich die volle Partizipation von Personen

mit Kommunikationsbeeinträchtigungen ermöglichen. Beeinträchtigungen der Kommunikation können motorische und/oder kognitive Ursachen haben und sich sowohl auf die linguistische, die operative, die soziale und die strategische Domäne beziehen (Light/McNaughton 2014). Die kommunikativen Kompetenzen, die die betreffenden Personen in der Gesprächssituation mobilisieren können, werden zudem von psychosozialen Faktoren, wie Selbstvertrauen oder Motivation, beeinflusst (ebd.). Talker sollen also zusätzliche Möglichkeiten eröffnen, um Informationen zu teilen, Bedürfnisse und Wünsche ausdrücken zu können und soziale Beziehungen aufzubauen und zu pflegen (Light/McNaughton 2014; Fornefeld 2010). Obwohl Talker einen wichtigen Beitrag zur Ermöglichung von Teilhabe in Gesprächssituationen leisten, entstehen durch die Beteiligung eines nicht-menschlichen Dritten bzw. Aktanden (Ankrich/Latour 2006) in der Kommunikationssituation besondere Herausforderungen in der Gesprächsorganisation. Der folgende Beitrag beschreibt, welche Auswirkungen diese Herausforderungen auf die Erforschung atypischer Gesprächssituationen haben und zeigt an empirischen Beispielen, wie diese bearbeitet werden können.

2. Erhebungskontext im weiteren Sinn

Wilkinson (2019) skizziert in einem Überblick drei typische Problemfelder der Gesprächsorganisation in atypischen Gesprächssituationen (z. B. bei Aphasie, Demenz, kognitive Entwicklungsbeeinträchtigung, motorische Kommunikationsbeeinträchtigung), die auf die Kommunikation mit einem Talker gleichsam zutreffen.

Das erste Problemfeld umfasst Auswirkungen von Verzögerungen im Gesprächsverlauf auf die Gesprächsorganisation: So führt die Verwendung von Kommunikationshilfsmitteln zu einer Verlangsamung der Äußerungsproduktion (2–26 Wörter pro Minute im Vergleich zu 80–120 Wörtern bei lautsprachlicher Kommunikation, Kraat 1985). Gesprächspartner*innen ohne Kommunikationshilfsmittel können also in derselben Zeit eine deutlich höhere Anzahl und Vielfalt von Redebeiträgen in das Gespräch einbringen. Die Verzögerung kann so massiv sein, dass es zu einem Zusammenbruch der zeitlichen, sequenziellen und thematischen Struktur des Gesprächs kommt oder zumindest das wechselseitige Verstehen erschwert ist (Auer/Hörmeyer 2017; Higginbotham/Engelke 2013; Higginbotham/Wilkins 1999; Light et al. 1985a). Zudem produzieren Sprecher*innen ohne Kommunikationsbeeinträchtigung häufiger eine neue Gesprächswendung als Sprecher*innen mit einem Talker (Clarke 2016; Light et al. 1985a; Light et al. 1985b; Light et al. 1985c; Kraat 1985). Dadurch haben erstere einen größeren Anteil an der Kontrolle des Gesprächsverlaufs (Sundqvist/Plejert/Rönnberg 2010).

Das zweite Problemfeld beschreibt die Anforderung, eine für das Gegenüber verständliche (bzw. interpretierbare) Äußerung zu formulieren. Aufgrund der verlangsamten Äußerungsgeschwindigkeit werden aber mit einem Talker häufig nur kurze Aussagen gemacht und mit körpereigenen Kommunikationsformen wie Gestik, Mimik und Vokalisationen ergänzt (Clarke/Kirton 2003; Light et al. 1985c; Antaki et al. 2017). Die komplexe, multimodale Struktur derartiger Äußerungen wird von Wickenden (2011a) auch als ein spezifisches ›Kreol‹ bezeichnet, die nur teilweise allgemein üblichen Erwartungen an eine verständliche Äußerung entspricht. Sie muss von Gesprächspartner*innen ohne Talker situativ interpretiert, gegebenenfalls ergänzt und kommunikativ abgesichert werden. Der Co-Konstruktion von Äußerungen über mehrere Turns hinweg (Goodwin et al. 1995; Renner et al. 2019) kommt daher in Gesprächen mit einem Talker im Vergleich zu Gesprächen ohne Talker eine besondere Bedeutung zu.

Teilweise versuchen die Zuhörenden, das Gemeinte noch vor Beendigung eines turns zu antizipieren und unterbrechen den Redezug der Sprecher*in, um den Abschluss des Turns zu beschleunigen (Savolainen/Klippi/Launonen 2020; Bloch 2011). Wenn die Vervollständigung misslingt bzw. falsch erweitert wurde, entsteht die Notwendigkeit, einen Reparatur-Turn zu initiieren. Dies kann umfangreiche Nebensequenzen mit mehreren Turn-Wechseln nach sich ziehen, die lediglich der Verständnisabsicherung der ursprünglichen Äußerung dienen (Griffiths et al. 2015).

Das dritte Problemfeld, das Wilkinson (2019) skizziert, umfasst die kommunikative Handlung, die durch den Redebeitrag verbal ausgeführt wird. Da Äußerungen häufig ohne die in der mündlichen Kommunikation üblichen Floskeln, bildhaften Ausdrücke, ohne smalltalk oder Partikeln auskommen muss (und dadurch teilweise mehr der Schrift- als der Lautsprache verwandt ist), fehlen häufig Informationen über die individuelle oder emotionale Relevanz oder situationsspezifische Bedeutungszusammenhänge. Teilweise fehlen Personen, die einen Talker zur Kommunikation verwenden, ausreichende kommunikative Erfahrungen (Light/McNaughton 2014), um beispielsweise beim Setzen eines neuen Gesprächsthemas zu antizipieren, was für die andere Person interessant sein könnte oder was als bekannt vorausgesetzt werden kann. Die Annahme eines gemeinsamen Sets von Gesprächsstrategien und Methoden, die von den Gesprächspartner*innen geteilt werden und die die Produktion und Interpretation von sozialem Handeln organisieren, muss daher zumindest manchmal infrage gestellt werden (Wilkinson 2019).

Gegenüber konventionellen Gesprächssituationen stellen Gespräche unter Beteiligung eines Talkers ein kommunikatives Setting dar, in dem die Nutzer*innen natürlicher verbaler Sprache Vorteile in der Äußerungsproduktion besitzen. Zugleich kommt den weiteren Kommunikationsmodi eine besondere Bedeutung zum Verstehen einer Äußerung zu. Durch die Nutzung eines Talkers werden also

erwartbare Interaktionsvollzüge erheblich irritiert[1] und die Beteiligten müssen erhebliche Interaktionsarbeit leisten, um kommunikative Barrieren zu überwinden (Niediek 2020; Niediek 2022). Die Aufzeichnung und Analyse von Gesprächen per Video bietet hier erhebliche Möglichkeiten, steht aber auch vor Herausforderungen, die nachfolgend skizziert werden.

3. Herausforderungen des spezifischen Erhebungskontextes: Transkription von Mehrpersonengesprächen mit Talkern

Hitzler und Böhringer (2021) haben auf das spezifische Verhältnis von Video und Videotranskript für die Arbeit mit Videodaten in konversations- bzw. interaktionsanalytischen Studien hingewiesen. Die dort beschriebenen Herausforderungen für die Analyse von Videodaten aus natürlichen Erhebungssettings (wie z. B. einem Klassenraum) zeigen sich auch am Gegenstand der Gespräche mit Talkern.

Die Lautstärke der Sprachausgabe eines Talkers muss gegenüber der Lautstärke von Umgebungsgeräuschen angepasst werden. Diese ist aber im Vergleich zur menschlichen Stimme deutlich begrenzt. Individuelle und situative Betonungen oder emotionale Färbungen (Aufregung, Wut, Verzweiflung) können mit konventionellen Sprachausgaben nur begrenzt zum Ausdruck gebracht werden. Synthetische Stimmen lagen bis vor wenigen Monaten (KI erweitert dieses Feld erheblich) für den deutschen Sprachraum nur in geringer Anzahl vor, sodass verschiedene Talker in einer Audioaufzeichnung kaum voneinander unterschieden werden können. Zudem entsteht eine gewisse Latenz bzw. Verzögerung zwischen der Eingabe (direkt auf dem Touch Panel des Geräts, durch Eye-Tracking oder Verwendung externer Schalter und Sensoren) und der Ausgabe, sodass die Zuordnung von Bild und Ton bei der Transkription erschwert ist.

Häufig sind Gestik und Mimik der Personen aufgrund motorischer Beeinträchtigungen gegenüber allgemeinen Erwartungen verändert – sodass eine interpretierende Annotation von mimischen Ausdrücken (wie »lächelnd«, »fragender Blick«) nur sehr vorsichtig erfolgen kann.

Der videographische Blick mit mehreren Kameras, wie er in Studien zur Multimodalität kommunikativer Interaktion (z. B. Mondada 2007) häufig eingesetzt wird, kann zusätzliche Daten über den Prozess des Formulierens ermöglichen, wie er bei gesprochener Sprache so nicht möglich ist: Der schwebende Finger über dem Display kann sowohl ein Suchen nach einer passenden Formulierung

1 Das Fachgebiet, das sich der Förderung und Analyse der Kommunikation mit alternativen Kommunikationsmitteln widmet, wird im deutschsprachigen Raum ›Unterstützte Kommunikation‹ genannt. Die internationale Bezeichnung lautet ›Augmentative and Alternative Communication‹. Communication mit einem Talker wird als ›Aided Communication‹ bezeichnet.

oder dem passenden Wort, genauso wie ein innehalten und Zuhören bedeuten. Eine Textzeile erlaubt der unterstützt sprechenden Person, vor dem Auslösen zu prüfen, was das Gerät sprechen wird, sodass sie sich selbst korrigieren kann, bevor Gesprächspartner*innen die Äußerung hören. Wenn also die Benutzung der Talkeroberfläche aufgezeichnet und annotiert werden kann, bietet dies die Möglichkeit, nicht nur das Ergebnis, sondern auch die Produktion des Gesprächs mitzuverfolgen. Bei Mehrpersonengesprächen mit Talker wird dann je eine Kamera für jede Person-Talker-Dyade und eine Kamera für das übergreifende Interaktionsgeschehen benötigt (in Beispiel 1 also vier Kameras und vier Mikrophone). Einige Geräte erlauben das nachträgliche ›Auslesen‹ der Sprachdatenaufzeichnung des Talkers, sodass dieser Transkriptionsschritt entfällt und das stille Ausprobieren, Fehlerkorrekturen usw. quasi in einer eigenen Datensorte aufgezeichnet wird und nachverfolgt werden kann.

Die Aufzeichnungen müssen anschließend synchronisiert und die Transkripte aufeinander abgestimmt werden. Von Tetzchner und Basil (2011) schlagen dazu vor, die Transkription in Spalten zu organisieren, sodass eine Spalte für die hörbaren Äußerungen (Sprachausgabe) und eine Spalte für die Konstruktion (Spracheingabe) auf dem Talker genutzt werden kann.

Auch hier lassen die enormen technologischen Entwicklungen im Bereich der automatisierten Datenanalyse durch künstliche Intelligenz eine Reihe neuer Möglichkeiten für die Aufbereitung und Synchronisierung von unterschiedlichen Datensorten erhoffen.

Verschiedentlich wird aber davor gewarnt (Hitzler/Böhringer 2021; Reichertz 2013), das Ideal des »allumfassenden Blicks« allzu intensiv zu verfolgen, da mehr Daten nicht zwangsläufig zu mehr Erkenntnis führen müssen. Gefordert sind damit vorausgehende methodologische Entscheidungen, welche Dimensionen in die Analyse einbezogen werden sollen (und welche nicht). Aus der Perspektive der Konversationsanalyse sind eben (nur) jene Daten relevant, die durch den Interaktionsverlauf Relevanz erhalten.

Diese grundlegenden Problemstellungen der Analyse von atypischer Kommunikation mit einem Talker stellen eine mögliche Erklärung dafür dar, warum zwar Videoaufzeichnungen und ihre Analyse mit (erweiterten) konversationsanalytischen Zugängen (u. a. Schmitt 2005, 2006; vom Lehn 2018) vor allem international ein übliches methodisches Vorgehen in dem Forschungsfeld darstellen (z. B. Bloch 2011; Clarke/Bloch/Wilkinson 2013; Auer/Hörmeyer 2017; Savolainen 2023). Vorliegende Studien beschränken sich aber häufig auf dyadische Interaktionssituationen zwischen einer Person, die mit einem Talker spricht und einer Person ohne Talker (z. B. Sprachtherapie für unterstützt sprechende Kinder, Eltern-Kind-Gespräche (z. B. Sennott et al. 2016; Sigurd Pilesjo/Noren 2019; Esbensen/Sigurd Pilesjö 2024) und arbeiten häufig mit vorgegebenen Tasks für die Dyade (z. B. Neuvonen et al. 2022), sodass diese Situationen nur eingeschränkt als natürliche Kommunikation im Sinne der sogenannten ›talk-in-interaction

studies‹ (siehe auch Schegloff 1987; Sigurd Pilesjo/Rasmussen 2011; Higginbotham/Engelke 2013) aufgefasst werden können.

4. Spezifischer Analysekontext

Im Folgenden wird entlang von zwei Forschungsbeispielen herausgearbeitet, wie sich diese Herausforderungen auf die Transkription audiovisueller Daten auswirken und wie damit die Transkription zum elementaren Bestandteil der Analyse wird.

Die Beispiele stammen aus Spontandatensammlungen im Rahmen kleinerer Lehr-Lernforschungsprojekte mit Studierenden und einer aus Mitteln der Professur finanzierten Vorstudie. Die Interaktionskontexte, in denen die Daten entstanden sind, wurden also nicht durch eine gezielte Aufgabe für die Interaktionsbeteiligten beeinflusst, insofern lassen sich die Studien als fokussierte Video-Ethnographien (Knoblauch/Vollmer 2022) bezeichnen. Alle Interaktionsbeteiligten wussten um die Aufnahme und haben vorab ihr Einverständnis gegeben. Bei minderjährigen Beteiligten wurde das Einverständnis der Erziehungsberechtigten eingeholt, darüber hinaus wurden alle Namen verfremdet und die Fotogramme bearbeitet. Die sichtbare Kamera und die Anwesenheit von Forscher*innen im Raum haben daher eine unbeabsichtigte, aber unvermeidbare wechselseitige Beeinflussung zur Folge. Diese erscheint aber für die in diesem Beitrag adressierten Fragestellungen weniger erheblich.

Forschungsbeispiel 1: Iterative Transkription

Für das erste Beispiel wurden im Jahr 2018 Interaktionssituationen während zweier Workshop-Wochenenden videografiert, in denen unterstützt sprechende Jugendliche und Erwachsene zusammen Freizeit verbracht oder an Bildungsthemen gearbeitet haben. Insgesamt wurden rund neun Stunden Videomaterial aufgezeichnet. Nach ausführlicher Sichtung und Sequenzierung des Datenmaterials wurde allerdings deutlich, dass nur ein sehr kleiner Teil des Materials mit herkömmlichen Verfahrensweisen für die weitere Analyse genutzt werden kann. Es stellt sich die Frage, wie derartiges Videomaterial erschlossen werden kann, um die sich etablierenden Interaktionsordnungen rekonstruieren zu können. Dazu wurde ein Verfahren entwickelt, das vorläufig als ›iterative Transkription‹ bezeichnet werden soll und im Folgenden anhand einer ›Ankerszene‹ skizziert wird.

Es handelt sich um eine Arbeitsgruppensituation, in der sich die drei Teilnehmer*innen Jonas, Berit und Anna über ihre Erfahrungen mit dem Talker unterhalten (Abb. 1). Diana, Sarah und Rita besitzen keine Kommunikationsbeeinträchtigung und sollen eine assistierende Rolle in der Situation einnehmen.

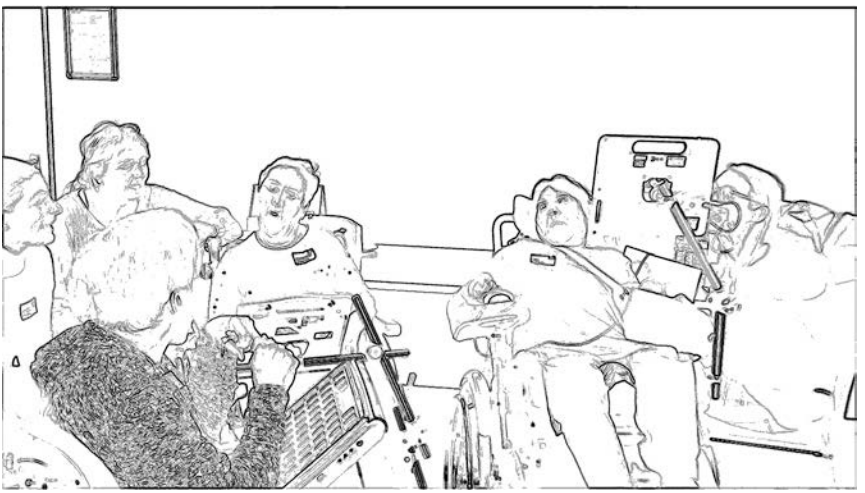


Abbildung 1: Beispiel 1, Fotogramm (verfremdet) | im Hintergrund von links nach rechts: Diana, Rita, Anna, Berit, Sarah, im Vordergrund links: Jonas

Die Sequenz ist durch viele Nebengeräusche und Stimmen gekennzeichnet, weil sich noch weitere Personen mit und ohne Talker im Raum befinden. Die Aufnahme erscheint auch nach wiederholtem Anschauen und Anhören wie ein großes Rauschen. Die Frage »What the hell is going on here?« (Geertz 1983, zit. n. Amann/Hirschauer 1997, S. 20) erscheint zunächst kaum dechiffrierbar. Trotzdem wurde diese Sequenz für eine detaillierte Analyse ausgewählt, weil einzelne Turns und Gesten eine hohe interaktive Dichte insbesondere unter den Talker-Nutzer*innen vermuten ließen. Für eine ausführliche inhaltliche Darstellung der Sequenz sei auf Niediek (2020) verwiesen.

Zunächst wurde ein GAT2-*Minimaltranskript* angefertigt (Selting et al. 2009; vgl. Meyer/Meier zu Verl i. d. B.). Dieses Transkriptionssystem bereitet die Analyse von Gesprächsdaten vor und ermöglicht eine zunehmende Detaillierung im Verlauf der weiteren Analyseschritte.

Tabelle 1: Beispiel 1, Auszug GAT2 Minimaltranskript

Rita	(übergang) manchmal muss auch zwischen (1.7) ↑njea weil du wenn duB braucht zeit ( ) Anna ARbeitet, und in arbeit, (1.9) du nimmst den talker nicht viel oder, das ist praktisch (0.5) (sam tag) jadeswegen (.) reicht es vom akku (.) aber (du) macht (--)ja
Sarah	((unverständlich 4.0))
Berit	*Warum*
Sarah	(wegen der Steck)
Anna	*wenn* (3.3) *ich bin* (3.8) *sommerstadt*
Rita	↑die wissen nicht was sommerstadt ist( )
Diana	was (.) genau was machst du in sommerstadt?
Rita	( ) das ist eine erholungs un(unverständlich 3.2)
Jonas	Ahhh
Anna	Ö(( ))entlichkeitsarbeit
Diana	nicht deine Arbeit? (2.3) wohnst du in sommerstadt?

Para- und nonverbale Zeichen und Rezeptionssignale werden in GAT2 durch entsprechende Transkriptionskonventionen dokumentiert, Mimik, Gestik und Körperhaltung werden allerdings lediglich beschreibend annotiert (Selting et al. 2009). Dem Vorschlag von Higginbotham und Engelke (2013) folgend wurden Ergänzungen der GAT2-Regeln vorgenommen (vgl. Tab. 1) und mit einem Talker Gesprochenes durch Kursivschrift gekennzeichnet. Das ergänzte GAT2-Transkript ermöglicht einen guten Nachvollzug der gesprochenen Sprache und der Bezugnahmen der Sprecher*innen auf die weiteren Beteiligten. Es fehlen aber alle Äußerungen, die nicht unmittelbar im Zusammenhang mit einer verbalen Äußerung stehen, sondern allein gestisch, mimisch oder durch Bewegung hervorgebracht werden. Das GAT-Transkript zeigt nur erfolgreiche Versuche der Turn-Übernahme, dokumentiert aber nicht, welche gestisch/mimischen oder körperlichen Versuche der Turn-Übernahme insbesondere die talkernden Gesprächsteilnehmer*innen hervorbringen. Zudem geht die Parallelität und Sequenzialität der einzelnen Handlungen verloren.

Für die Analyse dieser Sequenz fehlen im Transkript beispielsweise Informationen über die Beteiligung von Berit, die sich kaum sprachlich äußert, aber – so wird in den weiteren analytischen Schritten deutlich, dennoch zentral für den Interaktionsverlauf ist. Ihre Körperhaltung und Mimik, mit der sie durchaus die Sprecherpositionen der anderen bestätigen, verstärken oder auch untergraben könnte (vgl. Schmitt 2005) wird nicht mit transkribiert.

Für analytische Fragen, die sich den Folgen der Asymmetrie verbaler und non-verbaler Kommunikation bei atypischer Kommunikation widmen und Irritationen, Konflikte und Barrieren, die dadurch hervorgebracht werden in den Blick nehmen, ist daher das verbal-orientierte Transkript nicht ausreichend.

Daher wurde im zweiten Schritt ein Partiturtranskript mit der Transkriptionssoftware ›EXMARaLDA‹ (Schneider 2002) erstellt. Für Reichertz (2013)

ist die Erstellung eines *Partiturtranskripts* ein Arbeitsmittel in der Datenanalyse, das in jeder Hinsicht erweiterbar sein sollte. Die »Transkription eines Bildes [ist demnach] nicht nur eine Einzelanfertigung, sondern auch eine Einzelfallanalyse« (ebd., S. 48). Beschreibungen und zeitliche Zuordnungen wurden in dem Beispiel dementsprechend fortwährend überprüft und gegebenenfalls korrigiert. So können sprachliche und körperliche Aktivitäten aller Interaktionsbeteiligten in ihrer zeitlichen Ordnung erfasst und mehrere Interaktionsstränge in der Gruppe identifiziert werden. Die Transkription des gesprochenen Wortes, ebenso wie die Sprachausgabe der Talker wurde zur Basis genommen und als erste Zeile jedes Transkriptionsblocks (Zeile [v]) gesetzt (vgl. Tab. 2). Die Annotation von Blickbewegung und Mimik folgt in der Zeile [Blick] und Bewegungen des Rumpfes und der Extremitäten in der Zeile [Bewegung]. Räumliche und materielle Gegebenheiten sowie gegebenenfalls Veränderungen des Bildausschnitts und der Kamerapositionierung werden in einer abschließenden Kommentarzeile erfasst.

Durch das Partiturtranskript werden parallele Interaktionsstränge und wechselseitige Bezugnahmen der Interaktionspartner*innen nachvollziehbar. Die Möglichkeiten der Software (langsames, schnelles, frame-by-frame, stummes Abspielen) eröffnet eine präzise Dokumentation der zeitlichen Struktur von gleichzeitigen und nachfolgenden Handlungen (vgl. Knoblauch/Wilke i. d. B.). Durch die Ergänzung von Stills bzw. Fotogrammen bei einer späteren Veröffentlichung lässt sich eine gewisse Näherung an die Videoaufzeichnung erreichen. Die annähernd gleichwertige Behandlung der verschiedenen Zeilen macht allerdings analytische Entscheidungen der Segmentierung komplexer und durch die zeitliche Genauigkeit gerät die interaktive Aufeinander-Bezugnahme aus dem Blick.

Tabelle 2: Beispiel 1, Partiturtranskript

[5]	9 [15.9]	10 [17.0]	11 [19.1]	12 [21.1]
JONAS [Blick]	dreht Kopf zu Rita			schaut zu Rita
JONAS [Bewegung]	tupft sich mit Handrücken der linken Hand den Mund			tupft sich mit
RITA [v]	(übergang)		manchmal muss auch zwischen	(1.7) ↑njea
RITA [Blick]	nickt, schaut zu Anna			schaut zu Anna
RITA [Bewegung]	macht Kreisbewegung mit rechtem Arm		macht Kippbewegung mit rechter Hand	erhebt rechten Arm auf
ANNA [Blick]	Stirnrunzeln		hebt die Augenbrauen, leichtes Mund geöffnet	Nicken, schaut zu Rita
BERIT [Blick]	dreht Kopf in mehreren Teilbewegungen zu Rita		schaut zu Rita	
SARAH [Blick]	nickt mehrfach	schaut zu Rita	blickt auf Blatt Papier rechts neben sich	
SARAH [Bewegung]			lehnt sich zurück, hält Stift in der rechten Hand	beugt sich über das

[6]	..	13 [23.9]	14 [25.2]	15 [26.0]
JONAS [Blick]				
JONAS [Bewegung]	Handrücken der linken Hand über den Mund			
RITA [v]	weil du wenn du	Anna braucht zeit	( )	Anna
RITA [Blick]		schüttelt den Kopf, schaut nach links [außerhalb des Bildes?]	[au-	schaut zu Jonas, nickt
RITA [Bewegung]	halbe Höhe			
ANNA [Blick]				
BERIT [Blick]		dreht Kopf zu Sarah		schaut auf ihr
SARAH [v]		(unverständlich 4.0)		
SARAH [Blick]		hebt den Blick		blickt zu Berit, dreht
SARAH [Bewegung]	Blatt [schreibt?]			beugt sich vor,

[7]	..	16 [28.0]	17 [28.9]
JONAS [Blick]			schaut zu Rita
RITA [v]	ARbeitet,	und in arbeit,	(1.9)
RITA [Blick]		nickt	schüttelt den Kopf, schaut Richtung V
RITA [Bewegung]			bewegt nach oben offene Hand in und
ANNA [Blick]			schaut zu Rita
BERIT [Blick]	Display	schaut zu Anna, öffnet den Mund	schaut zu Rita
SARAH [v]			
SARAH [Blick]	Blick Richtung Display von Berit		schaut zu Rita
SARAH [Bewegung]	verschränkt die Arme tippt mit dem Stift an den Mund		rollt Stift vor dem Mund

[8]	..	18 [30.9]	19 [32.9]	20 [33.7]
JONAS [Blick]				
RITA [v]		du nimmst den talker nicht viel oder,	das ist praktisch	(0.5)
RITA [Blick]	und Berit [?]	schaut zu Anna, nickt mit dem Kopf	nickt zu Anna	schaut zu Anna
RITA [Bewegung]	her	weist mit rechter Hand auf Talker von Anna		
ANNA [Blick]		schaut zu Rita, hebt die Augenbrauen		schaut nach oben,

[8]	..	18 [30.9]			19 [32.9]	20 [33.7]
BERIT [Blick]						
SARAH [Blick]					senkt den Kopf, schaut zu Rita	
SARAH [Bewegung]					stützt Kinn auf der rechten Hand auf	

[9]	..	21 [34.3]	22 [35.5]	23 [37.8]
JONAS [Blick]				
RITA [v]		(sam tag) ja	deswegen (.) reicht es	vom akku (.) aber (du)
RITA [Blick]			schaut zu Sarah?	schaut zu Sarah?
ANNA [Blick]		blin- schaut zelt oben	nach Blick schweift umher	schaut zu Rita, Augenbrauen gehoben
BERIT [Blick]				
SARAH [Blick]			schaut auf Display von Berit	schaut zu Rita
SARAH [Bewegung]				tippt mit dem Stift an ihr Kinn

[10]	..	24 [41.7]	25 [43.7]	26 [44.2]
JONAS [Blick]				dreht Kopf nach rechts, schaut zu Berit oder Sarah [?] atmet
RITA [v]		macht (--) ja		
RITA [Blick]		schaut auf den Boden, nickt leicht	schaut zu Jonas	schaut zu Sarah
RITA [Bewegung]				beugt sich vor, schaut zu Sarah
ANNA [v]				
ANNA [Blick]			schaut auf ihr Display	suchender [?] Blick über das Display
BERIT [v]			WARUM	
BERIT [Blick]		schaut zu Rita		schaut zu Anna
BERIT [Bewegung]		tippt auf Schalter		
SARAH [v]				(wegen der Steck)
SARAH [Blick]				
SARAH [Bewegung]				legt Kinn auf Stift u Hand ab

[11]	..	27 [47.8]	28 [49.5]	29 [53.8]	30 [55.7]
JONAS [Blick]		tief aus	schaut zu Anna		
RITA [Blick]		nickt	lächelt		
RITA [Bewegung]		schaut auf Display von Anna			
ANNA [v]		wenn (3.3)	ich bin	(3.8)	SOMMERSTADT

[11]	..	27 [47.8]	28 [49.5]	29 [53.8]	30 [55.7]
ANNA [Blick]		Blick bewegt sich über das Display			schaut nach
ANNA [Bewegung]					habt linken
BERIT [Blick]			senkt den Blick etwas [Display von Anna?]	aufschaut auf Display von Anna	hebt Blick
SARAH [v]					
SARAH [Blick]		nickt leicht			schaut auf
SARAH [Bewegung]		stützt Kinn auf die Hand			dreht sich

[12]	..	31 [57.3]	32 [58.5]
JONAS [Blick]			
JONAS [Bewegung]		senkt rechten Arm	bewegt rechten Arm auf
RITA [v]			↑die wissen
RITA [Blick]			schaut auf Display von Anna
ANNA [v]			
ANNA [Blick]	oben, blinzelt mehrfach	schaut nach oben, blinzelt, schaut nach unten	schaut auf ihr Display,
ANNA [Bewegung]	Arm halbe höhe, hebt den Daumen [Geste für »gut«?]	klopft sich mit der linken Hand auf die Brust	
BERIT [Blick]	etwas, schaut zu Anna		
SARAH [Blick]	das Blatt rechts vor sich		
SARAH [Bewegung]	nach links zum Arbeitsblatt, öffnet Stift	beginnt auf Papier zu schreiben	

[13]	..	33 [60.4]
JONAS [Blick]		schaut zu Anna
JONAS [Bewegung]	und ab	nimmt rechten Arm an den Mund
RITA [v]	nicht was sommerstadt ist	( )
RITA [Blick]		
RITA [Bewegung]		beugt sich zu Anna
ANNA [Blick]	leichtes Kopfschütteln	schaut nach links, hebt den Kopf, zieht Augenbrauen zusammen [fragender Blick?]
BERIT [Blick]		
SARAH [Blick]		
SARAH [Bewegung]		setzt den Stift ab, hebt den rechten Arm

[14]	34 [62.4]	35 [63.4]	36 [64.3]
RITA [Blick]	auf Annas Display [?]		
RITA [Bewegung]	lehnt sich etwas zurück	beginnt, sich mit der linken Hand im Gesicht zu kratzen	kratzt sich bis zum linken
ANNA [Blick]	Hals leicht übertreckt, schaut zu Diana [?]		dreht sich zum Display ihres
BERIT [Blick]	blinzelt, schaut zu Diana [?]		
SARAH [Blick]			schaut zu Rita oder Diana [?], presst
SARAH [Bewegung]	nimmt linken Arm auf den Schoß	lehnt sich zurück	
DIANA [v]	was (.) genau was	machst du in sommerstadt?	

Die Komplexität des Partiturtranskriptes erschwert es in der Folge, die verschiedenen Haupt- und Nebenstränge der Interaktion zu identifizieren und die sich im Verlauf etablierenden Interaktionsordnungen zu verstehen.

Daher wurde schließlich eine dritte Form der Datenaufbereitung genutzt, die sich stärker an ethnographischen Verlaufsprotokollen orientiert. Sie wird nachfolgend als *Gerichtete Beschreibung* bezeichnet, weil sie einerseits die verbalen Äußerungen abbildet, aber auch Beschreibungen aller para- und nonverbalen Äußerungen enthält, die nicht nur zeitlich, sondern auch sinnhaft aufeinander folgen (vgl. Tab. 3).

Tabelle 3: Beispiel 1, Gerichtete Beschreibung

Rita nickt und schaut zu Anna: (unverständlich) *manchmal muss auch zwischendurch* (macht Kreisbewegung mit dem rechten Arm) *naja weil du wenn du* (schüttelt den Kopf) *Anna braucht zeit* (halb zu Diana gewandt) *Anna arbeitet und in arbeit* (schaut zu Jonas).
Alle schauen zu Rita.
Rita (schaut zu Anna): *du nimmst den talker nicht viel oder praktisch (selten)*
Anna schaut nach oben Augenbrauen angehoben, ein Zeichen für Zustimmung?
Rita: *ja deswegen reicht es vom akku aber du machst ja*
Berit (schaut zu Anna, mit dem Talker): *warum*
Anna (mit dem Talker): *wenn ich bin sommerstadt*. Anna schaut nach oben, tippt sich mit dem Finger auf die Brust.
Rita: *die wissen nicht was sommerstadt ist*. Rita schaut zu Anna, schüttelt leicht mit dem Kopf.
Anna schaut zu Diana, runzelt die Stirn.
Diana: *was, genau was machst du in sommerstadt*.
Anna schaut zu Diana, dann auf das Display ihres Talkers
Rita (schaut erst zu Diana, dann zu Jonas): *das ist eine erholungs- und* (unverständlich)
Anna schaut zu Rita dann zu Diana, Rita schüttelt leicht den Kopf, sagt etwas Unverständliches zu Anna.
Anna (mit dem Talker): *Ö(ffentlichk)eitsarbeit*.
Berit und Jonas schauen zu Anna und Rita, Sarah macht Notizen.

Die Sequenz erweist sich schließlich als hochkomplexe Interaktionspraxis (vgl. Niediek 2020), die von den Beteiligten kompetent und durch individuell sehr unterschiedliche Beiträge etabliert wird. So wurde durch die Analyse des hier vorgestellten Beispiels deutlich, wie Berit, die massive motorische Einschränkungen hat, zunächst über längere Zeit die Turn-Übernahme nicht gelingt, weil Rita und Diana, die mit dem Mund sprechen, Anna in eine Nebensequenz zur Ergänzung und Erklärung verwickeln, welche Bedeutung »sommerstadt« hat. Erst nachdem Berit eine Nebensequenz mit ihrer Assistentin Sarah vollzieht, kann diese durch einen Blick zu Sarah diese wiederum zur Turn-Übergabe an Berit motivieren, sodass Berit schließlich ihren Redebeitrag platzieren kann. Allerdings nicht, indem sie selbst spricht, sondern indem Sarah zunächst für sie spricht. Schließlich gelingt es Berit, das Gespräch durch das getalkerte Wort »Warum« wieder auf das ursprüngliche Thema zurückzuführen (warum Anna ihren Talker nicht am Arbeitsplatz verwenden kann). Die Gerichteten Beschreibungen ermöglichten, den Sequenzverlauf der verschiedenen Nebengespräche zu rekonstruieren.

Diese Gerichtete Beschreibung (Tab. 3) stellt eine »Justierungsleistung« (Breidenstein et al. 2020, S. 46) dar, die weitere analytische Schritte ermöglicht, weil erst die (vorläufige) Identifikation mehrerer, parallel verlaufender Interaktionen eine weitergehende Analyse ermöglicht. Sie stellt einen Versuch dar, einen Interaktionsverlauf zu rekonstruieren und sich auf die dafür relevanten Daten zu konzentrieren. Damit enthält diese Form der Aufbereitung im Vergleich zu den beiden vorausgegangenen Transkriptionsformen die größten interpretativen Anteile. Dabei werden bewusst Informationen ausgelassen und Daten sanft geglättet, um einen ersten verstehenden Nachvollzug des Geschehens zu ermöglichen. Anders als in klassischen ethnografischen Studien teilnehmender Beobachtung wird aber eine Gerichtete Beschreibung erst durch die vertiefte Auseinandersetzung mit den vorhergehenden Transkriptionen möglich und kann durch diese validiert werden. Erst auf der Basis des sorgfältigen Partiturtranskripts lassen sich einzelne Interaktionsfelder rekonstruieren und für eine weitergehende Interpretation nutzbar machen. So können die Beteiligungen von Akteur*innen auch jenseits von verbalen oder getalkerten Beiträgen rekonstruiert und die vorfindbaren Interaktionsordnungen herausgearbeitet werden.

Forschungsbeispiel 2

Die zweite Sequenz entstand kurz vor den Sommerferien im Jahr 2023 in einer Klasse der Sekundarstufe 1 an einer Förderschule mit dem sonderpädagogischen Schwerpunkt geistige Entwicklung. Die Aufnahme ist in einem Lehr-Lernforschungsprojekt entstanden, in dem Studierende jeweils eine unterstützt kommunizierende Fokus-Schüler*in und ihre alltäglichen Interaktionssituationen in den Blick genommen haben. Die nachfolgende Sequenz entstand in der

Frühstückspause im Klassenraum. Die Schüler*innen sitzen an ihren Pulten, die Lehrkräfte am Pult oder daneben. Der Kameraausschnitt ist in diesem Beispiel auf die Schülerin Valerie (SU, vorne links im Bild) und ihre unmittelbaren Sitznachbarn (S1, hinten links und S2 rechts im Bild) gerichtet (vgl. Abb. 2). Die beiden beteiligten Lehrkräfte (LK Frau Mager und LM männlich), sind nicht auf dem Video zu sehen.

Abbildung 2: Beispiel 2, Fotogramm (verfremdet)

Frau Mager verwendet in der Sequenz den alten Talker von Valerie, hat aber selbst keine Kommunikationsbeeinträchtigung. Dies erschwert, wie bereits zuvor beschrieben, die klare Zuordnung von Talker-Äußerungen und Geräuschen zum Bild. Sie muss durch die zeitliche Abfolge erschlossen werden. Daher wurde auch hier zunächst ein GAT/Partiturtranskript mithilfe der Transkriptionssoftware EXMARaLDA erstellt. In der nachfolgenden Transkriptpassage sind die verbalen Äußerungen der beteiligten Personen und die der beiden Talker eingeblendet (vgl. Tab 4). Die Verbalzeilen [v] und die Talkerzeilen [t] wurden kursiv gesetzt. Sprachliche Äußerungen der Talker wurden in Großbuchstaben gesetzt. Wortäußerungen von weiteren Schüler*innen im Klassenraum, die zufällig aufgezeichnet wurden, aber nicht dem Gesprächsverlauf zugeordnet werden konnten, sind ebenso ausgeblendet wie die Partiturzeilen zur Mimik und Gestik von Schüler S1 und S2, da sie aufgrund der Kamerapositionierung unvollständig bleiben mussten und daher keinen sinnvollen analytischen Beitrag leisten.

Die Sequenz beginnt (vgl. Tab. 4) mit einem Kommentar des Mitschülers S1, nachdem Frau Mager ihren Talker aufgrund eines Softwareproblems neu starten musste.

Tabelle 4: Beispiel 2, selektives Partiturtranskript

[1]	0 [00:00.0] 1 [00:12.9]	2 [00:13.4]
SU [v]		((schmatztendes
SU [m]	schaut zunächst kauend und mit neutralem Gesichtsausdruck nach vorn zu den Lehrkräften,	
SU [g]	sitzt zurückgelehnt auf ihrem Stuhl, hält ihr Brot in der Hand, beugt sich da beugt sich vor	

[2]	.. 3 [00:15.7]	4 [00:17.3]
SU [v]	Geräusch))((schmatzendes Geräusch))	((schmatzendes Geräusch))
SU [t]	*Biep*	
SU [m]	dann schaut sie konzentriert auf den Talker, dabei kaut sie weiter ihr Brot	
SU [g]	tippt mit dem rechten Zeigefinger auf die Talkeroberfläche	
S1 [v]		äh frau mager braucht doch

[3]	..	5 [00:24.1]	6 [00:26.2]
SU [v]	*(mmh, mh)*		
SU [t]		*piep*	*SPAGHETTI*
SU [m]			schaut
SU [g]			bedient mit
LK [v]		*(natürlich brauch) ich nen talker*	
S1 [v]	*keinen talker der alte kann (auf Schrott) geräumt werden*		

[4]	..	7 [00:27.9]
SU [t]		*biep biep*
SU [m]	konzentriert auf den Talker	schaut auf ihren Talker, kaut dabei langsam
SU [g]	ihrem linken Zeigefinger die Sprachausgabe des Talkers	tippt mit dem linken Zeigefinger auf

[5]	.. 8 [00:33.1]
SU [t]	
SU [m]	mit geöffnetem Mund
SU [g]	verschiedene Positionen auf dem Talker
S1 [v]	*ich glaube valerie macht nur wenn (.) (valerie) frau mager talker*

[6]	.. 9 [00:37.1] 10 [00:39.1]	11 [00:40.2] 12 [00:40.6] 13 [00:41.5] 14 [00:42.2]
SU [v]	Hä	
SU [t]	*mööp*	*biep* *mööp*
SU [m]	schaut auf ihren Talker	schaut konzentriert auf ihren Talker, dabei ist ihr Mund leicht
SU [g]		bedient mit ihrem linken Zeigefinger die Sprachausgabe des
LK [v]		*das nächste mal nehm wir_n talker mit zum pommes bestelln*
S1 [v]	*hat*	

[7]	15 [00:44.0]	16 [00:44.0]	17 [00:45.8]
SU [m]	geöffnet	schaut von ihrem Talker auf zur LK, neutraler Gesichtsausdruck	schaut die LK mit neutralem
SU [g]	Talkers	tippt nicht mehr auf ihrem Talker	sitzt am Tisch, Hände nicht
LK [v]	++++++		*und morgen auch (.) talker*

[8]	..	18 [00:47.8]
SU [t]		*ESSICH (0.5)*
SU [m]	Gesichtsausdruck an, Mund leicht geöffnet, senkt dann den Kopf und den Blick	schaut auf ihren
SU [g]	sichtbar, Brotdose geöffnet	Hände nicht sichtbar,
LK [v]	*Mitnehm*	

[9]	..	19 [00:49.2*]20 [00:49.3*]	
SU [t]		*biep*	*biep SPAGHETTIESSICH*
SU [m]	Talker		Schaut Richtung LK, grinsen?
SU [g]	Auslösen der Sprachausgabe mit linker Hand?		Hand, greift den Verschluss der Brotdose

[10]	21 [00:51.0]	22 [00:53.0]
SU [v]	((lachen))	
SU [m]	blickt hoch, beißt in den Verschluss der Brotdose	blickt auf, grinst breit, schaut seitlich rechts
SU [g]	senkt den Kopf zur Brotdose	hebt rechten Arm, ein Stück Brot in der Hand
LK [v]		*bist du sicher dass es das is was du sagen wolltest*

[11]	23 [00:55.9]	24 [00:58.8]	25 [01:00.8]
SU [v]	((schmatzendes Geräusch))		
SU [m]	kaut, neutraler Gesichtsausdruck schaut auf ihr Getränk, kaut dabei langsam		trinkt ihr Getränk und
SU [g]	Brot wechselt in die linke Hand	greift zu ihrem Getränk und trinkt	
S1 [v]		*hä nimmt er morgen auch talker mit (0.5) auch bei eisdiele*	

[12]	..	26 [01:02.8]
SU [m]	schaut dabei in Richtung Kamera schaut beim Trinken erst nach vorn zur LK, dann auf ihren Tisch,	
SU [g]		stellt ihr Getränk ab
S1 [v]		
S2 [v]		*ja ER nimmt auch*

[13]	..	27 [01:04.0]	28 [01:05.0]
SU [m]	neutraler Gesichtsausdruck	schaut wieder zur LK	schaut erst zur LK und
SU [g]			schüttelt ihre rechte Hand
LK [v]		*wir hatten den auf dem sportplatz nich und valerie konnte leider nich sagen dass*	
S2 [v]			

[14]	..	29 [01:09.3]30 [01:10.8]	31 [01:11.6]32 [01:11.6]
SU [m]	dann auf ihr Brot, scheint zuzuhören	schaut erst auf ihr Brot, dann zur LK	schaut nach vorn auf einen
SU [g]	aus, führt dann ihr Brot zum Mund	beißt von ihrem Brot ab	
LK [v]	*sie gerne pommes essen möchte*	*das war blöd*	
LM [v]			das war voll peinlich
S1 [v]		*ja das weiß ich*	

[15]	..	33 [01:13.4]34 [01:13.6]	35 [01:14.4]36 [01:15.6]
SU [v]	hm		
SU [m]	unbestimmten Punkt kaut, schaut nach vorn und unten, scheint zuzuhören		
SU [g]		lässt die Hand, in der sie das Brot hält, sinken neigt den Kopf zweimal leicht nach	
LK [v]		*da hat sie sich hinter mir versteckt*	
LM [v]			*mhm*

[16]	..	37 [01:16.5]	38 [01:20.2]	39 [01:22.1]
SU [m]				
SU [g]	rechts unten, dann wieder nach vorn	hält den Kopf gerade	neigt den Kopf leicht zur rechten Seite, hält	
LK [t]			*ich esse*	
LM [v]		*oha warum macht er das*		
S1 [v]				
S2 [v]				*keine ahnung warum ER es auch macht*

[17]	.. 40 [01:24.5]	41 [01:25.1]42 [01:25.2]	43 [01:26.2]44 [01:26.5]45 [01:26.6]	46 [01:27.2]
SU [m]				
SU [g]	ihn dann wieder gerade	greift sich mit der linken Hand ins Haar	streicht sich mit der linken	
LK [t]			*spa::ghetti*	
LM [v]	((kurzes, schnaubendes Lachen))			
S1 [v]		((lacht laut))		*warum er es sie macht*
S2 [v]				

[18]	47 [01:28.5]48 [01:28.6]49 [01:29.1]	50 [01:30.1]51 [01:30.2]	52 [01:32.1]
SU [m]			
SU [g]	Hand durch ihr Haar, zieht leicht an einer Strähne	lässt die Strähne los	lässt die linke Hand
LK [v]	*ja warum sie das macht*		
LK [t]		*ich es spa::ghettieis*	
LM [v]	*genau*		

[19]	53 [01:33.2]	54 [01:33.5]	55 [01:34.5]	56 [01:37.0]
SU [m]		hört auf zu kauen und schaut aufmerksam zur LK	schaut aufmerksam nach vorn, Mund	
SU [g]	wieder sinken	beide Hände liegen auf ihrem Schoß, Kopf ist leicht nach rechts geneigt		
LK [t]		*morgen (0.4) ich esse spa::ghettieis morgen*		
LM [v]				*ja ?* (0.4) *SPA::ghettieis ?*

[20]	..	57 [01:39.2]	58 [01:41.1]
SU [m]	leicht geöffnet	kaut langsam	kaut, schaut zur LM
SU [g]		kratzt sich mit der linken Hand im Gesicht	streicht sich mit der linken Hand durch
LM [v]			*ja hat er doch gesahagt (0.3) spaghettieis*

[21]	59 [01:44.1]	60 [01:44.2]	61 [01:45.3]	62 [01:47.6]	63 [01:48.8]
SU [m]				schaut auf ihren Talker, dann nach vorn, kaut dabei mit	
SU [g]	die Haare	hält ihr Brot in der rechten Hand	beugt sich nach vorn, tippt mit dem rechten Mittelfinger		
LM [v]		((unverständlich))	*aber spa::ghettieis ist tatsächlich +++*		
S1 [v]		lacht laut			

Das Beispiel verweist auf die besondere Bedeutung der beiden Talker in dieser Interaktionssequenz: Die Frage, welcher »biep« und welche Äußerung welchem Gerät und damit welcher Sprecherin zugeordnet wird, erfolgt über die sequenzielle Struktur der beobachtbaren mimischen und gestischen Aktivitäten von Valerie. Bieps und Äußerungen, denen keine Bewegung von Valerie zugeordnet werden können, so die Schlussfolgerung, müssen zum Talker der Lehrerin Frau Mager zugeordnet werden, weil keine anderen Talker im Raum waren. Dennoch bleibt diese Zuordnung fehleranfällig, da nicht jede Geste von Valerie eindeutig als Berührung des Displays identifiziert werden kann. Umgekehrt wird über das akustische Signal »biep«, das dem Tippen auf ein Feld nachfolgt, die Zuordnung bestätigt. Ton und Bild sind also wechselseitig aufeinander zu beziehen.

Der Blick auf den Gesprächsverlauf zeigt, dass der Einsatz eines Talkers nicht nur einfach die menschliche Stimme ersetzt, sondern einen eigenen Status erhalten sollte, der in diesem Beispiel durch die separate »Talker«-Zeile im Partiturtranskript eingelöst wird: Während Schüler S1 und Frau Mager zunächst über die Talker als Geräte oder Hilfsmittel im Alltag im Sinne einer Sache sprechen (Segment 4–5), nutzt Valerie den Talker, um ein neues Thema zu eröffnen (Segment 6–20). In Segment 9–13 wird einerseits durch das »möp«-Geräusch des Talkers, andererseits durch die paraverbale, gestische und mimische Reaktion von Valerie deutlich, dass das Gerät auf ihren Input scheinbar nicht in der erwarteten Weise reagiert. Zu vermuten ist, dass sie eine alternative Strategie entwickelt, um die gewünschte Äußerung zu produzieren. Der akustische Output »ESSICH« in Segment 18 scheint ihr ausreichend nahe dem zu sein, was sie sagen möchte, sodass sie den Befehl gibt, noch einmal die gesamte Aussage »SPAGHETTIES-SICH« in Segment 20 zu produzieren. Ihre anschließende Belustigung lässt vermuten, dass ihr klar ist, dass das gesprochene Wort nur so ähnlich klingt, wie gewollt, aber eine ganz andere Bedeutung hat. Während Frau Mager sie auf den Fehler hinweist, übergeht der Schüler S1 das vermeintliche Missverständnis und bleibt bei dem vermutlich vorrangig gemeinten Thema »Spaghettieis«. Die zweite Lehrkraft greift das Thema in gewandelter Form auf, indem er darauf hinweist, wie problematisch es für Valerie ist, wenn sie den Talker nicht dabei hat, um

sagen zu können, was sie möchte. Frau Mager formuliert nun ihrerseits auf dem Talker eine Aussage (Segment 43–55) »ich esse Spa::ghettieeis morgen«, allerdings gelingt es ihr ebenfalls nicht, das Wort Spaghettieis in der üblichen Prosodie mit dem Gerät zu produzieren, was ihr Kollege entsprechend kommentiert.

Die Sequenz verdeutlicht, dass der Transkription von Talker-Stimmen eine besondere Bedeutung zukommt, die einerseits in der Zuordnung des akustischen Outputs zu einem oder einer Sprecher*in liegt. Es eröffnet sich eine sequenzielle Struktur in der wechselseitigen Verschränkung von Bild und Ton – während bei verbaler Sprache die Mundbewegung und die gesprochene Sprache weitgehend ineinander fallen.

Andererseits zeigt sich, dass auch bei der Transkription einer Talker-Stimme die Aussprache und Prosodie berücksichtigt werden muss. Die Sprachausgabe gibt Wörter zum Teil verzerrt oder in ungewohnter Sprachmelodie wieder und beeinflusst so den Gesprächsverlauf, sowohl im Hinblick auf die Notwendigkeit der Co-Konstruktion von Inhalten (Renner et al. 2019) aber auch im Hinblick auf Änderungen im thematischen Verlauf des Hauptthemas.

5. Folgerungen für die Bedeutung der Transkription von Mehrpersonen-Gesprächen mit Talker

Die generellen Probleme der Arbeit mit videographischen Daten machen es auch in der Analyse von Gesprächen mit Talker erforderlich, dem Transkript einen epistemischen Status beizumessen, wie Ayas (2015) fordert. Jede Transkription setzt analytische Entscheidungen voraus, durch die eine selektive und vom jeweiligen Erkenntnisinteresse geleitete Darstellungsform entsteht, die produktiv für den weiteren Analyseprozess ist. Die Beispiele zeigen, dass die Notwendigkeit der Ergänzung von Transkriptionskonventionen (von Tetzchner/Basil 2011) weitergehende methodologische Fragen aufwirft.

So ist zu entscheiden, ob Kommunikationshilfsmitteln ein quasi eigenständiger Status zugeschrieben werden muss – also spricht die Person oder das Gerät. Insbesondere, wenn das Gerät unerwartet oder nicht reagiert, lässt sich der Talker-Output kaum als Sprechakt direkt der Person zuschreiben, sondern es erhält den Status eines nicht-menschlichen Aktanden (Ankrich/Latour 2006). Die Transkription der akustischen Signale des Talkers in einer separaten Partiturzeile ermöglicht eine gewisse Distanzierung, ohne die Zuordnung zu einer Person aufzugeben.

Die Beispiele verdeutlichen, wie Bild und Ton in einem sequenziellen Zusammenhang stehen, sodass auf Videoaufzeichnungen bei Gesprächen mit Talker auch dann nicht verzichtet werden kann, wenn der Fokus der Auswertung letztlich auf den verbalen/akustisch wahrnehmbaren Äußerungen besteht. Je mehr

Personen und Geräte beteiligt sind, desto wichtiger wird auch die gestische und mimische Beteiligung der Akteure. Durch den Telegrammstil, der für Äußerungen mit einem Talker typisch ist, lassen sich die Interaktionsverläufe, gescheiterte Turn-Übernahmen oder auch Wendungen im Gespräch aufgrund von gestischen und mimischen Beiträgen nur durch eine entsprechende Annotation des nonverbalen Interaktionsgeschehens dokumentieren. Zugleich führen mehr Daten auch zu einem Verlust von Verstehensprozessen und zu mehr Intervention des Erhebungsdesigns in das Interaktionsgeschehen.

Hitzler und Böhringer (2021) weisen darauf hin, dass der Anspruch vollständiger Transkription weder einlösbar sei noch mit den Prämissen der Konversationsanalyse erforderlich wäre. Sie unterscheiden Arbeits- und Darstellungstranskript: Während das Arbeitstranskript die Präzisierung einer am Videomaterial gemachten Entdeckung ermöglicht, fokussiert das Darstellungstranskript auf das Phänomen (mit entsprechender Reduzierung von Komplexität gegenüber der videographierten Situation).

In dem ersten Forschungsbeispiel wurde diese Grundidee aufgegriffen, indem ein Prozess mehrfacher Transkriptionsgänge vorgestellt wurde, der auch als *iterative Transkription* (Niediek et al. 2022) bezeichnet werden kann. Der Schritt der *Gerichteten Beschreibung* hat allerdings nicht nur Darstellungsfunktionen, sondern stellt selbst ein analytisches Mittel dar, indem er Horizonte für die weitere Analyse aufspannt. Es geht um die Identifikation von Haupt- und Nebensequenzen, Bestätigungen und Korrekturen, die gleichzeitig stattfinden, die aber in der Gerichteten Beschreibung zeitweilig ausgeblendet werden, um einzelne Ebenen für die weitere Analyse hervorzuheben. Das ausführliche Partiturtranskript stellt dafür parallel zum Video die zentrale Basis dar.

Es zeigt sich, dass die Möglichkeiten der Datenanalyse wesentlich durch die Form der Datenaufbereitung und Transkription geprägt sind. Daher muss ihr in der Analyse atypischer Gesprächssituationen eine besondere Aufmerksamkeit gewidmet werden und umfangreiche zeitliche und personelle Ressourcen eingeplant werden. Unter den Bedingungen der Forschungsökonomie drittmittelfinanzierter Forschung ist das oft nur schwer zu realisieren. Auch dies mag eine Erklärung dafür sein, warum atypische Gespräche bislang nur selten Gegenstand qualitativ-rekonstruktiver Forschung ist.

Allerdings eröffnet die Auseinandersetzung mit den Daten mögliche Beiträge für die Weiterentwicklung von analytischen Zugängen zur Organisation atypischer Gespräche, sodass weitere Schritte in diese Richtung ertragreich erscheinen. Der Blick auf außergewöhnliche Gesprächssituationen kann damit nicht nur einen Beitrag zum Verstehen von Teilhabe- und Ausschlussprozessen in Interaktionssituationen haben, sondern möglicherweise auch einen Beitrag zur Methodenentwicklung leisten.

Literatur

Amann, Klaus/Hirschauer, Stefan (1997): Die Befremdung der eigenen Kultur. Ein Programm. In: Hirschauer, Stefan/Amann, Klaus (Hrsg.): Die Befremdung der eigenen Kultur. Zur ethnographischen Herausforderung soziologischer Empirie Frankfurt am Main: Suhrkamp, S. 7–52.

Ankrich, Madleine/Latour, Bruno (2006): Zusammenfassung einer zweckmässigen Terminologie für die Semiotik menschlicher und nicht-menschlicher Konstellationen. In: ANThology. Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie. Bielefeld: transcript, S. 398–405.

Antaki, Charles/Crompton, Rebecca J./Walton, Chris/Finlay, W. M. L. (2017): How adults with a profound intellectual disability engage others in interaction. In: Sociology of Health & Illness 39, H. 4, S. 581–598. https://doi.org/10.1111/1467-9566.12500.

Auer, Peter/Bauer, Angelika/Hörmeyer, Ina (2020): How can the ›autonomous speaker‹ survive in atypical interaction? The case of anarthria and aphasia. In: Wilkinson, Ray/Rae, John, P./Rasmussen, Gitte (Hrsg.): Atypical Interaction. The Impact of Communicative Impairments within Everyday Talk. Cham: Palgrave Macmillan, S. 373–408.

Auer, Peter/Hörmeyer, Ina (2017): Achieving Intersubjectivity in Augmentative and Alternative Communication (AAC): Intercorporeal, Embodied, and Disembodied Practices. In: Meyer, Christian/Streeck, Jürgen/Jordan, J. Scott (Hrsg.): Intercorporeality: Emerging Socialities in Interaction, Foundations of Human Interaction. New York: Oxford Academic, S. 323–360. https://doi.org/10.1093/acprof:oso/9780190210465.003.0013.

Ayas, Ruth (2015): Doing Data: The Status of Transcripts in Conversation Analysis. In: Discourse Studies 17, H. 5, S. 505–528.

Bloch, Steven (2011): Anticipatory other-completion of augmentative and alternative communication talk: a conversation analysis study. In: Disability & Rehabilitation 33, H. 3, S. 261–269. http://doi.org/10.3109/09638288.2010.491574.

Breidenstein, Georg/Hirschauer, Stefan/Kalthoff, Herbert/Nieswand, Boris (2020): Ethnografie. Die Praxis der Feldforschung. 3. Auflage. München: UVK.

Clarke, Michael (2016): Co-construction, asymmetry and multimodality in children's conversations. In: Smith, Martine/Murray Janice (Hrsg.): The Silent Partner? Language, Interaction and Aided Communication. Havant: J & R Press, S. 177–198.

Clarke, Michael/Bloch, Steven/Wilkinson, Ray (2013): Speaker Transfer in Children's Peer Conversation: Completing Communication-aid-mediated Contributions. In: Augmentative and Alternative Communication 29, H. 1, S. 37–53. https://doi.org/10.3109/07434618.2013.767490.

Clarke, Michael/Kirton, Andrea (2003): Patterns of interaction between children with physical disabilities using augmentative and alternative communication systems and their peers. In: Child Language Teaching & Therapy 19, S. 135–151. https://doi.org/10.1191/0265659003ct248oa.

Esbensen, Annette/Sigurd Pilesjö, Maja (2024): Children's Reformulations During Speech-Language Assessment: Ways to Show Competence in Children With Developmental Language Disorder«. In: Communication and Medicine 19, H. 2, S. 126–136. https://doi.org/10.1558/cam.23829.

Fornefeld, Barbara (2010): Alle reden von Bildung für alle – Sind alle noch gemeint? Bildungsanspruch für Menschen mit Komplexer Behinderung. In: Musenberg, Oliver/Riegert, Judith (Hrsg.): Bildung und geistige Behinderung. Bildungstheoretische Reflexionen und aktuelle Fragestellungen. Oberhausen: Athena, S. 260–281.

Goodwin, Charles/Schlegloff, Emanuel A./Ochs, Elinor et al. (1995): Co-Constructing Meaning in Conversations with an Aphasic Man. In: Research on Language and Social Interaction 28, H. 3, S. 233–260.

Griffiths, Sarah/Barnes, Rebecca/Britten, Nicky/Wilkinson, Ray (2015): Multiple repair sequences in everyday conversations involving people with Parkinson's Disease. In: International Journal of Language and Communication Disorders 50, H. 6, S. 814–829.

Higginbotham, D. Jeffery/Engelke, Christopher R. (2013): A Primer for Doing Talk-in-interaction Research in Augmentative and Alternative Communication. In: Augmentative and Alternative Communication 29, H. 1, S. 3–19. https://doi.org/10.3109/07434618.2013.767556.

Higginbotham, D. Jeffery/Wilkins, David P. (1999): Slipping through the timestream: Social issues of time and timing in augmented interactions. In: Kovarsky, Dana/Maxwell, Madleine/Duchan, Judith, F. (Hrsg.): Constructing (in) competence: Disabling evaluations in clinical and social interaction. London: Psychology Press, S. 49–82.

Hitzler, Sarah/Böhringer, Daniela (2021): »Conversation is simply something to begin with«: Methodologische Herausforderungen durch Videodaten in der qualitativen Sozialforschung am Beispiel der Konversationsanalyse. In: Zeitschrift für Soziologie 50, H. 2, S. 79–95.

Hörmeyer, Ina (2015): Der Einsatz von Körper und Maschine in der Unterstützten Kommunikation. Eine konversationsanalytische Untersuchung. Empirische Kommunikationsforschung im Gesundheitswesen (EKiG). Band 4. Mannheim: Verlag für Gesprächsforschung.

Knoblauch, Hubert/Vollmer, Theresa (2022): Fokussierte Ethnographie. In: Poferl, Angelika/Schröer, Norbert (Hrsg.): Handbuch Soziologische Ethnographie. Wiesbaden: Springer VS, S. 353–366.

Kraat, Arlene W. (1985): Communication between aided and natural speakers. Madison, Wisconsin.

Light, Janice/Collier, Barbara/Parnes, Penny (1985a): Communicative interaction between young nonspeaking physically disabled children and their primary caregivers: Part I – discourse patterns. In: Augmentative and Alternative Communication 1, H. 2, S. 74–83.

Light, Janice/Collier, Barbara/Parnes, Penny (1985b): Communicative interaction between young nonspeaking physically disabled children and their primary caregivers: Part II – Communicative function. In: Augmentative and Alternative Communication 1, H. 3, S. 98–107.

Light, Janice/Collier, Barbara/Parnes, Penny (1985c): Communicative interaction between young nonspeaking physically disabled children and their primary caregivers: Part III – Modes of communication. In: Augmentative and Alternative Communication 1, H. 4, S. 125–133.

Light, Janice/McNaughton, David (2014): Communicative Competence for Individuals who require Augmentative and Alternative Communication: A New Definition for a New Era of Communication? In: Augmentative and Alternative Communication 30, H. 1, S. 1–18. https://doi.org.10.3109/07434618.2014.885080.

Mondada, Lorenza (2007): Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. Discourse Studies 9, H. 2, S. 194–225. https://doi.org.10.1177/1461445607075346.

Neuvonen, Kirsi A./Smith, Martine M./Launonen, Kaisa/von Tetzchner, Stephen (2022): Communication partner strategies in negotiation for meaning in interactions involving aided communication. In: Clinical Linguistics & Phonetics 37, H. 12, S. 1104–1123. https://doi.org/10.1080/02699206.2022.2148130.

Niediek, Imke (2020): Mehr als Worte: Praktiken der Kommunikationsassistenz in gemeinsamen Arbeits- und Lernsituationen. In: Müller, Falco/Munsch, Chantal (Hrsg.): Jenseits der Intention: Ethnografische Einblicke in Praktiken der Partizipation. Weinheim/Basel: Beltz Juventa, S. 184–198.

Niediek, Imke (2022): Kommunikationsassistenz in Gesprächssituationen mit Menschen mit komplexem Unterstützungsbedarf. In: Qualifizierung für Inklusion 4, H. 1. o. S. https://doi.org/10.21248/qfi.81.

Niediek, Imke/Gerland, Juliane/Dobslaw, Gudrun (2022): Teilhabe multimodal. In: Wansing, Gudrun/Schäfers, Markus/Köbsell, Swantje (Hrsg.): Teilhabeforschung – Konturen eines neuen Forschungsfeldes: Beiträge zur Teilhabeforschung. Wiesbaden: Springer VS, S. 281–301. https://doi.org/10.1007/978-3-658-38305-3_15.

Reichertz, Jo (2013): Der bildliche Ausdruck als vertextbare Kommunikation? In: Soziale Welt 64, H. 1/2, S. 37–50. http://www.jstor.org/stable/24754650 (01.08.2024).

Renner, Gregor/Hörmeyer, Ina/Hoffer, Lena (2019): Ko-Konstruktion erkennen und verstehen – eine Analyse verschiedener Ko-Konstruktionstechniken in der Unterstützten Kommunikation. In: Sprache – Stimme – Gehör 43, H. 3, S. e1-e7. https://doi.org/10.1055/s-0043-109146.

Savolainen, Irina (2023): Speaking partners' talk in transitions between symbols in conversations that use speech-output technologies. In: Augmentative and Alternative Communication 40, H. 2, S. 86–99. https://doi.org/10.1080/07434618.2023.2283854.

Savolainen, Irina/Klippi, Anu/Launonen, Kaisa (2020): Coconstructing in Conversations Using a Communication Book. In: Journal of Interactional Research in Communication Disorders 9, H. 2, S. 141–171. https://doi.org/10.1558/jircd.36668.

Schegloff, Emanuel A. (1987): Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis. In: Social Psychology Quarterly 50, H. 2, S. 101–114.

Schmitt, Reinhold (2005): Zur multimodalen Struktur von turn-taking. In: Gesprächsforschung Online, H. 6, S. 17–61.

Schmitt, Reinhold (2006): Videoaufzeichnungen als Grundlage für Interaktionsanalysen. In: Deutsche Sprache 1, H. 1, S. 18–31.

Schmitt, Reinhold (2012): Zur Multimodalität von Unterstützungsinteraktion. In: Deutsche Sprache 40, H. 4, S. 343–371.
Schneider, Werner (2002): Annotationsstrukturen in Transkripten. DV-technische Strukturanforderungen für Annotate exemplifiziert an EXMARaLDA. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion, H. 3, S. 192–236. http://www.gspraechsforschung-ozs.de/heft2002/px-schneider.pdf (22.12.2020).
Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar et al. (2009): Gesprächsanalytisches Transkriptionssystem (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion, H. 10, S. 353–402. http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf (01.08.2024).
Sennott, Samuel C./Light, Janice C./McNaughton, David (2016): AAC Modeling Intervention Research Review. In: Research and Practice for Persons with Severe Disabilities 41, H. 2, S. 101–115.
Sigurd Pilesjo, Maja/Noren, Niklas (2019): Facilitators' use of a communication device following children's aided turns in everyday interaction. In: Journal of Interactional Research in Communication Disorders 10 H. 1, S. 67–98.
Sigurd Pilesjo, Maja/Rasmussen, Gitte (2011): Exploring interaction between a non-speaking boy using aided Augmentative and Alternative Communication and his everyday communication partners. features of turn organization and turn design. In: Journal of Interactional Research in Communication Disorders 2, H. 2, S. 183–213.
Sundqvist, Annette/Plejert, Charlotta/Rönnberg, Jerker (2010): The role of active participation in interaction for children who use augmentative and alternative communication. In: Communication and Medicine 7, H. 2, S. 165–175.
Tuma, René/Schnettler, Bernd/Knoblauch, Hubert (2013): Videographie: Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.
vom Lehn, Dirk (2018): Ethnomethodologische Interaktionsanalyse. Videodaten analysieren und die Organisation von Handlungen darstellen. Weinheim/Basel: Beltz Juventa.
von Tetzchner, Stephen/Basil, Carmen (2011): Terminology and Notation in Written Representations of Conversations with Augmentative and Alternative Communication. In: Augmentative and Alternative Communication 27, H. 3, S. 141–149.
Wickenden, Mary (2011a): Whose voice is that?: Issues of identity, voice and representation arising in an ethnographic study of the lives of disabled teenagers who use Augmentative and Alternative Communication (AAC). In: Disability Studies Quarterly 31, H. 4., o. S. http://dx.doi.org/10.18061/dsq.v31i4.
Wilkinson, Ray (2019): Atypical interaction: Conversation analysis and communicative impairments. In: Research on Language and Social Interaction 52, H. 3, S. 281–299. https://doi.org/10.1080/08351813.2019.1631045.

Digitale Materialitäten in multizentrischen Interaktionen

Mikroethnographische Verfahren zur Untersuchung gruppenförmigen Arbeitens im Tablet-gestützten Unterricht

Matthias Herrle, Matthias Proske, Aline Puzicha und Anne Zimmer

Zusammenfassung: In unserem Beitrag beleuchten wir methodologische und forschungspraktische Herausforderungen, die sich bei der mikroethnographischen Untersuchung des Interaktionsgeschehens im digital-mediatisierten Unterricht stellen. Im Zentrum steht dabei die Frage nach Zugängen zur Produktion unterrichtlichen Wissens in Tablet-gestützten Gruppenarbeiten. Um die Organisation multizentrischer Interaktionen adäquat nachzuvollziehen, erweitern wir die in der Unterrichtsforschung oft übliche Zwei-Kamera-Strategie um gruppenbezogene Aufzeichnungen und Screenrecordings der Tablets. Die so erzeugten Daten ermöglichen es, die Bedeutung und Funktion der dynamischen Materialität digitaler Technologien für die Herstellung von Wissensprodukten detaillierter zu rekonstruieren. Auf diese Weise können tiefe Einblicke in Eigenlogiken schüler*innenseitiger Wissensproduktion gewonnen werden. Allerdings ist die durch das Erhebungssetup gesteigerte Sichtbarkeit des Geschehens für Unterrichtsforschende nicht zu konfundieren mit der wechselseitigen Wahrnehmung der Unterrichtsbeteiligten. Das ›Kamerawissen‹ ist nicht identisch mit dem Teilnehmer*innenwissen, was bei der Analyse stets zu berücksichtigen ist.

Schlüsselwörter: videobasierte Unterrichtsforschung, Interaktionsanalyse, Mikroethnographie, digitale Medien/Materialitäten, Screenrecording

1. Einleitung

Vor dem Hintergrund unseres Forschungsprojekts »Tablets im Unterricht (TabU)«[1] widmen wir uns in diesem Beitrag methodologischen Fragen, methodischen

1 Das Projekt »Tablets im Unterricht (TabU): Soziomediale Organisation unterrichtlicher Wissensproduktion in Phasen gruppenförmigen Arbeitens« (Projektnummer 528534827, Laufzeit 2024–2026) wird von der Deutschen Forschungsgemeinschaft (DFG) gefördert.

Möglichkeiten und forschungspraktischen Herausforderungen, die sich ergeben, will man auf der Basis audiovisueller Daten nachvollziehen, wie digitale Technologien in ihrer Ding- und Medienhaftigkeit (vgl. Proske/Rabenstein/Reis 2024) in das schulische Interaktionsgeschehen einbezogen werden und sich als relevant für Praktiken und Prozesse der Wissensproduktion (vgl. Carvalho i. d. B.) im Tablet-gestützten Unterricht erweisen. Dabei rückt ein Unterrichtsformat in den Blick, bei dem die Nutzung und Einbindung von Tablets besonders im Vordergrund steht: Gruppenarbeiten. Im Unterschied zu klassen*öffentlichen* Formaten, wie Vorträge oder Lehrgespräche, sind gruppenunterrichtliche Settings durch das simultane Prozessieren pluraler Interaktionszusammenhänge im Klassenzimmer gekennzeichnet: Zeitgleich befassen sich Gruppen in variierender Aufmerksamkeitszentrierung mit der Erzeugung von Wissensprodukten. Dabei beziehen sich die Beteiligten in ihrem Interagieren aufeinander und auf digitale Materialitäten, die als Ressourcen, Gestaltungsobjekte und eigensinnige Mitspieler bedeutsam werden können (vgl. Nevile et al. 2014; Davis 2020). Für Ansätze mikroethnographischer Interaktionsforschung (vgl. LeBaron 2008; Herrle 2020) stellt sich für die Untersuchung solcher Settings die Anforderung, das gegenüber traditionellem Unterricht erweiterte Spektrum multimodaler Phänomene videographisch zu erfassen und geeignete Formen im Umgang mit den audiovisuellen Daten zu entwickeln, um nachvollziehen zu können, wie unterrichtliche Wissensproduktion mit digitalen Materialitäten in multizentrischen Interaktionen funktioniert.

Im *zweiten Abschnitt* dieses Beitrags verorten wir zunächst den von uns verfolgten Ansatz mikroethnographischer Unterrichtsinteraktionsforschung im Bereich ethnographischer bzw. videographischer Schul- und Unterrichtsforschung. Im *dritten Abschnitt* beschreiben wir, mit welchen neuen Herausforderungen Forschende bei der videobasierten Erhebung Tablet-gestützten Unterrichts konfrontiert werden und illustrieren, in welcher Weise wir damit in unserem Projekt bei der Aufzeichnung des Interaktionsgeschehens im (Gruppen-)Unterricht umgehen. Im *vierten Abschnitt* schildern wir, wie die vielfältige und komplexe audiovisuelle Datenbasis im Rahmen mikroethnographischer Analyseverfahren zu unterschiedlichen erkenntnisgenerativen Zwecken reduziert werden kann. Wir skizzieren dabei sowohl das Vorgehen bei der Datenanalyse als auch Möglichkeiten der Datenpostproduktion und -aufbereitung, die genutzt werden, um Forschungsbefunde zu generieren und zu dokumentieren. *Abschließend* bilanzieren wir forschungsmethodische Erträge und Herausforderungen.

Projektleitung: Matthias Herrle und Matthias Proske; Projektmitarbeiterinnen: Aline Puzicha und Anne Zimmer sowie Melina Eckhardt, Laurens Gallas, Lukas Horst und Maike Süßmuth.

2. Mikroethnographische Unterrichtsinteraktionsforschung

Mikroethnographie ist Feldforschung. Als Gegenstand fungieren Phänomene der sozialen Welt in ihrer natürlichen Umgebung. Die Datenerhebung basiert auf dem Sammeln von Informationen aus erster Hand, dem Sich-Begeben in authentische Situationen zur Beobachtung von Praktiken im Umgang mit Handlungsproblemen (vgl. Breidenstein et al. 2020, S. 36 f.).

Mikroethnographie ist aber auch laborintensiv. Im Unterschied zu ethnographischen Ansätzen, bei denen der längere Feldkontakt als Ressource zur dichten Beschreibung und Aufschlüsselung der Funktionsweise der beobachteten Praxis fungiert (vgl. ebd., S. 38), werden bei mikroethnographischen Untersuchungen – als Variante »fokussierter Ethnographie« (vgl. Knoblauch 2001) – kurzfristige Feldaufenthalte genutzt, um das Interaktionsgeschehen mithilfe audiovisueller Aufzeichnungen zu dokumentieren und sich in großer Detailschärfe mit Aspekten der multimodalen Konstitution des Interaktionsgeschehens zu befassen (vgl. Erickson 1992). An die Stelle des wiederholten Feldaufenthalts und der Erzeugung von Beobachtungsprotokollen tritt hier die wiederholte Betrachtung audiovisueller Daten und die Realisierung verschiedener Formen der Postproduktion und Datenaufbereitung. Was im Feldaufenthalt zunächst als Entlastung und Erweiterung der begrenzten Wahrnehmungskapazität menschlicher Beobachter*innen in einer komplexen Umgebung erscheint, erweist sich später als Selektionsproblem. So werden Forscher*innen in den von ihnen erzeugten audiovisuellen Daten mit einer Vielzahl wahrnehmbarer Phänomene konfrontiert, was die Frage aufwirft, worauf sie in welcher Weise ihre Aufmerksamkeit richten sollen – und worauf nicht. Insofern sind Aufzeichnungen weniger als Endprodukte der Datenerhebung zu betrachten, die bereits einen Fokus auf bestimmte Geschehensabläufe und ihre Organisation legen. Vielmehr handelt es sich um Rohdaten, die je nach Fragestellung arrangiert, aufbereitet und vor allem wiederholt aus verschiedenen Perspektiven betrachtet werden können (und müssen!), um die Organisation des Geschehens nachzuvollziehen: »The videotape itself is not data. It is a resource for data construction, an information source containing potential data out of which actual data must be defined and searched for« (Erickson 2006, S. 178). Um solche audiovisuellen Rohdaten zu erzeugen, ist es notwendig, möglichst viel von dem zu registrieren, was sich für die Beteiligten im Geschehen als relevant erweist. Insofern wird sich hier an ›Vollständigkeit‹ im konstitutionslogischen Sinne als Maxime der Datenerhebung orientiert.[2] Da vorab allerdings nur begrenzt prognostiziert werden kann, wo sich im

2 Dass dies als Maxime gilt, verweist zugleich darauf, dass realiter immer auch mit Ungenauigkeiten und blinden Flecken umgegangen werden muss, die im Erhebungsalltag entstehen. Umso wichtiger erscheint es, dass Strategien ethnographischen Beobachtens das Geschehen begleiten, womit Phänomene registriert werden können, die sich dem Kamerafokus entziehen und für die Analyse wertvolle Kontextinformationen bereitstellen können.

Klassenzimmer relevante Ereignisse abspielen, wird das Geschehen im Unterricht i. d. R. überblicksartig aufgezeichnet – sich orientierend an der räumlichen Positionierung von Lehrkraft und Schüler*innen (vgl. Herrle/Breitenbach 2016).

Mikroethnographie interessiert sich für die multimodale Organisation von Interaktionen. Dafür erweist sich die audiovisuelle Konservierung des Geschehens als unerlässlich. Besonders einflussgebend für mikroethnographische Zugänge auf der Grundlage audiovisueller Daten sind die seit Mitte des 20. Jahrhunderts im US-amerikanischen Raum sich herausbildenden Ansätze der Context Analysis (vgl. Kendon 1990) und der Conversation Analysis (vgl. Maynard 2013). Untersuchungen, die vor diesem Hintergrund ab den 1960er Jahren zur Interaktionsorganisation im Klassenzimmer durchgeführt wurden (vgl. Cazden 1986), fungieren heute als zentrale Grundlagen der Unterrichtsforschung (vgl. Proske/Rabenstein/Meseth 2021). Die systematische Nutzung audiovisueller Daten zur Rekonstruktion der Interaktionsorganisation im Klassenzimmer hat sich im deutschsprachigen Raum allerdings erst seit den 2000er Jahren im Zusammenhang mit der Akzeptanz, der Erschwinglichkeit und der vereinfachten Handhabbarkeit digitaler Aufzeichnungsgeräte verbreitet. Mittlerweile haben im Bereich qualitativ-rekonstruktiver Unterrichtsforschung verschiedene Ansätze an Kontur gewonnen (vgl. Herrle/Dinkelaker 2016; Corsten et al. 2020). Während etwa Ansätze der dokumentarischen Methode audiovisuelle Unterrichtsaufzeichnungen nutzen, um Erkenntnisse über handlungsleitende Orientierungsmuster zu generieren, werden Videoaufzeichnungen in praxisanalytischen Ansätzen zur Rekonstruktion routinierter Praktiken und Subjektpositionierungen eingesetzt. Zentral für die Mikroethnographie erweist sich dagegen ein Interesse an der sozialen Organisation von Unterrichtsinteraktionen. Praktiken, die Unterrichtsbeteiligte im Umgang miteinander und mit Dingen in ihrer Umgebung realisieren, werden daraufhin untersucht, welche Anforderungen mit ihnen im Interaktionskontext bearbeitet werden.

3. Vorgehensweisen zur Erhebung audiovisueller Daten

In zahlreichen Videoforschungsprojekten zur Interaktion zwischen Lehrkräften und Schüler*innen wird das Unterrichtsgeschehen mit einer »Zwei-Kamera-Strategie« (Beeli-Zimmermann/Wannack/Staub 2020, Abs. 5 ff.) erfasst. Rechnung getragen wird damit der historisch gewachsenen sozial-räumlichen Ordnung im Klassenzimmer und den damit verbundenen Handlungserwartungen. Diese zentrierte Ordnung gruppenförmiger Beschulung, die sich durch eine dauerhaft komplementäre und räumlich frontale Anordnung von Lehrkraft und Schüler*innen auszeichnet (vgl. Caruso 2019, S. 157 ff.), erweist sich als beharrlich, wird allerdings spätestens seit der Verbreitung individualisierter Unterrichtsformate in den letzten 20 Jahren (vgl. Rabenstein et al. 2018) dynamischer. Die

»Dezentralisierung der Ordnung« und die »Vervielfältigung und Ausdifferenzierung von Aktivitätszentren« (Reh et al. 2015, S. 323) im individualisierten Unterricht sowie in Formaten wie Gruppen-, Einzel- oder Stationenarbeit stellt infrage, ob ein adäquater Einblick in die Organisation des Geschehens im Klassenzimmer durch die Verwendung von zwei Kameras, welche sich weiter allein an der räumlichen Gegenüberanordnung von Lehrkraft und Schüler*innen orientiert, erreichbar ist. Da relevante Aktivitäten von Schüler*innen wie Lehrkräften sich dem klassenöffentlichen Geschehen entziehen, scheinen neue Zugänge zur videobasierten Erhebung solcher Aktivitäten notwendig.

Zur Steigerung von Unsichtbarkeiten und des Verborgen-Bleibens von Aktivitäten der Schüler*innen und Lehrkräfte trägt der unterrichtliche Einbezug von mobilen digitalen Endgeräten bei. Betroffen davon sind sowohl die Interaktionsbeteiligten selbst als auch die forschenden Beobachter*innen, was Avgustis und Oloff (2024) als »participant opacity« einerseits und »analytical opacity« andererseits beschreiben (S. 85). Womit genau sich Schüler*innen befassen, ist zwar aus der Distanz ebenso intransparent, wenn sie mit analogen Dingen, wie etwa Schulbüchern oder Heften, hantieren. Durch die Verwendung einer prinzipiell unbegrenzten Anzahl an Apps und Internetseiten, die genutzt und aufgerufen werden können, erweitern mobile Endgeräte den Rahmen möglicher Engagements jedoch erheblich: »Due to their inherent multi-functionality and mobility, smartphones can be used in a variety of individual and joint actions, which are often of low visibility to co-present others« (ebd.). Insofern steht die videobasierte Erhebung des Interaktionsgeschehens im digital-mediatisierten Unterricht vor der Herausforderung, die dynamische Materialität von *mobile digital devices* bereits bei der Erhebung des Geschehens zu berücksichtigen. Um die Aktivitäten Einzelner im Umgang mit digitalen Materialitäten und deren interaktive Organisation auf der Basis von Videodaten nachzuvollziehen, bedarf es eines Erhebungssetups, das über die traditionelle »Zwei-Kamera-Strategie« hinausgeht. Zu berücksichtigen sind dabei die spezifischen Bedingungen, unter denen sich das mediatisierte Alltagsgeschehen realisiert. Dafür sind gute Feldkenntnisse unabdingbar.

In unserem Projekt hat sich gezeigt, dass das Phänomen, für das wir uns interessieren – die Erzeugung von Wissensprodukten in Tabletklassen im Zusammenhang mit Gruppenarbeiten – im Zeitverlauf durch unterschiedliche soziomaterielle und soziomediale Konstellationen zwischen Dingen (z. B. Tablets, Netzwerkinfrastruktur, Beamer), Schüler*innen und Lehrkräften im Klassenzimmer hervorgebracht wird: Einzelaktivitäten von Schüler*innen im Umgang mit ihrem Tablet (Personensphäre), Tablet- und Lehrkraft-begleitete Interaktionen zwischen Schüler*innen in parallel prozessierenden Kleingruppen (Gruppensphäre) und klassenöffentliche Interaktionen zwischen Lehrkraft und Schüler*innen, begleitet ebenfalls durch Tablets sowie durch Beamer-Projektionen (Klassensphäre). Alle drei Konstellationen gehen sowohl mit unterschiedlichen

Anforderungen an das Handeln und (wechselseitige) Wahrnehmen der Beteiligten einher als auch an die registrierende Konservierung dieses Geschehens im Medium audiovisueller Daten. Man kann die drei Sphären zwar analytisch voneinander unterscheiden, was für eine Konzipierung des Erhebungssetups auch sinnvoll erscheint (vgl. Abb. 1). Faktisch aber überlagern sich die Sphären, was insbesondere im Zusammenhang der Datenpostproduktion, -aufbereitung und -analyse zu berücksichtigen (vgl. Abschnitt 4) und bereits während der Erhebung zu reflektieren ist.

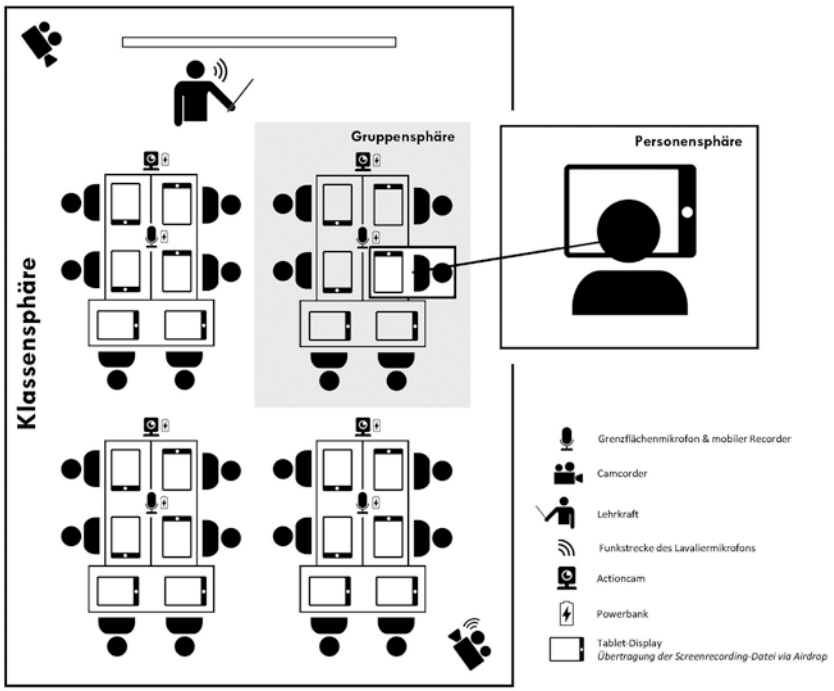


Abbildung 1: Setup zur Erhebung multizentrischer Interaktionen

3.1 Videographie der Klassensphäre

Die Klassensphäre ist dadurch gekennzeichnet, dass das Interagieren der Beteiligten auf der uns interessierenden Vorderbühne des Geschehens klassenöffentlich wahrnehmbar gemacht wird – in unserem Fall z. B. bei Instruktionen zur gruppenförmigen Aufgabenbearbeitung oder der Präsentation von Wissensprodukten, die in Gruppenarbeiten erzeugt wurden. Diese zentrierte Interaktion zwischen Lehrkraft und Schüler*innen kann durch eine herkömmliche »Zwei-Kamera-Strategie« aufgezeichnet werden, um nachzuvollziehen, wer wann mit wem öffentlich wahrnehmbar verbal und körperlich interagiert und wie sich auf

welche klassenöffentlich wahrnehmbaren digitalen Materialitäten (z. B. die Tablet-Projektion über den Beamer) bezogen wird. Die zwei Kameras haben wir auf Stativen gegenüber im Raum angeordnet und auf die Schüler*innengruppe einerseits und die Lehrkraft andererseits ausgerichtet. Zudem haben wir die Lehrkraft mit einem Lavalier-Mikrofon ausgestattet, das per Funksender mit der auf sie ausgerichteten Kamera verbunden ist. Dabei ist es wichtig, dass die Lehrkraft-Kamera durch ein Teammitglied bedient wird, um sie bei Bewegung der Lehrkraft im Raum zu justieren und gegebenenfalls kurz auf Projektionen zu zoomen, damit nachvollzogen werden kann, was dort für die Beteiligten wahrnehmbar ist.[3] Für die Kamera, die auf die Schüler*innengruppe ausgerichtet und die bestenfalls mit einem Weitwinkelobjektiv ausgestattet ist, ist ein Justieren während des Unterrichts i. d. R. nicht erforderlich.

Zugänglich wird auf diese Weise all das, was visuell und akustisch klassenöffentlich geteilt wird. Verborgen bleibt hingegen all das, was jenseits des klassenöffentlich Geteilten auf der Hinterbühne – insbesondere in intimen und persönlichen Distanzverhältnissen – geschieht und mit besonderen Anforderungen für die Etablierung geteilter akustischer, visueller oder haptischer Räume (vgl. Breidenstein 2006, S. 39 ff.) verbunden ist. Auch der Einzug mobiler Endgeräte ins Klassenzimmer geht mit veränderten Möglichkeiten der Realisierung von Selbstbeschäftigung und der Etablierung von Gelegenheiten für »co-ocurring spatially unlimited multimodal interaction« (Sahlström/Tanner/Valasmo 2019, S. 329) einher, die sowohl den forschenden Beobachter*innen als auch der Lehrkraft und anderen Schüler*innen verborgen bleiben.

3.2 Videographie der Gruppensphäre

Die Gruppensphäre ist dadurch gekennzeichnet, dass eine klassenöffentliche Wahrnehmung des (Inter-)Agierens von Gruppenbeteiligten in weiten Teilen aufgehoben ist. In parallel-prozessierenden Kleingruppen ist das Handeln der Schüler*innen in solchen Konstellationen u. a. auf die aufgabenbezogene Herstellung von Wissensprodukten ausgerichtet. Die raum-körperliche Konstellation der Gruppen – konzentrisch auf Stühlen um zusammengeschobene Tische versammelt – ermöglicht den Beteiligten einen geteilten akustischen Raum (vgl. Breidenstein 2006, S. 50 ff.) zu etablieren und nach außen hin abzuschirmen. Ebenfalls wird die haptische Bezugnahme i. d. R. auf die am Tisch verteilten Gruppenmitglieder begrenzt. Visuell können sich die Gruppenmitglieder wechselseitig zwar gut wahrnehmen. Phänomene, die sie auf ihren Tablets in Erscheinung

3 Bei solchen Zooms gilt es genau abzuwägen, wann und für wie lange sie vollzogen werden. Generell sollte allerdings auf ein Zoomen und Schwenken der Lehrkraftkamera verzichtet werden.

bringen, sind allerdings nur für sie selbst und eventuell direkte Platznachbar*innen sichtbar. Ihr Teilen erfordert, dass der Tabletscreen den anderen gegenüber gezeigt oder medial vermittelt wird.

Zwar ist es möglich (und notwendig), mit den beiden Raumkameras zu überblicken, wo Gruppen im Klassenzimmer verortet sind und wie sich im Raum bewegt wird (insbesondere in Bezug auf die Lehrkraft) – insofern ist dies auch eine wichtige Ressource, auf die im Bedarfsfall zurückgegriffen werden kann. Akustisch und visuell kann mit diesem Setup allerdings nicht nachvollzogen werden, wie die Gruppenbeteiligten miteinander und in Bezug auf ihre Tablets interagieren. Um dies nachzuvollziehen, nutzen wir in unserem Projekt ultra-weitwinklige Action Cams, die wir vor den Gruppentischen auf kleinen Stativen montiert haben. Um zu registrieren, wie und worüber dort gesprochen wird, haben wir die Gruppentische zudem mit mobilen Audiorekordern und Grenzflächenmikrofonen ausgestattet. Sowohl die Video- als auch die Audioaufzeichnungsgeräte werden batteriebetrieben (mithilfe von Powerbanks) eingesetzt.

Zugänglich wird auf diese Weise all das, was akustisch sowohl gruppenöffentlich und auch im intimen Bereich zwischen benachbarten Gruppenmitgliedern gesprochen wird. Visuell können Gestik, Mimik, Körper- und Blickausrichtung der Beteiligten sowie ihr Hantieren mit Dingen beobachtet werden. Verborgen bleibt hingegen zumeist all das, was jenseits des geteilten akustischen und visuellen Raums im persönlichen Territorium der Schüler*innen im Umgang mit ihrem Tablet wahrnehmbar gemacht wird – sowohl für die Forscher*innen als auch für die Gruppenmitglieder selbst.

3.3 Videographie der Personensphäre

Die Personensphäre bezieht sich auf all das, was Schüler*innen innerhalb ihres persönlichen Territoriums im Umgang mit ihrem Tablet (visuell) wahrnehmbar machen. Um nachvollziehen zu können, wie die Schüler*innen mit ihrem Tablet hantieren, welche Apps sie nutzen, was sie in welcher Weise eingeben, was in welcher Weise angezeigt wird und wie sie sich auf das Angezeigte beziehen sowie um verstehen zu können, was davon in welcher Weise zum Interaktionsgegenstand der Gruppe (und später gegebenenfalls der Klassenöffentlichkeit) gemacht wird, ist es notwendig, die Tabletbildschirme aufzuzeichnen.

In unserem Projekt haben wir dafür die betriebssystemeigene Screenrecording-Software genutzt, die auf den iPads der Schüler*innen zur Verfügung steht. Zugänglich wird auf diese Weise all das, was auf der Bildschirmoberfläche der Schüler*innen visuell wahrnehmbar gemacht wird. Tonsignale (etwa von abgespielten Videos oder vom Sprechen der jeweiligen Schüler*innen) können dabei zwar prinzipiell auch erhoben werden. Um allerdings den Speicherplatz der auf diese Weise produzierten Dateien nicht übermäßig zu beanspruchen (und

aufgrund der Zugänglichkeit verbaler Schüler*innenäußerungen durch das Erhebungssetup für die Gruppensphäre), wurde in unserem Projekt auf die Tonaufzeichnung beim Screenrecording verzichtet.

Keine Auskunft erhält man auf Basis von Screenrecordings darüber, auf welche Objekte genau die Aufmerksamkeit der Schüler*innen bei der Fixierung ihrer Tablets ausgerichtet ist. Hierfür wäre es notwendig, die Augenbewegungen der Schüler*innen, etwa mittels Eye-Tracking (vgl. Carvalho; Reitstätter/Pesen/vom Lehn i. d. B.), nachzuvollziehen. Dies würde allerdings das ohnehin technisch anspruchsvolle Erhebungssetup zusätzlich verkomplizieren, die Invasivität der Erhebung steigern und eventuell mit dem Problem verbunden sein, zusätzliche Software auf die Tablets aufspielen zu müssen, was erfahrungsgemäß mit erheblichen Hürden verbunden ist – auch hinsichtlich der Einwilligung in die Datenerhebung.

Screenrecordings und ihre bedarfsweise Synchronisation mit audiovisuellen Daten der Gruppen- und Klassensphäre (vgl. Abschnitt 4.2) versprechen völlig neue Einblicke in die Organisation des individuellen und gruppeninternen Agierens und Interagierens im Zusammenhang der mediengestützten Erzeugung von Wissensprodukten (vgl. Mroz 2014). Interessiert man sich für das Interaktionsgeschehen ›in the wild‹, ist diese Steigerung des Potenzials neuer Einblicke zugleich auch mit der Steigerung des Potenzials an Ungewissheiten und technischer Schwierigkeiten verbunden. Das hat einerseits damit zu tun, dass vorab der Akku der Tablets aufgeladen und genügend freier Speicherplatz vorhanden sein muss, um die produzierte Datei zu generieren und abzulegen. Andererseits ist die Erzeugung eines Screenvideos, das später mit den Daten der Gruppen- und Klassensphäre synchronisiert werden kann, daran gebunden, dass die Schüler*innen das Screenrecording selbsttätig zu Beginn der Gruppenarbeitsphase starten und zudem aktiv einen Synchronisationspunkt[4] kreieren, was wiederum störanfällig ist (z. B. durch die Notwendigkeit, den Bildschirm permanent im On-Modus zu halten). Nach Abschluss der Unterrichtsstunde kann sich die Übertragung der Screenrecording-Dateien als problematisch erweisen, wenn dafür nicht genügend Zeit eingeräumt wird. Nach verschiedenen Versuchen im Rahmen der Projektpilotierung mittels schuleigener und selbst eingerichteter Cloudlösungen, erwies sich die Übertragung der Dateien via AirDrop als geeignete Lösung. Hierzu ist es allerdings erforderlich, pro erhobenem Gruppentisch ein MacBook in räumlicher Nähe der Geräte zu positionieren und die Schüler*innen zu bitten, ihre Dateien (möglichst einzeln) zu senden. Schließlich sollte beachtet werden, dass eine informierte Einwilligung in die Videoerhebung bei der Verwendung

4 Die Produktion eines Synchronisationspunktes ist notwendig, um unterschiedliche Video- und Audiodateien im Nachhinein zusammenzuschneiden. Die Produktion einer sog. Transiente, die in der grafischen Darstellung der Audiospur bei Videobearbeitungsprogrammen klar erkennbar ist, kann z. B. durch ein lautes In-die-Hände-Klatschen erzeugt werden.

von Klassen-, Gruppen- und Bildschirmkamera nicht immer leicht zu erzielen und an ein vertrauensvolles Verhältnis gebunden ist, für dessen Etablierung sich für uns insbesondere die Überzeugung sogenannter Gate-Keeper (insbesondere die Schulleitung, aber auch Fachlehrkräfte) als hilfreich erwiesen hat.

4. Verfahren zur Postproduktion, Aufbereitung und Analyse audiovisueller Daten

Ausgangspunkt mikroethnographischer Analyse von Unterrichtsvideos ist das ethnomethodologische Erkenntnisinteresse, »to discover the ›methods‹ that persons use in their everyday life in society in constructing social reality and also to discover the nature of the realities they have constructed« (Psathas 1968, S. 509). Welche Bedeutung und Funktion digitalen Dingen für Methoden und Praktiken der Erzeugung von Unterrichtsrealitäten zukommt, wie sie als Ressourcen oder Gestaltungsobjekte (vgl. Nevile et al. 2014), als eigensinnige Mitspieler (vgl. Davis 2020) oder als botenhafte Medien (vgl. Krämer 2008) fungieren, ist aus dieser Perspektive eine Frage, die sich auf Basis der videographisch beobachtbaren, handlungspraktischen und semantischen Bezugnahme von Lehrkräften und Schüler*innen im Kontext spezifischer Situationen aufklären lässt. Dafür ist es notwendig, eine Beobachtungsgrundlage zu generieren, die es ermöglicht, jene Phänomene in den Blick zu nehmen, deren Relevanz für das Geschehen durch die Bezugnahme der Akteure im Umgang miteinander und mit (digitalen) Dingen in räumlicher Umgebung erkennbar wird. Vor diesem Hintergrund gilt es, die Komplexität der erhobenen Daten (vgl. Abschnitt 3) erkenntnisinteressenbezogen zu reduzieren. Welche Anforderungen sich dabei konkret für welche Formen der Postproduktion und Datenaufbereitung ergeben, stellen wir im Folgenden im Zusammenhang der Skizzierung methodischer Verfahren der Datenanalyse dar.

Kennzeichnend für die Strukturierung des Analyseprozesses ist eine Bewegung »from the top down – from whole to part« (Erickson 1982, S. 218; vgl. auch Herrle 2020). Zunächst wird versucht, nachzuvollziehen, wie der gesamte Unterrichtsablauf durch die Beteiligten in eine Abfolge einzelner Interaktionssegmente gegliedert wird (Abschnitt 4.1). Dann werden selektiv und geleitet durch die Fragestellung einzelne Segmente in ihrer Organisation analysiert (Abschnitt 4.2), um schließlich die mikrologische Konstitution multimodaler Interaktionspraktiken (Abschnitt 4.3) zu rekonstruieren, die Akteure dort im Umgang mit Anforderungen der unterrichtlichen Herstellung von Wissensprodukten in Bezug aufeinander und auf (digitale) Dinge realisieren. Auf allen Ebenen ist dabei die Durchführung minimaler und maximaler Kontrastierungen (vgl. Glaser/Strauss 2005) leitend, um das auf diese Weise erkennbar werdende Spektrum an Formen der unterrichtlichen Einbettung von Gruppenarbeiten, an Varianten der

ablaufbezogenen Organisation gruppeninterner Wissensproduktion und Praktiken im Umgang mit kontextbezogenen oder kontextübergreifenden Interaktionsanforderungen – insbesondere im Umgang mit digitalen Dingen – auszuschärfen.

4.1 Kontextualisierung von Gruppenarbeiten im Interaktionsablauf von Unterrichtsstunden

Um nachvollziehen zu können, wie Unterrichtsbeteiligte durch die Nutzung verbal-sprachlicher, körperlicher und dinglicher Ressourcen füreinander Differenzen in der Abfolge von Interaktionskontexten im Ablauf von Unterrichtsstunden erkennbar machen, ist es notwendig, »the occurrence of major shifts in participants (in and out of scene), of sustained postural and interpersonal distance configurations, and of major topics and/or speaking/listening activities« (Erickson 2006, S. 184) zu betrachten. Dies dient als Grundlage zur Erzeugung von Interaktionsinventaren, die es ermöglichen, Gruppenarbeiten im Interaktionsablauf von Unterrichtsstunden zu lokalisieren und ihre Einbettung in das Unterrichtsgeschehen zu bestimmen.

Grundlage hierfür ist die Videographie der Klassensphäre (vgl. Abschnitt 3.1) – also die audiovisuelle Aufzeichnung des Agierens der Lehrkraft einerseits und des Agierens der Klasse andererseits. Im Rahmen der Postproduktion synchronisieren wir diese beiden Aufnahmen und erstellen ein zwei-perspektivisches Video. Um Segmente im Ablauf des Interaktionsgeschehens rekonstruieren zu können, ist es hilfreich, sich dieses Video zunächst in beschleunigter Geschwindigkeit (z. B. achtfach) anzusehen. Dabei werden in manchen Fällen bereits Diskontinuitäten in der Positionierung und Ausrichtung der Akteure erkennbar, was als Anzeichen für Kontextwechsel fungieren kann. Mit Blick auf diese Dimension, auf den thematischen Bezug des Geschehens und auf die soziale Adressierung bzw. auf die Muster des Sprecher*innenwechsels wird das Video dann wiederholt in Echtzeit betrachtet, um Einheiten und Differenzen im Ablauf des Geschehens sowie die hierarchische Organisation von Segmenten und Sub-Segmenten nachzuvollziehen, die durch Übergangsphasen miteinander verbunden sind.

Als Form der Datenaufbereitung resultiert daraus eine tabellarische Übersicht über die segmentäre Ablaufordnung des Geschehens. Dabei sollte sich die Benennung von Segmenten an jenen Merkmalen orientieren, die sich in ihnen reproduzieren und die sie von anderen Segmenten unterscheidbar machen, was in einer eigenen Spalte beschrieben werden kann. Festgehalten werden sollte zudem die Dauer von Segmenten und Übergängen sowie erste Auffälligkeiten und Besonderheiten hinsichtlich der Untersuchungsfrage. Für unser Interesse an der Interaktion mit digitalen Dingen erweist es sich z. B. als besonders relevant, zu notieren, welche Art digitaler Technologie zu welchen Phasen in bestimmter Weise

genutzt wird. Der Nachvollzug der segmentären Differenzierung des Unterrichtsgeschehens erweist sich für unser Projekt als wichtig, um zu verstehen, wie gruppenförmiges Arbeiten in den Klassenunterricht eingebettet ist. So hat sich z. B. gezeigt, dass Phasen gruppenförmigen Arbeitens durch Segmente gerahmt werden, in denen vorab die Bereitschaft zur Herstellung von Wissensprodukten klassenöffentlich etabliert wird und im Nachhinein Wissensprodukte klassenöffentlich dargestellt und begutachtet werden (vgl. Herrle/Hoffmann/Proske 2022).

4.2 Arrangements von Schüler*innen, Lehrkräften und digitalen Materialitäten im Ablauf von Gruppenarbeiten

Um nachvollziehen zu können, wie sich das Geschehen innerhalb bestimmter Segmente – d. h. in unserem Projekt in parallel verlaufenden Gruppenarbeiten – organisiert, ist erneut eine detaillierte Betrachtung des Interaktionsablaufs notwendig. Dabei interessiert uns, welche Arrangements bzw. Konstellationen zwischen Schüler*innen, Lehrkräften und digitalen Dingen in den Gruppenarbeiten mehr oder weniger dauerhaft zur Aufführung gebracht werden und den segmentären Ablauf des Geschehens bestimmen. Erneut ist es dabei notwendig, die Aufmerksamkeit auf die Ausrichtung von Körpern und Blicken, auf den thematischen Bezug des Agierens sowie auf die Art und Weise zu lenken, in der sich die Beteiligten aufeinander und auf (digitale) Dinge beziehen. Zudem ist es hilfreich, auf Kontinuitäten und Wechsel jener Phänomene zu achten, die auf den Tabletbildschirmen der Schüler*innen angezeigt werden.

Grundlage für die Analyse sind die Video- bzw. Audioaufzeichnungen der Gruppen- und der Personensphäre (vgl. Abschnitt 3.2 und 3.3). Diese gilt es im Rahmen der Postproduktion zu synchronisieren und ein Video zu produzieren, das sowohl das audiovisuelle Geschehen in der Gruppe als auch die Screenrecordings aller Tablets synchron zugänglich macht (siehe Abb. 2). Auch diese Datengrundlage gilt es beschleunigt, realzeitlich und wiederholt zu betrachten, um Kontinuitäten und Wechsel im Gruppengeschehen identifizieren zu können. Bei der Analyse zu beachten sind dabei allerdings zwei Umstände, die einerseits aus der Art des Erhebungssetups und andererseits aus der spezifischen Sozialität von Gruppenarbeiten resultieren: (a) die Relation von forschungs- und akteurseitiger Wahrnehmung und (b) die simultane Sequentialität von Aktivitäts-/Interaktionssträngen im Geschehen.

(a) Wie bereits mit Verweis auf den Begriff »participant opacity« (Avgustis/Oloff 2024, S. 85) in Abschnitt 3 beschrieben, trägt der unterrichtliche Einbezug mobiler digitaler Endgeräte zur Steigerung von Unsichtbarkeiten im Interaktionsgeschehen bei. So ist für die Schüler*innen in Gruppenkonstellationen i. d. R. nicht erkennbar, was andere Gruppenmitglieder

gerade auf ihrem Tablet machen und wahrnehmen – von direkten Platznachbar*innen abgesehen. Um darauf gruppenöffentlich referieren zu können, ist es notwendig, dass das individuell Wahrnehmbare verbal beschrieben, der jeweilige Tabletbildschirm gezeigt oder der Bildschirminhalt medial geteilt wird. Durch das hier genutzte Erhebungssetup – die Screenrecordings der Tablets – entsteht jedoch eine grundsätzliche Differenz zwischen der Wahrnehmung der Teilnehmer*innen des Geschehens und der Beobachter*innen im ›Labor‹. Diese Differenz lässt sich als Unterschied zwischen »participant opacity« und »analytical transparency« fassen (vgl. Abb. 2). Für die Interaktionsforscher*innen wird sichtbar, welche Wahrnehmungsfelder die einzelnen Schüler*innen auf ihren Tabletbildschirmen erzeugen. Während dieser Umstand forschungsseitig genutzt werden kann, um den Umgang mit digitalen Dingen im Rahmen individueller Aktivitätsstränge – etwa bei Recherchetätigkeiten in Gruppenarbeiten – zu beschreiben, steht diese Ressource den beteiligten Schüler*innen nicht offen. Vielmehr entscheiden sie als Gruppenmitglieder situativ, was von dem für sie selbst Sichtbaren sie in welcher Weise und zu welchem Zeitpunkt als Referenz für das Interaktionsgeschehen in der Gruppe wahrnehmbar machen.

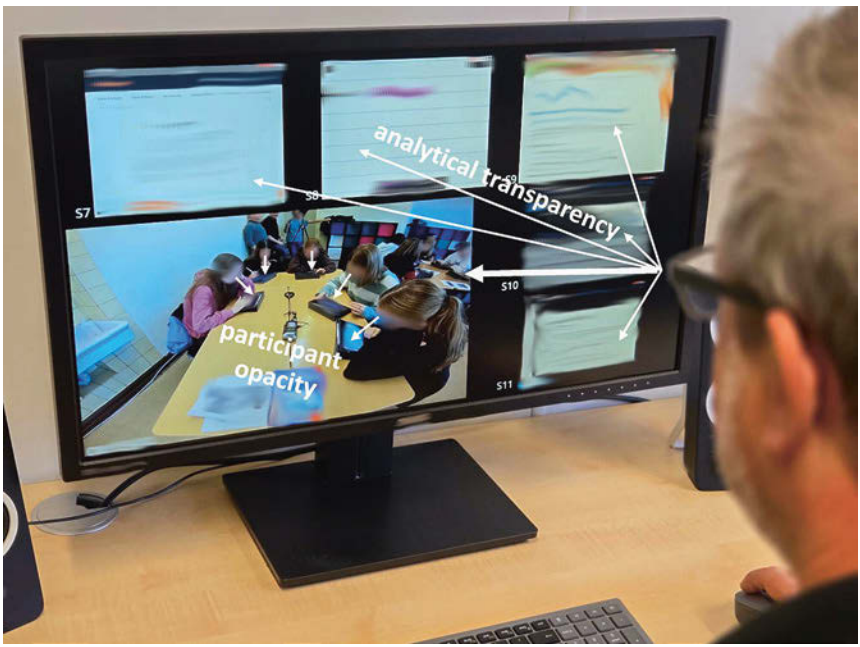


Abbildung 2: Differente Wahrnehmungsverhältnisse zwischen Beteiligten und Beobachtenden

(b) Im Unterschied zu weiten Teilen des klassenöffentlichen Unterrichts, vollziehen sich Gruppenarbeiten nicht immer als geteiltes Geschehen. Vielmehr weisen unsere Beobachtungen darauf hin, dass Momente geteilter Aufmerksamkeit im Sinne eines »fully-focused gathering« (Goffman 1963, S. 91) hier eher als Ausnahmen zu betrachten sind. Der sich in unseren Daten zeigende Normalfall im oft arbeitsteilig organisierten Gruppenprozess besteht darin, dass sich die Beteiligten zwar auf einen gemeinsamen Aktivitätsrahmen verständigt haben (z. B. bestimmte Rechercheaufträge zu bearbeiten). Das Agieren der Beteiligten selbst verläuft allerdings parallel und dezentriert – im Sinne eines »multifocused gathering« (ebd.): Manche befassen sich allein mit ihrem Tablet, manche kontaktieren dabei immer wieder ihre*n Platznachbar*in, andere arbeiten eng zusammen. Es hat sich in unserer Untersuchung als sinnvoll erwiesen, die so entstehende simultane Sequentialität von Aktivitätssträngen und die Art ihrer Verflechtung im Hinblick auf Aufmerksamkeitszentrierung (zentriert, teilzentriert, dezentriert) und enaktierte Aktivitätsrahmen (geteilt, parallel, diffus) zu differenzieren, um Arrangements zwischen Schüler*innen, Lehrkräften und digitalen Materialitäten sowie ihren segmentären Ablauf detailliert beschreiben zu können.

Befunde, die hierbei erzielt werden, können ebenfalls tabellarisch festgehalten werden. Für Phasen, die analytisch von besonderem Interesse sind, bietet es sich zudem an, das Agieren der einzelnen Beteiligten spalten- oder zeilenförmig zu notieren (vgl. etwa Dinkelaker 2010), um so (Dis-)Kontinuitäten der wechselseitigen Bezugnahme – etwa in Phasen individuellen Recherchierens – nachvollziehen zu können. Auf diese Weise ist es einerseits möglich, Aufschluss darüber zu erhalten, in welche Phasen der Ablauf des gesamten Geschehens in einer Gruppe segmentiert ist – z. B. für die Durchführung einer Rechercheaufgabe im Politikunterricht in grober Einstellung: (1) Herstellen von Produktionsbereitschaft, (2) Individuelle Suchbewegungen, (3) Inspektion durch die Lehrkraft, (4) Besprechen von Vorgehen und Zuständigkeiten, (5) Individuelle Suchbewegungen, (6) Übermittlung und Zusammenführung von Wissensfragmenten. Andererseits ist es aber auch möglich, das Agieren einzelner Schüler*innen im Umgang mit ihrem Tablet und in Bezugnahme auf ihre Mitschüler*innen ablaufbezogen zu beschreiben, woran eine sequenzanalytische Bestimmung der Bedeutung und Funktion bestimmter Aktivitäten als Interaktionspraktiken im Umgang mit bestimmten Anforderungen anschließen kann (vgl. Abschnitt 3.3). So lässt sich das Recherchieren eines Gruppenmitglieds im Politikunterricht exemplarisch etwa durch die Abfolge folgender Aktivitäten beschreiben: (1) Aufrufen einer webbasierten Suchmaschine, (2) Eintippen von Suchphrasen, (3) Auswählen einer ersten angezeigten Website, (4) Inspizieren dieser ersten Website, (5) Auswählen einer zweiten angezeigten Website, (6) Inspizieren der zweiten Website,

(7) Rückkehr zur ersten Website, (8) Partizipation an der gemeinsamen Besprechung des weiteren Vorgehens.

4.3 Praktiken im Umgang mit digitalen Materialitäten

Um die Form und Funktion von Praktiken beschreiben zu können, die Schüler*innen realisieren, um im Umgang miteinander und mit digitalen Dingen jene Anforderungen zu bearbeiten, die sich ihnen für die Erzeugung von Wissensprodukten in Gruppenarbeiten stellen, ist es notwendig, die Verkettung von Aktivitäten sequenzanalytisch zu untersuchen. Dabei orientieren wir uns an einer ethnomethodologisch-konversationsanalytisch ausgerichteten Perspektive (vgl. Housley 2012), die danach fragt, welche Probleme die Akteure durch welche Art und Weise ihres Agierens bearbeiten: »why that now?«.[5] Grundlage für Sequenzanalysen sind immer nur kleine Ausschnitte aus dem Geschehen, die für die Untersuchung der jeweiligen Fragestellung als besonders aufschlussreich betrachtet werden.

Um auf der Grundlage audiovisueller Daten entscheiden zu können, welche Phänomene berücksichtigt werden müssen, orientiert sich die Analyse an erkennbaren Relevanzsetzungen der Akteure – also daran, worauf sie sich beziehen. Das betrifft auch wahrnehmbare Phänomene auf den Tabletbildschirmen: »what we have to investigate emerges from the way in which the participants themselves display a particular field to be consequential and relevant through the orientation of their bodies and the organization of their actions« (Goodwin 2000, S. 1508).

Um die Bedeutung und Funktion multimodaler Äußerungen und Aktivitäten zu rekonstruieren, ist es notwendig, den analytischen Blick darauf zu richten, wie im Interaktionsgeschehen an sie angeschlossen wird. Die Ermittlung der Bedeutung körperlicher oder bildlich-vermittelter Äußerungen ist in manchen Fällen allerdings mit dem Problem verbunden, dass aufgrund nonverbal verfasster Anschlussreaktionen keine eindeutige Bestimmung erfolgen kann: »This is already because bodily displays are implicit and therefore often very ambiguous« (Deppermann 2013, S. 4). Im Umgang damit sind die Forschenden auf die Alltagspraxis zurückgeworfen, Bedeutungshypothesen zu formulieren, deren Gültigkeit am Fortgang des Interaktionsgeschehens geprüft wird.

Besondere Herausforderungen ergeben sich für die videobasierte Analyse von Interaktionen, wenn nicht nur eine Vielzahl an Personen eine Vielzahl an Äußerungsmodalitäten nutzt, um zu interagieren – wie es für Schulunterricht typisch ist –, sondern nun auch noch digitale Dinge in solche Prozesse eingebunden sind, die als dynamische Materialitäten prinzipiell eine unbegrenzte Anzahl an

5 Zur Differenzierung ethnomethodologischer Grundannahmen und ihrer Relevanz für die videobasierte Sequenzanalyse von Interaktionen vgl. etwa Tuma (2018).

Phänomenen in Erscheinung bringen können: »Dynamic materiality is demonstrated in the ways users can touch, tap or slide to move from one screen to another – while transferring images or text – or tap hyperlinks to access web page« (Walsh/Simpson 2013, S. 149). Solche Dinge partizipieren am Interaktionsgeschehen, indem Akteure auf sie Bezug nehmen. In ihrer Materialität sind ihnen einerseits »Affordances« (Davis 2020) im Sinne von Aufforderungsstrukturen eingeschrieben. Andererseits entstehen ihre Bedeutung und Funktion immer erst in spezifischen konturierten Situationen und Kontexten, in denen sie Verwendung finden (vgl. Proske/Rabenstein/Reis 2024). Die Frage, *wie* (digitale) Dinge am Interaktionsgeschehen partizipieren, beschäftigt Ansätze videobasierter Interaktionsforschung – in wechselnden Konjunkturen – bereits seit über 30 Jahren (vgl. Jordan/Henderson 1995, S. 75). Der Umgang mit Tablets wird in unserem Projekt untersucht, indem wechselseitige Referenzen zwischen Tablet und dem Agieren der Schüler*innen in den Blick gerückt werden. So kann nachvollzogen werden, wie Tablets als Zugang zu Wissensressourcen, als Medium des interindividuellen Austauschs, als Medium zur Darstellung und Kommunikation von Wissen oder als Ding mit eigener Agenda mit dem Interaktionsgeschehen verwoben sind und auf diese Weise am Prozess der Herstellung von Wissensprodukten beteiligt werden.

Für diese Rekonstruktion fungiert die Video- und Audiographie der Gruppen- und der Personensphäre (vgl. Abschnitt 3.2 und 3.3) wiederum als Datenbasis. Sie wird genutzt, um im Rahmen der Postproduktion einen Zusammenschnitt der verschiedenen Quellen zur simultanen Abbildung von Gruppenvideo und Screenrecordings zu ermöglichen. Zudem erweist es sich oft als hilfreich, ein Verbaltranskript des Geschehens zu produzieren, um nachvollziehen zu können, was von wem wann und wie gesprochen wird. Die Analyse selbst erfolgt allerdings primär auf Grundlage des Videomaterials, das realzeitlich, verlangsamt und wiederholt betrachtet wird, um nachvollziehen zu können, welche Aktivitäten sequenziell aufeinander folgen und welche Bedeutung und Funktion den Aktivitäten der Beteiligten als Praktiken im Umgang miteinander und mit ihrem Tablet zugewiesen wird. Je nach Bedarf ist es dabei möglich, Bildausschnitte – z. B. das Tablet einzelner Schüler*innen – im Videoeditor zu vergrößern. Als hilfreich kann es sich auch erweisen, relevante Momente von Abfolgen via Screencapture als Still festzuhalten und mit dem Verbaltranskript der betreffenden Stelle zu unterlegen (z. B. rudimentär in PowerPoint). Um Befunde einer solchen Analyse wissenschaftlich zu kommunizieren, produzieren wir Skizzen, die all jene Phänomene abbilden, die sich in der visuellen Dimension des Geschehens als relevant für das (Inter-)Agieren der Beteiligten und die darauf fußenden Interpretationen erweisen – verbunden gegebenenfalls mit entsprechenden Verbaltranskriptstellen. Screenshots der Tablets werden allerdings als anonymisierte Originale eingefügt. Exemplarisch illustriert dies Abbildung 3. Sie dokumentiert die Interaktion einer Schülerin mit digitalen Dingen während einer Gruppenarbeit im Unterricht.

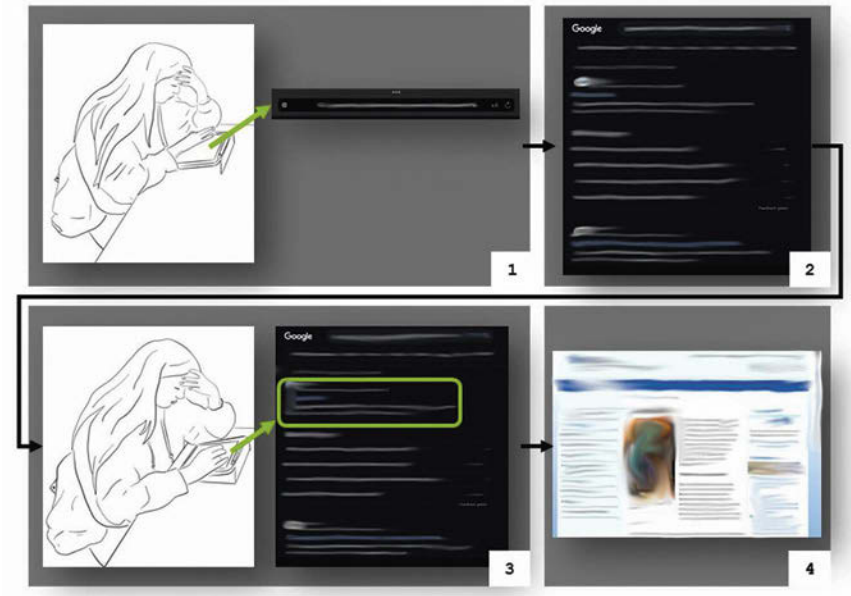


Abbildung 3: Sequenz individueller Wissensproduktion im Gruppenarbeitskontext

Als initialer Turn fungiert hier die Eingabe einer Frage in das Suchfeld der Suchmaschine Google (1). Diese ›reagiert‹, indem sie eine listenförmige Anordnung von Links als Wahrnehmungsfeld offeriert (2). In ihrem daran anschließenden Turn wird erkennbar, dass die Schülerin den ersten Eintrag auf der Liste als Antwort auf ihre Frage interpretiert, indem sie mit ihrem Finger draufklickt (3) und die digitale Materialität die Gestalt einer neuen Website annimmt (4). Diese fungiert fortan als Wahrnehmungsfeld zur Suche nach Wissensobjekten. Erkennbar wird an diesem Fall, wie Schüler*innen Praktiken des Suchens nach geeigneten Wahrnehmungsfeldern und Wissensobjekten realisieren und wie sie digitale Materialitäten als Ressourcen zur Erzeugung von Wissensprodukten nutzen. Erkennbar wird aber auch, wie das visuell-vermittelte Antwortverhalten von Suchmaschinen Affordanzen generiert und digitale Materialitäten als Mitspieler bei der Erzeugung von Wissensprodukten relevant werden.

5. Ertrag und Herausforderungen

Versucht man den forschungsmethodischen Ertrag der voranstehenden (Selbst-) Beobachtung unseres Forschungsansatzes zu pädagogischen Interaktionen im Kontext des (digitalen) Wandels zu bilanzieren, dann ist zunächst auf die Beeinflussungsrelation von Bedingungen des Untersuchungsgegenstandes und der Art der Beschaffenheit der zu erhebenden Daten hinzuweisen. Zwei Bedingungen

bzw. gegenstandsbezogene Transformationsprozesse sind für die methodische Konzipierung der Datenerhebung folgenreich: Zum einen zwingt die steigende Bedeutung dezentrierter Interaktionsformate bei der Bearbeitung schulischer Aufgaben – die erziehungswissenschaftlich unter den Stichworten der Individualisierung oder Differenzierung von Unterricht diskutiert werden – dazu, das Erhebungssetup ebenfalls zu dezentrieren. Diese Dezentrierung ist unabdingbar, wenn man in mikroethnographischer Einstellung die simultane Sequentialität der Interaktionsprozesse in den Blick bekommen möchte, auf denen die soziale Organisation von unterrichtlicher Kooperation und Wissenskommunikation fußt. Die für die videobasierte Unterrichtsforschung prägende »Zwei-Kamera-Strategie« ist hierfür nicht hinreichend.

Ebenso folgenreich für die Datenerhebung sind Veränderungen im »medialmaterielle(n) Gefüge« (Bettinger/Jörissen 2021, S. 6) von Unterricht, die mit der digitalen Transformation von und im Unterricht einhergehen (vgl. Proske/Rabenstein/Reis 2024). Insbesondere der Einbezug von mobile digital devices, wie Tablets, in die schüler*innenseitige Bearbeitung von Aufgaben stellt nicht nur neuartige Herausforderungen für die Hervorbringung von Unterrichtswissen, sondern wirft auch die Frage auf, welche Relation zwischen dem besteht, was auf den einzelnen Screens der Tablets der Schüler*innen passiert, und der Art und Weise, wie gestellte Unterrichtsaufgaben bearbeitet werden und der damit verbundene Prozess der Wissensproduktion verläuft. Um diese Frage zu beforschen, sind *Screenrecordings* – als neuartiger Typ audiovisueller Daten in der Unterrichtsforschung – essenziell. Sie ermöglichen es erstens, die *Bezugsprobleme* sichtbar zu machen, vor deren Bearbeitung Schüler*innen im wissensbezogenen Umgang mit der dynamischen Materialität des Digitalen stehen. Das Potenzial von Screenrecordings besteht hier darin, einen Zugang zur schüler*innenseitigen Arbeit an den Dingen des Unterrichts zu ermöglichen – einen Zugang, der schwieriger wäre, würden die Schüler*innen mit analogen Dingen hantieren. Zweitens erlauben es Screenrecordings, die *Selektivitäten* zu rekonstruieren, auf deren Folie das auf dem Screen *individuell* Sicht- und Hörbare Teil des *kollektiven* Bearbeitungsprozesses wird – oder auch nicht.

Beide Veränderungen im Erhebungssetup führen nicht nur zu einer Ausweitung der audiovisuellen Datengrundlage unserer Studie, sie haben auch einen analytischen Effekt in der Beforschung von Unterricht: Sie steigern die audiovisuelle Sichtbarkeit der sozialen und wissensbezogenen Organisation unterrichtlicher Interaktionsprozesse – im Anschluss an Avgustis/Oloff (2024) könnte man auch von einer ›increasing analytical transparency‹ sprechen. Diese ist aber nicht zu konfundieren mit der Teilnehmer*innenperspektive. Das Kamerawissen, auf das die Analyse Zugriff hat, entspricht nur partiell dem Teilnehmer*innenwissen.

Nicht zu leugnen ist auch, dass die Multiplizierung der audiovisuellen Datengrundlage die Aufgabe, die Selektivität der Interaktionsorganisation zu rekonstruieren, deutlich komplexer macht, weil nur in sehr kleinteiliger Analyse z. B.

erkennbar gemacht werden kann, welche individuellen Bezugnahmen auf das Tablet Teil der Gruppeninteraktion werden und welche nicht. Eine zentrale Herausforderung besteht vor allem darin, in der Rekonstruktion mit der »dynamic materiality« flüchtiger Screens und interaktiver Bezugnahme Stand zu halten.

Insofern unser Projekt, das wir hier in seinem forschungsmethodischen Vorgehen zum Beobachtungsgegenstand gemacht haben, noch am Anfang steht, können wir auf diese Herausforderungen noch keine finalen Antworten geben. Vielmehr stehen wir fortwährend vor der Aufgabe, der gesteigerten Komplexität der erhobenen audiovisuellen Daten, die selbst eine Reaktion auf die Dynamisierung des Untersuchungsfeldes ist (bezüglich multizentrischer Interaktion und dynamischer digitaler Materialitäten), durch die Entwicklung angepasster Forschungsmethoden gerecht zu werden.

Literatur

Avgustis, Iuliia/Oloff, Florence (2024): Collecting and Analysing Multi-Source Video Data. Grasping the opacity of smartphone use in face-to-face encounters. In: Haddington, Pentti/Eilittä, Tiina/Kamunen, Antti/Kohonen-Aho, Laura/Oittinen, Tuire/Rautianen, Iira/Vatanen, Anna (Hrsg.): Ethnomethodological Conversation Analysis in Motion. Emerging Methods and New Technologies. New York: Routledge, S. 85–110.

Beeli-Zimmermann, Sonja/Wannack, Evelyne/Staub, Sabina (2020): Videobasierte Unterrichtsforschung: Aufnahmen mit zwei Kameras – und dann? In: Forum: Qualitative Sozialforschung 21, H. 2, Art. 20. http://dx.doi.org/10.17169/fqs-21.2.3298.

Bettinger, Patrick/Jörissen, Benjamin (2021): Medienbildung. In: Sander, Uwe/von Gross, Friederike/Hugger, Kai-Uwe (Hrsg.): Handbuch Medienpädagogik. Wiesbaden: Springer VS, S. 1–13.

Breidenstein, Georg (2006): Teilnahme am Unterricht. Ethnographische Studien zum Schülerjob. Wiesbaden: VS Verlag für Sozialwissenschaften.

Breidenstein, Georg/Hirschauer, Stefan/Kalthoff, Herbert/Nieswand, Boris (2020): Ethnografie. Die Praxis der Feldforschung. 3. Auflage. München: UVK.

Caruso, Marcelo (2019): Geschichte der Bildung und Erziehung: Medienentwicklung und Medienwandel. Paderborn: Ferdinand Schöningh.

Cazden, Courtney B. (1986): Classroom Discourse. In: Wittrock, Merlin C. (Hrsg.): Handbook of Research on Teaching. New York: Macmillan, S. 432–463.

Corsten, Michael/Pierburg, Melanie/Wolff, Dennis/Hauenschild, Katrin/Schmidt-Thieme, Barbara/Schütte, Ulrike/Zourelidis, Sabrina (Hrsg.) (2020): Qualitative Videoanalyse in Schule und Unterricht. Weinheim und Basel: Beltz Juventa.

Davis, Jenny L. (2020): How Artifacts Afford. The Power and Politics of Everyday Things. Cambridge: MIT Press.

Deppermann, Arnulf (2013): Multimodal interaction from a conversation analytic perspective. In: Journal of Pragmatics 46, H. 1, S. 1–7.

Dinkelaker, Jörg (2010): Simultane Sequenzialität. Zur Verschränkung von Aktivitätssträngen in Lehr-Lernveranstaltungen und zu ihrer Analyse. In: Corsten, Michael/Krug, Melanie/Moritz, Christine (Hrsg.): Videographie praktizieren. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 91–117.

Erickson, Frederick (1982): Audiovisual Records as a Primary Data Source. In: Sociological Methods & Research 11, H. 2, S. 213–232.

Erickson, Frederick (1992): Ethnographic microanalysis of interaction. In: LeCompte, Margaret D./Millroy, Wendy L./Preissle, Judith (Hrsg.): The handbook of qualitative research in education. San Diego: Academic Press Inc., S. 201–225.

Erickson, Frederick (2006): Definition and Analysis of Data from Videotape: Some Research Procedures and Their Rationales. In: Green, Judith L./Camilli, Gregory/Elmore, Patricia B. (Hrsg.): Handbook of complementary methods in education research. London: Routledge, S. 177–191.
Glaser, Barney G./Strauss, Anselm L. (2005): Grounded Theory. Strategien qualitativer Forschung. 2. Auflage. Bern: Huber.
Goffman, Erving (1963): Behavior in public places: Notes on the social organization of gatherings. New York: Free Press.
Goodwin, Charles (2000): Action and embodiment within situated human interaction. In: Journal of Pragmatics 32, S. 1489–1522.
Herrle, Matthias (2020): Ethnographic Microanalysis. In: Huber, Matthias/Froehlich, Dominik E. (Hrsg.): Analyzing Group Interactions. A Guidebook for Qualitative, Quantitative and Mixed Methods. Abingdon: Routledge, S. 11–25.
Herrle, Matthias/Breitenbach, Sebastian (2016): Planung, Durchführung und Nachbereitung videogestützter Beobachtungen im Unterricht. In: Rauin, Udo/Herrle, Matthias/Engartner, Tim (Hrsg.): Videoanalysen in der Unterrichtsforschung: Methodische Vorgehensweisen und Anwendungsbeispiele. Weinheim und Basel: Beltz Juventa, S. 30–49.
Herrle, Matthias/Dinkelaker, Jörg (2016): Qualitative Analyseverfahren in der videobasierten Unterrichtsforschung. In: Rauin, Udo/Herrle, Matthias/Engartner, Tim (Hrsg.): Videoanalysen in der Unterrichtsforschung. Methodische Vorgehensweisen und Anwendungsbeispiele. Weinheim und Basel: Beltz Juventa, S. 76–129.
Herrle, Matthias/Hoffmann, Markus/Proske, Matthias (2022): Unterrichtsgestaltung im Kontext digitalen Wandels: Untersuchungen zur soziomedialen Organisation Tablet-gestützter Gruppenarbeit. In: Zeitschrift für Erziehungswissenschaft 25, S. 1389–1408. https://doi.org/10.1007/s11618-022-01099-8.
Housley, William (2012): Ethnomethodology, conversation analysis and educational settings. In: Delamont, Sara/Jones, Angela (Hrsg.): Handbook of Qualitative Research in Education. Cheltenham: Edward Elgar, S. 448–459.
Jordan, Brigitte/Henderson, Austin (1995): Interaction Analysis: Foundations and Practice. In: The Journal of the Learning Sciences 4, H. 1, S. 39–103.
Kendon, Adam (1990): Some context for Context Analysis: a view of the origins of structural studies of face-to-face interaction. In: Kendon, Adam (Hrsg.): Conducting interaction: Patterns of behavior in focused encounters. Cambridge: Cambridge University Press, S. 15–49.
Knoblauch, Hubert (2001): Fokussierte Ethnographie. In: Sozialer Sinn 2, H. 1, S. 123–141.
Krämer, Sybille (2008): Medium, Bote, Übertragung. Kleine Metaphysik der Medialität. Frankfurt am Main: Suhrkamp.
LeBaron, Curtis (2008): Microethnography. In: Donsbach, Wolfgang (Hrsg.): The International Encyclopedia of Communication. Volume VII. Malden: Blackwell, S. 3120–3123.
Maynard, Douglas W. (2013): Everyone and No One to Turn to: Intellectual Roots and Context for Conversation Analysis. In: Sidnell, Jack/Stivers, Tanya (Hrsg.): The Handbook of Conversation Analysis. Chichester, West Sussex: Wiley-Blackwell, S. 11–31.
Mroz, Aurore P. (2014): Process Research Screen Capture. In: Chapelle, Carol A. (Hrsg.): The Encyclopedia of Applied Linguistics. John Wiley & Sons, S. 1–7. https://doi.org/10.1002/9781405198431.wbeal1450.
Nevile, Maurice/Haddington, Pentti/Heinemann, Trine/Rauniomaa, Mirka (2014): On the interactional ecology of objects. In: Nevile, Maurice/Haddington, Pentti/Heinemann, Trine/Rauniomaa, Mirka (Hrsg.): Interacting with objects. Language, materiality, and social activity. Amsterdam: John Benjamins, S. 3–26.
Proske, Matthias/Rabenstein, Kerstin/Meseth, Wolfgang (2021): Unterricht als Interaktionsgeschehen. Konstitution, Ordnungsbildung und Wandel. In: Hascher, Tina/Helsper, Werner/Idel, Til-Sebastian (Hrsg.): Handbuch Schulforschung. Wiesbaden: Springer VS, S. 1–27. https://doi.org/10.1007/978-3-658-24734-8_42-1.
Proske, Matthias/Rabenstein, Kerstin/Reis, Oliver (2024): Digitale Technologien als eigensinnige Akteure der Transformation von und im Unterricht. In: Zeitschrift für interpretative Schul- und Unterrichtsforschung 13, S. 3–17.
Psathas, George (1968): Ethnomethods and Phenomenology. In: Social Research 35, H. 3, S. 500–520.

Rabenstein, Kerstin/Kunze, Kerstin/Martens, Matthias/Idel, Til-Sebastian/Proske, Matthias/Strauß, Svenja (Hrsg.) (2018): Individualisierung im Unterricht. Transformationen – Wirkungen – Reflexionen. Bad Heilbrunn: Klinkhardt.

Reh, Sabine/Idel, Till-Sebastian/Rabenstein, Kerstin/Fritzsche, Bettina (2015): Ganztagsschulforschung als Transformationsforschung. Theoretische und empirische Erträge des Projekts. In: Reh, Sabine/Fritzsche, Bettina/Idel, Till-Sebastian/Rabenstein, Kerstin (Hrsg.): Lernkulturen. Rekonstruktion pädagogischer Praktiken an Ganztagsschulen. Wiesbaden: Springer VS, S. 297–336.

Sahlström, Fritjof/Tanner, Marie/Valasmo, Verneri (2019): Connected youth, connected classrooms. Smartphone use and student and teacher participation during plenary teaching. In: Learning, Culture and Social Interaction 21, S. 311–331.

Tuma, René (2018): Video-Interaktionsanalyse. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 423–444.

Walsh, Maureen/Simpson, Alyson (2013): Touching, tapping … thinking? Examining the dynamic materiality of touch pad devices for literacy learning. In: Australian Journal of Language and Literacy 36, H. 3, S. 148–157.

Eye-Viewing als Methode der qualitativen Sozialforschung

Durch Eye-Tracking-Technologie erweiterte Videographie zur Erforschung von (medienbezogenen) Praktiken in Schule und Unterricht

Isabel Neto Carvalho

Zusammenfassung: Der Beitrag erörtert die Notwendigkeit, methodologische Ansätze in der Unterrichtsforschung zu erweitern, um schulische Praktiken in einer Kultur der Digitalität besser analysierbar zu machen. Trotz zahlreicher Studien zur Mediennutzung, fehlt es an einem Verständnis dafür, wie Medien die Praktiken im Klassenzimmer verändern. Ethnographische Ansätze, insbesondere videobasierte Unterrichtsbeobachtungen, bieten Möglichkeiten, diese Veränderungen in Lehr- und Lernprozessen zu analysieren. Dennoch gibt es eine Diskrepanz zwischen dem, was beobachtbar ist, und dem, was während der Interaktion mit digitalen Artefakten geschieht. Im Beitrag wird gezeigt, wie der Einsatz von Eye-Tracking-Technologie dabei helfen kann, Mensch-Mensch-Maschine-Praktiken zu fokussieren, indem die beobachteten Akteure zu Ko-Konstrukteuren der Forschung gemacht werden. Bei dieser als »Eye-Viewing« bezeichneten Methode werden videoethnographische Verfahren mit Eye-Tracking-Brillen kombiniert, um ein tieferes Verständnis der Perspektive der Akteure zu gewinnen. Der Beitrag zeigt, wie diese Methode (nicht nur) in der medien- und schulpädagogischen Forschung zur Anwendung kommen kann und diskutiert die potenziellen Chancen und Herausforderungen, die sie mit sich bringt.

Schlüsselwörter: Eye-Viewing, digitale Medien, Ethnographie, medienbezogene Praktiken, Schule

1. Einleitung

Um soziale Praxis rund um Digitalität und Bildung und die mit ihr verbundenen gesellschaftlichen Herausforderungen deutlicher nachzeichnen zu können, braucht es in der sozialwissenschaftlichen Forschung einen Perspektivwechsel weg von gängigen Blicken (z. B. auf lediglich die »Nutzung« digitaler Medien) hin zu medienbezogenen Praktiken von Lehrpersonen, der einen »wissensabhängigen,

kulturell spezifischen Umgang mit […] Medien« (Reckwitz 2008, S. 165) ins Zentrum stellt. So fehlt es – trotz zahlreicher Studien zu Materialität in der Schule – nach wie vor an empirischer Klarheit darüber, wie pädagogische Praktiken im Kontext von Technologie- und Datenintegration innerhalb und außerhalb des Klassenzimmers genau entstehen. Einer der Gründe dafür ist, dass viele Studien sich auf den instrumentellen Mediengebrauch im Klassenzimmer fokussieren, ohne den wachsenden Einfluss neuer epistemischer Akteure in Form von Medien angemessen zu berücksichtigen. Die Rolle der digitalen Medien im Klassenzimmer hat sich jedoch über die eines bloßen »Hilfsmittels« hinaus entwickelt (Cramer 2015; Macgilchrist 2019). Vielmehr sind sie nun in Praktiken involviert, manchmal sogar ohne direkt »anwesend« zu sein (Fawns/O'Shea 2018) und werden zunehmend zu Akteuren in diesen »Lehrer-Lerner-Ding-Interaktionen« (vgl. Allert/Asmussen 2017).

Angesichts dieser Überlegungen liegt es auf der Hand, dass Methoden der Sozialforschung entwickelt werden müssen, die die neuen sozialen Realitäten mit ihrer tiefen Verankerung digitaler Infrastrukturen in Bildungssettings reflektieren (z. B. Macgilchrist 2021). Eine Möglichkeit, die Sprache-Körper-Raum-Ding-Verwobenheit von Praktiken phänomensensibel zu erforschen und dabei auch Perspektiven der Digitalität einzubeziehen, ist ein videoethnographischer Ansatz. Videoethnographische Methoden (vgl. Dindar et al. 2017; Neto Carvalho/Schiefner-Rohs/Troxler 2022) bieten den Vorteil, dass Setzungen im und durch das Feld geschehen (Dinkelaker 2018) und es ermöglichen, das »Noch-nicht-in-Worte-Gefasste« zu beschreiben (vgl. Mohn 2008). So können Normativitäten, die in gängige Begrifflichkeiten und routinisierte Praxen eingebettet sind, offengelegt und damit der Reflexion und dem Verstehen zugänglich gemacht werden.

Um diese Praxen auf einem feingranularen Niveau und unter starker Berücksichtigung der Perspektiven der Akteure sichtbar zu machen, habe ich in den letzten Jahren den Nutzen von Eye-Tracking-Technologie als Ergänzung ethnografisch-orientierter Videographie exploriert. Im folgenden Abschnitt wird diese Methode des sogenannten »Eye-Viewing«[1] vorgestellt. An einem Fallbeispiel wird der Beitrag der Frage nachgehen, wie durch Eye-Viewing die bisher unsichtbaren oder schwer fassbaren Mensch-Maschine-Interkationen im Unterricht empirisch detaillierter und phänomensensibler sichtbar gemacht werden können.

1 Mit dieser Bezeichnung grenze ich die Methode zu experimentell-quantifizierenden Eye-Tracking-Verfahren z. B. der Neuropsychologie und Werbewirksamkeitsforschung ab.

2. Videographie medienbezogener Praxis

Obwohl die Videographie medienpädagogischer Praxis ein relativ junger Forschungszweig ist, bezieht sie sich auf Verfahren (z. B. ethnographischer Film oder sozialwissenschaftliche Foto- und Filmanalyse), die bereits im späten 19. Jahrhundert in der Forschung eingesetzt wurden (vgl. Tuma/Schnettler/Knoblauch 2013; vgl. Knoblauch/Wilke i. d. B.). Heute lassen sich sehr unterschiedliche Arten der Videographie pädagogischer Praktiken unterscheiden (siehe z. B. Dinkelaker/Herrle 2009). So nutzen einige Forscher:innen den Einsatz von Videographie im Klassenzimmer, um Einblicke in die sozio-mediale Organisation von Unterrichtsinteraktionen und den Einsatz digitaler Materialitäten zu gewinnen (Herrle/Hoffmann/Proske 2022; vgl. Herrle/Proske/Puzicha/Zimmer i. d. B.). Außerdem nutzen sie diese Methode, um die Produktion von Leistung in digitalisierten Unterrichtssettings zu beschreiben (z. B. Spiegler et al. 2023). Daneben gibt es Ansätze, die stärker in der ethnografischen Forschungstradition stehen. Videoethnographische Ansätze der Unterrichtsbeobachtung ermöglichen einen mikrologischen, weitgehend offenen Blick auf Lehr- und Lernprozesse aller Art, auch auf solche mit digitalen Medien (vgl. Coleman 2010). Dennoch bleibt die Diskrepanz zwischen dem, was von außen wahrgenommen werden kann und der tatsächlichen Dynamik digitaler Artefakte im Interaktionskontext undurchsichtig, selbst im Fall der Videoethnographie, bei der die Forscher:innen das Subjekt typischerweise von außen betrachten, gefiltert durch die Kamera. Die Interaktion zwischen digitalem Gerät und User bringt allerdings auch stille Beziehungen hervor, die im Gebrauch intransparent sind (Bettinger 2023; Krämer 2008) und oft nur in Momenten der Störung oder des Zusammenbruchs empirisch sichtbar werden (ebd.; aber auch Alirezabeigi/Masschelein/Decuypere 2020).

Eine Möglichkeit, den Fokus auf Epiphänomene zu reduzieren, besteht darin, Medien aus einer praxeologischen Perspektive mit Bezug auf die Soziomaterialität zu verstehen. Das bedeutet, Technologien »nicht als willfähige und neutrale Werkzeuge zu begreifen, die in den Händen der Lehrer und Schüler ohne weiteres ihren Dienst verrichten« (Röhl 2013, S. 2), sondern Technologien auch empirisch mit ihrem konstitutiven Beitrag zur Praxis ernst zu nehmen (vgl. ebd.). Digitale Medien werden somit nicht auf bloße Hilfsmittel im Unterricht reduziert, sondern als Akteure gesehen, die Verhaltensweisen ermöglichen und begrenzen (vgl. Kalthoff/Cress/Röhl 2016).

Dieser praxeologisch-theoretische Ansatz lässt sich durch einen affordanztheoretischen ergänzen (z. B. Bareither 2023). Das Konzept der Affordanz wird seit einigen Jahren in der medienbezogenen Forschung (insbesondere im Rahmen der »digitalen Ethnographie«, vgl. Hopkins 2016) verwendet, um die in Technologien und Medien eingeschriebenen Handlungsaufforderungen und -beschränkungen zu beschreiben (ebd., S. 2). In Anlehnung an McLuhan (1992)

wird davon ausgegangen, dass jedes Medium die Macht hat, den Nutzer:innen seine eigenen Postulate aufzuerlegen (Zillien 2008). Einerseits ist die Affordanz eines Mediums unabhängig von den Bedürfnissen, der Aufmerksamkeit oder dem Urteilsvermögen der Beobachtenden zu verstehen: »For example, a glass of water affords drinking whether or not I am thirsty, a ball affords throwing, whether or not anybody sees it, and a pit affords falling even if it is concealed by brush« (Gaver 1991, S. 80, zit. n. Zillien 2008, S. 166). Gleichzeitig erfordert das Erkennen von Affordanzen Erfahrung: »Deshalb nimmt ein Mensch – der in einer Gesellschaft, die Stühle als Sitzgelegenheiten definiert, aufgewachsen ist – direkt wahr, dass ein Sessel ebenfalls die Affordanz des Darauf-Sitzens hat« (Zillien 2008, S. 4). Das Affordanzkonzept setzt also eine wechselseitige Bezugnahme von Subjekt und Objekt voraus und überwindet damit den Gegensatz zwischen beidem: »By cutting across the subjective/objective barrier, Gibson's affordances introduce the idea of the actor-environment mutuality; the actor and the environment make an inseparable pair« (McGrenere/Ho 2000).

Nimmt man diese theoretischen Thesen ernst, rücken empirische Fragen nach dem »Wie« der Praxis in den Vordergrund: Wie genau läuft die Interaktion zwischen Lehrpersonen und Medien ab? Welche pädagogischen Praktiken entstehen in dieser Mensch-Maschine-Interaktion? Mit meinem empirischen Ansatz will ich die sogenannten »stillen Daten« (vgl. Hirschauer 2001) in der Interaktion zwischen menschlichen und nicht-menschlichen Akteuren sichtbar machen, um Einblicke in das Stille/Passive/Unausgesprochene des kulturellen Alltags und vor allem in die Konstitution von Unterricht aus der Perspektive der Digitalität zu geben.

3. Erhebungskontext: Eye-Viewing als ethnographisch-inspirierte Praktiken- und Affordanzanalyse

Um dies auf einer feingranularen Ebene zu erreichen, setze ich Eye-Tracking-Technologie (vgl. Reitstätter/Pesen/vom Lehn i. d. B.) ein, um medial konstituierte Praktiken zu visualisieren. Bislang wurde die Eye-Tracking-Technologie vor allem in der quantitativen Bildungsforschung eingesetzt, insbesondere in Bezug auf Fragen der Professionalisierung (z. B. Keller et al. 2022) oder in der Lehr-Lernforschung (z. B. Wiedbusch et al. 2023; Mayer et al. 2023). Studien, die mithilfe der Eye-Tracking-Technologie *mobile* Daten im schulischen Umfeld erheben und damit potenziell auch die soziale Situation berücksichtigen, wurden bisher nur vereinzelt durchgeführt: Cortina et al. (2018) und Kleiner, Rakoczi und Krimm (2012) nutzen Eye-Tracking-Technologie, um das Unterrichtsgeschehen aus der Perspektive der Lehrkraft zu analysieren. Sie betonen insbesondere die (veränderten) Möglichkeiten für (zukünftige) Lehrkräfte, ihren eigenen Unterricht zu reflektieren. In jüngerer Zeit gibt es auch fachdidaktisch orientierte

Studien, die mit mobiler Eye-Tracking-Technologie arbeiten (z. B. Sprenger 2021 für die Mathematikdidaktik, Esau/Fletcher 2018 für die Technikdidaktik, Bock/Dresing 2021 für die Deutschdidaktik) und i. d. R. Lese- und Aufzeichnungsstrategien analysieren. Diese Videodaten werden dann wieder quantitativ ausgewertet (z. B. in Bezug auf die Häufigkeit des Blickkontakts).

Allerdings vernachlässigt auch diese Perspektive die soziale Situation, wenn wir Praktiken im Feld sicht- und verstehbar machen wollen. Mithilfe von Eye-Tracking-Brillen können jedoch die Augenbewegungen und visuelle Blickfelder der Brille-tragenden Person zeitsynchron aufgezeichnet und in Bezug zu anderen Datenquellen – z. B. Hand-, Stand- und 360°-Kameras – gesetzt werden. Durch die Verbindung dieser unterschiedlichen Aufnahmenmodi entsteht dann ein Video, das sowohl die Perspektive als auch die Fokussierung der Akteure wiedergibt (vgl. Cortina et al. 2018). Die qualitative Analyse von *zwischenmenschlichen* Interaktionen durch Eye-Tracking-Videos wird in der Linguistik bereits praktiziert, insbesondere im Rahmen der Konversations- und Multimodalforschung (z. B. Holler/Kendrick 2015; Stukenbrock 2018). Kesselheim und Hottinger (2019) analysieren beispielsweise die Praktiken der Wissenskonstruktion durch »Mit-Mach«-Museumsbesucher:innen und beziehen neben Audio- und Videodaten auch durch Eye-Tracking gewonnene Videos in ihre Analysen ein, indem sie untersuchen, wie Objekte oder räumliche Elemente von Akteuren für ihre Museumsinteraktion in Anspruch genommen werden.

Aufbauend auf diesen etablierten qualitativen Methoden der Linguistik, hat die Eye-Tracking-Technologie auch innovatives Potenzial für ethnographische Erforschung von Interaktionen zwischen *menschlichen* und *nicht-menschlichen* Akteuren. Daher ergänze ich die videographische Forschung (die bisher mit Hand-, Stand- und 360°-Kameras durchgeführt wurde) mit Eye-Tracking-Technologie, um ›ein Sehen durch die Augen Anderer‹ zu ermöglichen. An dieser Stelle möchte ich betonen, dass es bei dem Vorgehen um das Sichtbar-Machen des Sehens und weniger um das Tracking geht: »Während es bei der Anwendung der Eye-Tracking-Technologie meist um das ›Verfolgen‹ von Blicken, um Blickpfade geht […] [,] geht es uns aber um das Bild der Brille und das bildliche Sehen auf Praktiken durch die Augen von Akteur*innen [sic!] in diesen Praktiken« (Neto Carvalho et al. 2022, S. 267).

Es wird daher bewusst darauf verzichtet, ausschließlich Eye-Tracking-Daten zu erheben, da dies, wie Kesselheim und Hottinger (2019) betonen, höchst problematisch wäre: »Anders als die Leseforschung, die die Blickrichtung plausibel als indirektes Maß für Interpretationsprozesse auffasst (Clifton Jr., Charles et al. 2016), beantwortet Eye-Tracking in Interaktionsstudien nur die Frage wohin geblickt wurde, aber weder was wirklich gesehen noch als was das Gesehene gesehen wurde« (ebd.). Erst die zusätzliche Dokumentation der Sprache und des Körpers der Akteure gibt Aufschluss darüber, »was die Blicke für die Interaktionsteilnehmer jeweils bedeuten; und wir zeigen, wie die Hinzunahme von

Audio und Video untersuchbar macht, wie ›sozial dargestellter Blick‹ und gemessener Blick auseinanderklaffen. Blick ist, wie wir zeigen werden, keineswegs nur eine Frage der Augenrichtung (wie in Eye-Tracking-Studien oft vorausgesetzt wird), sondern eine multimodale Gestalt, für die die Ausrichtung von Becken, Schultern und Kopf eine ebenso große Bedeutung haben kann wie die Blickrichtung« (ebd.).

Auf diese zweite Perspektive der Beobachtungen, kann also nicht verzichtet werden. Geertz (1983) stellt fest: »Schreiben ist die zentrale ethnographische Tätigkeit«. Beobachtungen werden während des Feldaufenthalts protokolliert, O-Töne von Feldakteuren transkribiert und am Schreibtisch werden diese Notizen zu ›dichten Beschreibungen‹ vertextet. Auch der hier vorgestellte Zugang zum Feld via Kamera und Eye-Tracking versteht sich als schreibende Tätigkeit, die ebenfalls einem *Dreischritt* folgt: Im Rahmen der Datenerhebung orientiere ich mich am kamera-ethnographischen Konzept von Mohn und Amann (2005, 2006a, 2006b; Mohn 2008), welche sich als Ethnograph:innen einen Zugang zum Feld durch die blickende Kamera als Federhalter verschaffen, d. h. sie benutzten quasi ihre Kamera wie einen Stift. In der Erhebungssituation werfe ich *erstens* also sogenannte »Blickschneisen«. Ich werfe sie in die Richtung, die sich aus der jeweiligen Forschungsfrage, aber auch aus dem, was das Feld als relevant markiert (z. B. Markierungen durch Worte, Gesten, die räumliche Gestaltung) ergibt. Die Handkamera (diese hat sich als Erhebungstechnologie in den doch sehr unvorhersehbaren schulischen Settings als besonders tragfähig erwiesen) folgt dem Blick und beschreibt bereits soziale Wirklichkeit.

4. Analysekontext: Experimentelle Schnitttechniken, szenische Beschreibungen und sequenzielle Analyse

Diese vom Videomaterial eingefangenen ›Geschichten‹ werden *zweitens* im Rahmen eines experimentellen Schnitts ausfindig gemacht. Die grundlegende Herausforderung jeder videographischen Forschung ist nämlich die angemessene Würdigung der so gewonnenen Fülle an Datenmaterial. Laut Luhmann (1992) haben Unterrichtsinteraktionen einen episodalen Charakter. Das Feld generiert also Prozesse, die eine narrative Struktur aufweisen. In einem vorläufigen Schritt der Vertextung dieser ›Geschichten‹, den ich bereits als Datenanalyse verstehe, werden Gliederungen bzw. nach dem Vorbild der dokumentarischen Methode sogenannte szenische Verläufe (vgl. Bohnsack 2009) zu den jeweiligen gefilmten Episoden angefertigt. Aus dem Videomaterial werden dann besonders dichte, komplexe Szenen herausgeschnitten und kurze Filmclips erstellt, die eine Geschichte erzählen. Die beiden von der Hand- bzw. Standkamera und der Eye-Tracking-Technologie produzierten Videos werden dafür in einem Zweibildverfahren nebeneinandergestellt, um sie in ihrer Synchronität analysierbar zu machen.

Nachdem die visuelle Datenkonstruktion erfolgt ist, schließt *drittens* unmittelbar die textuelle Datenkonstruktion an, indem von diesen ›Geschichten‹ sogenannte szenische Beschreibungen angefertigt werden. Unter einer szenischen Beschreibung wird eine detaillierte Vertextung einer zuvor begründet ausgewählten Sequenz verstanden (vgl. Neto Carvalho 2017). Die Beschreibung gestaltet sich dabei möglichst detailliert und gleichzeitig möglichst frei von den Vorannahmen der schreibenden Ethnograph:innen. Bohnsack nennt diese Art der Beschreibung in Anlehnung an Panofsky ›vorikonographisch‹: »Es ist diese vor-ikonografische oder – in der Sprache der Semiotik – die denotative Ebene, deren genaue Beobachtung und Beschreibung die wesentliche Grundlage der ikonologischen Interpretation und der dokumentarischen Bild- und Videointerpretation darstellt« (Bohnsack 2021, S. 575). Im Anschluss an Breidenstein liegt »das Qualitätskriterium im Rahmen der Strategie der Verschriftlichung einerseits in der Ausführlichkeit und Detaillierung des Schreibens und andererseits in der Reflexion und zunehmenden Fokussierung und analytischen »Verdichtung« des Schreibens« (Breidenstein 2010, S. 210). Der Vorteil der Verschriftlichung der Beobachtung liegt in der sprachlichen Erkundung eines Phänomens, das noch nicht in sprachlicher Form existiert und erst durch die Beschreibung in die Sprache gebracht wird. Praktiken und implizites, körperliches Wissen müssen erst aufgeschrieben werden, weil sie nicht in schriftlicher Form vorliegen und implizites Wissen sprachlos und stumm ist. Die als Fließtext verfasste Beschreibung umfasst »Blicke (Blickrichtungen, Blickschneisen, Blickkontakt), Körper (Bewegungen von Körpern im Raum, mit Dingen, Körperhaltungen im Verhältnis zueinander, Mimik und Gestik), Sprechakte […], Objekte, Dinge, Artefakte, die in Praktiken involviert sind, und Kontexte, in die die Praktiken eingebettet sind« (Neto Carvalho 2017, S. 119) aus der Perspektive beider Kameras. Es können auch Standbilder integriert werden, um eine hohe Dichte der Beschreibung zu ermöglichen. Sprechakte werden markiert und mit Anführungszeichen dargestellt.

Im Prozess der Bearbeitung der Videos werden also schon relevante Themen entdeckt und erste Thesen entwickelt, die bei der Verarbeitung der Daten eine Rolle spielen (vgl. Neto Carvalho 2017). Die entstandenen Szenen werden dann in Anlehnung an die Grounded Theory (vgl. Strübing 2014) offen kodiert. Ähnliche Codes interpretiere ich als *seriell* auftretende soziale Praktiken. Die Benennung der Praktiken kann als Kodiersystem verstanden werden, das z. B. in MAXQDA erfasst werden kann. Die szenischen Beschreibungen können dann vor dem Hintergrund der *sequenziellen* Struktur sozialer Praktiken mit einer sequenziellen Interpretationsmethode rekonstruiert werden (z. B. Wernet 2009).

Durch die Kombination der Aufnahmemodi der Handkamera und der (mobilen) Eye-Tracking-Technologie entstehen audiovisuelle Daten, die viel stärker

auf die Perspektive der jeweiligen Akteure fokussiert und damit empirisch zugänglich sind. Durch das Tragen einer Eye-Tracking-Brille werden die beobachteten Akteure zu Ko-Konstrukteuren von sozialwissenschaftlich relevanten Beobachtungsdaten. Insbesondere mediengestützte Praktiken, die einer Beobachtung von außen verborgen bleiben würden (z. B. das Verweilen auf bestimmten, möglicherweise irritierenden Texten oder Objekten auf Bildschirmen; das Ausprobieren neuer Programme an Laptop oder PC; die Eingabe von Noten durch eine Lehrerin in eine App auf ihrem Smartphone etc.), können so für die Forscher:innen sichtbar gemacht werden. Im Folgenden werden anhand eines ausgewählten Beispiels die möglichen Chancen und Herausforderungen dieses Ansatzes aufgezeigt.

5. Nutzungsszenarien: Praktiken der Unterrichtsvorbereitung sichtbar machen – Lehrmaterialerstellung als krisenhafte Aushandlung zwischen menschlichen und nicht-menschlichen Akteuren

Der nachfolgend dargestellte Fall wurde im Rahmen eines von der RPTU Kaiserslautern-Landau in den Jahren 2019 bis 2023 geförderten Projekts mit dem Titel »Kontextsensitive Videographie zur Erfassung digitaler Transformationsprozesse in Schule und Unterricht (EduGraphie)« erhoben. In dieser Zeit beobachten wir Lehrpersonen an vier weiterführenden Schulen unterschiedlicher Schulformen in Deutschland, zunächst pandemiebedingt an ihren Heimarbeitsplätzen, später auch im Unterricht vor Ort. Die Akteure werden zunächst im jeweiligen Feld teilnehmend beobachtet, dann mit einer Hand- oder Standbildkamera, und später tragen sie in diesem Forschungssetting eine mobile Eye-Tracking-Brille. An einem Fall wird im Folgenden – zugegeben verkürzt – herausgearbeitet, wie digitale Medien an Praktiken konstitutiv beteiligt sind und welchen Erkenntnisgewinn die Perspektive des Eye-Viewing bietet.

Wir beobachten einen Lehrer, der an einem Schreibtisch sitzt und Lehr- und Lernmaterialien für den Biologieunterricht vorbereitet (siehe Abb. 1). Schräg gegenüber von ihm ist eine Standkamera positioniert, hinter der die Kamerafrau steht. Der Lehrer selbst trägt eine Eye-Tracking-Brille. Das entstandene Video zeigt einen Büroraum mit vier rechteckigen Tischen in der Mitte des Raumes. Der Laptop ist aufgeklappt und an einen externen Monitor angeschlossen. Auf der Tastatur des Laptops ist ein Ordner zu sehen, auf dem ein aufgeschlagenes Buch liegt. Vor dem Monitor befindet sich eine kabellose Tastatur, daneben liegt eine kabellose Computermaus. Ein Tablet befindet sich zwischen der Tastatur und der Lehrkraft. Links neben der Tastatur befindet sich ein Smartphone (im Ruhezustand).

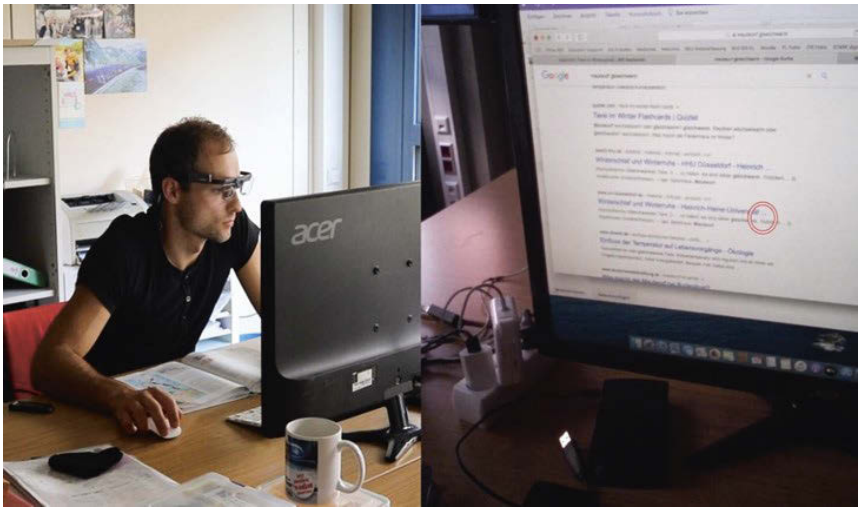


Abbildung 1: Praktiken der Krisenbearbeitung, sichtbar durch die Handkamera und die Aufnahmen aus der Perspektive der Eye-Tracking-Kamera | Kameraführung und Standbilderstellung: Isabel Neto Carvalho

Die folgende, für den Beitrag gekürzte Beschreibung, die Standkamera- und Eye-Tracking-Aufnahmen integriert, fasst nun das aus der Situation hervorgehende Geschehen zusammen:

> *Zunächst öffnet der Lehrer über seinen Laptop das OneNote-Programm, welches mittels PC-Monitor sichtbar ist. Sein Blick, erkennbar an einem roten Kreis (dem Fixationspunkt), wandert über die Dokumentnamen auf der linken Seite. »Jo, wad mach ich? (.) ähm. Bio«, sagt der Lehrer und klickt sich kurz durch das Programm. Es sind einzelne Textpassagen, Bilder und Tabellen sichtbar. Nach kurzem Auf- und Abscrollen klickt der Lehrer im Menüband auf eine neue Textseite, woraufhin eine leere Seite erscheint. Er dreht sich zum halbhohen Regal hinter sich um und schnappt sich eines der Bücher, das er kurze Zeit später zwischen Tastatur und Körper legt (er legt das Tablet auf den Ordner). Es handelt sich um ein Biologiebuch. Danach wandert sein Blick zwischen dem Buch und dem Dokument (das Textpassagen, Bilder und Tabellen beinhaltet) hin und her. Der Lehrer öffnet einen Internet-Browser. Die Google-Suchmaschine ist geöffnet. Er klickt in das Suchfeld, beugt sich über die Tastatur, sieht auf die Tasten, tippt ›maulwurf gleichwarm‹ ein und drückt die Enter-Taste. Im oberen Ergebnisfeld werden ein Teaser und ein Bild von einem Maulwurf präsentiert. Die Lehrkraft überfliegt den kurzen Text und klickt einen darunter deutlich blau eingefärbten und unterstrichenen Link zum Wikipedia-Artikel über Maulwürfe. Der Artikel wird sichtbar. Ohne diesen zu lesen, öffnet der Lehrer per Tastatureingabe ein Suchfeld und gibt ›therm‹ in das Suchfeld ein. Daraufhin wird ›0/0‹ angezeigt. »Alter«, flüstert der Lehrer und tippt auf den Zurück-Button des Webbrowsers. Zurück auf der Google-Ergebnisseite, überfliegt er die untereinanderstehenden*

Titel und klickt auf das vierte Ergebnis ›Winterschlaf und Winterruhe – Heinrich-Heine-Universität‹. Es öffnet sich ein PDF-Dokument. Wieder öffnet die Lehrkraft ein Suchfeld und gibt ›maulw‹ ein. Das angezeigte Ergebnis ist ›0/0‹. Sein Blick fährt über den Titel des Dokuments: ›Winterschlaf und Winterruhe. Ein Referat‹, seufzt, lehnt sich zurück und klickt auf die Zurück-Taste. In diesem Moment richtet er seinen Blick auf eine Grafik einer aufgeschlagenen Buchseite neben sich. Er beugt sich über die Seite und liest eine Passage mit der Zwischenüberschrift ›Winterschlaf‹. Sein Blick wandert zu einer anderen Passage mit der Zwischenüberschrift ›Wechselwarme Tiere‹. »Ah, ok«, sagt der Lehrer und öffnet auf dem Laptop das Programm OneNote.

Die zusammenfassende Interpretation konzentriert sich auf die Frage, wie sich im Zusammenspiel zwischen menschlichem Akteur und digitalen Medien mediale Praktiken der Unterrichtsvorbereitung zeigen: Die Lehrperson ist von Dingen umgeben, die digitale Kulturtechniken nahelegen. Durch deren räumliche Anordnung, die in den Videoaufnahmen sichtbar werden, kann auf unterschiedliche Wertigkeiten geschlossen werden: So sind digitale und analoge Artefakte (sichtbar sind Ordner und Buch) an der Praxis beteiligt, beide werden von dem Akteur auch einbezogen. Die oberflächliche Recherche mittels Suchmaschine erscheint als unübersichtlich, sperrig und am Ende nicht zielführend. Dies wird insbesondere in den Eye-Tracking-Aufnahmen evident, wenn dadurch sichtbar wird, wie digitale Medien genutzt werden. Die aufgeführte Praktik kann so gelesen werden, dass das digitale Artefakt den menschlichen Akteur zu einer speziellen Rechercheprak-tik auffordert, nämlich eine Dokumentensuche mit Schlagwort. Eine kongruente Praktik im Buch, z. B. die Suche per Glossar, konnte in der sich anschließenden Praktik für den Akteur »Buch« hingegen nicht beobachtet werden. Das Buch evoziert eine andere Informationsgewinnungsstrategie in der Praxis: die Orientierung an Textgestaltungselementen, wie uns ebenfalls die Blicke des Akteurs zeigen. Gleichwohl ist die Praktik des Recherchierens sowohl mit der Anordnung der Körper im Raum als auch mit der Anordnung der Suchergebnisse auf den angezeigten Webseiten verbunden. Nur direkt ›greif-‹ und ›sichtbare‹ Dinge werden in die Praktik einbezogen: Das Eye-Viewing macht wiederum deutlich, dass nur die ersten Treffer angeklickt und keine Texte gelesen werden, Informationen werden überflogen. Netzwerkpraktiken – bspw. durch Verlinkung im Wikipedia-Artikel oder der Reformulierung von Suchstichworten – werden nicht auf ihre Passung zur Praktik der Informationsgewinnung geprüft. Der Lehrer ist enttäuscht (von sich? von der Technik?). Das Seufzen und Zurücklehnen sind dessen hör- und sichtbare Zeichen. Und wie durch Ironie werden dann die im Buch gefundenen Informationen, die in den Blick geraten, als passend eingeschätzt.

In Bezug auf die eingesetzten Erhebungsinstrumente ist diese Szene interessant, weil das Suchverhalten des Lehrers in Kombination mit den Eye-Tracking-Daten und den Daten der Videokamera detailliert analysiert werden kann und in Kombination sich die Besonderheit der Situation erst erschließt. Das

Körperverhalten – insbesondere das Beugen und das Greifen nach etwas – in Kombination mit dem Blickverhalten (›kursorisches Suchen‹) machen die mediale Praktik des Recherchierens für die Forschenden gut sichtbar. In der »Suche« als Wissenspraktik (Macgilchrist et al. 2023) zeigt sich in der Rekonstruktion des dann Sichtbaren die Widerständigkeit digitaler Artefakte, die der Lehrperson durch die in sie eingeschriebene Suchpraktik nicht ad hoc das passende Ergebnis liefert. Das digitale Artefakt zeigt sich widerständig, während das Zusammenspiel mit dem analogen Artefakt (Schulbuch) nahtlos aufgeführt werden kann, wodurch es zum ›Ersatz‹ des Artefakts – und damit zu Mehraufwand – kommt, um die Praktik erfolgreich zur Aufführung zu bringen. Es liegt nahe, dass dies die bisherige, routinisierte Praxis des Recherchierens zu sein scheint. Sichtbar wird an dieser Szene deutlich eine buchkulturelle Prägung von Schule (Böhme 2006), da für die digitalisierungsbezogenen Praktiken in diesem Beispiel ähnliche Routinen im Zusammenspiel (noch) nicht verfügbar scheinen. Die beobachtete Praktik der Vorbereitung des Zeigens wird nicht nur durch Medien unterstützt, sondern wird zu einem gemeinsamen Arrangement. Der Laptop und damit der sich aufspannende Wissensraum nimmt zentral an der Praktik des Unterrichtvorbereitens teil. Der Lehrende wird quasi zum ›Sachersteller‹ und die Medien geben durch ihre Affordanzen bestimmte Handlungen vor.

6. Fazit und Ausblick

In diesem Beitrag wird die Frage aufgeworfen, wie Eye-Tracking-Technologie es ermöglicht, Medienpraktiken phänomensensibel zu visualisieren und damit das Instrumentarium der Qualitativen Videoanalyse in einer Kultur der Digitalität erweitert. Durch das Eye-Viewing-Verfahren wird ein Fall zur Vorstellung gebracht, der Öffnungs- und Schließungstendenzen der pädagogischen Praxis mit digitalen Medien zeigt: Digitale Medien öffnen einen Zugriff auf schier endlose Möglichkeiten an Tools und Informationen über die visuelle und auditive Aufbereitung der in Schule verhandelten Dinge und können gleichzeitig zu Orientierungslosigkeit führen. Im Gegensatz zu Lehrmittelprodukten fehlt den digitalen Artefakten die didaktische Struktur, die typischerweise durchdacht und auf das Wesentliche reduziert ist. Sie müssen angepasst oder ihnen muss sich angepasst werden – bisher etablierte und routinisierte Praktiken geraten dadurch ins Stocken. Erst mit Anpassung, Ersetzung oder Mehraufwand finden sie ihren routinisiert wirkenden Eingang in den Unterricht. Digitale Medien evozieren also »widerständige Praktiken« (vgl. Neto Carvalho/Schiefner-Rohs 2024), die aber nicht unbedingt empirisch von außen durch Beobachtung sichtbar werden, sondern sich in ›Mikro‹-Praktiken manifestieren. Die dargestellte Szene zeigt auch, dass digitale Medien nicht nur im ›Gebrauch‹ sind, sondern dass sie konstitutiv in die Praxis involviert sind, diese in eine bestimmte Richtung lenken und

zusätzlichen Aufwand oder Abbruch verursachen. Gleichzeitig kann eine Beziehung zwischen Artefakten (z. B. Stifte, Bücher und Bildschirme), digitalen Daten (z. B. Suchergebnisse, Software-Schnittstellen) und zugrundeliegenden Plattformen (z. B. Google, Wikipedia) empirisch aufgedeckt werden. Sichtbar geworden sind Elemente einer ›relationalen Agency‹ zwischen Lehrperson und digitalem Medium durch entsprechende Affordanzen und Handlungsofferten.

Mit der Methode des Eye-Viewing gelingt also ein Zugang, der noch stärker die »leisen Daten« (vgl. Hirschauer 2001) sichtbar macht, um so Erkenntnisse über stumme/passive/nicht-sagbare Dinge des kulturellen Alltags sowie die eher subtil präfigurierende Handlungsmacht digitaler Medien in sozialen Settings bzw. die medialen Prozessqualitäten, welche Sozialität mitstrukturieren, empirisch sichtbar machen zu können. Allerdings wird auch deutlich, dass die Erhebung sozio-medialer Praktiken nicht allein auf die Eye-Tracking-Technik beschränkt werden sollte: Denn erst in Kombination des Eye-Tracking-Materials mit dem Material der Videokamera, welches aus der Perspektive der Beobachtenden gewonnen wurde, kann die jeweilige Praktik der Lehrperson wirklich detailreich beschrieben werden. Die räumliche Komposition der Dinge der Unterrichtsvorbereitung, das Körperhandeln der Lehrperson, insbesondere das Etwas-Zeigen, das Über-etwas-Beugen und das Nach-etwas-Greifen sowie die Gestik und Mimik in Kombination mit dem Blickverhalten, machen Praktiken im Zusammenspiel mit Technologie sichtbar. Damit kann aufgezeigt werden, wo mediale pädagogische Praktiken selbstverständlich in den Alltag der Professionellen integriert werden und wo sie sich dagegen sperrig verhalten. Ein praxistheoretischer Zugang mittels Eye-Viewing macht es sowohl möglich, die Verwobenheit von Menschen und Artefakten (Allert/Asmussen 2017) als auch die Entstehung von Subversion und Widerständigkeit aller an einer Praktik beteiligten Akteure sichtbar zu machen, um so eine empirische Erweiterung von Theorien zu professionellem Medien- und Lehrhandeln zu leisten – fernab von kompetenz- und nutzungstheoretisch geprägten Diskursen. Bei der Verwendung von zwei mobilen Eye-Tracking-Brillen wird es zudem möglich sein, Praktiken der zwischenmenschlichen Bezugnahme mit ihrer Sprache-Körper-Raum-Ding-Verwobenheit detaillierter zu untersuchen.

Trotz des Potenzials der Eye-Tracking-Technologie, vielversprechende Anwendungen in der Bildungsforschung zu ermöglichen, bringt sie jedoch neue forschungspraktische Herausforderungen hervor. Die liegen z. B. in der Invasivität der Technologie, da diese Kameras das Erscheinungsbild der im Fokus der Aufmerksamkeit stehenden Teilnehmenden verändert und damit eine noch offensichtlichere Form der Öffentlichkeit schafft. In Lehr-Lern-Settings reagieren die beobachteten Akteure, dies kann ich im Anschluss an alle meine Erhebungen behaupten, folgerichtig stärker auf die Eye-Tracking-Technologie als auf andere Aufzeichnungsgeräte und gewöhnen sich langsamer an deren Anwesenheit. Daher sollte die Technologie je nach Intimität oder Vulnerabilität des Settings

vorsichtig eingeführt, unterschiedlich intensiv genutzt oder ganz darauf verzichtet werden. In meinen explorativen Studien kam es zudem immer wieder zum Verlust der Kalibrierung, einer Unterbrechung der WLAN-Verbindung und einem Qualitätsverlust durch Wechsel der Lichtverhältnisse.

Die Technologie stellt uns aber auch vor forschungsethische Herausforderungen: Wie Krug und Heuser (2018) beobachtet haben, werden Eye-Tracking-Brillen von Studienteilnehmenden oft nicht nur als invasiv, sondern sogar als schmerzhaft empfunden. Vor dem Hintergrund des aktuellen forschungsethischen Diskurses in der qualitativen Sozialforschung und ihrer eigenen Rekonstruktionen von Interaktionen zwischen Studienleitenden und Studienteilnehmenden plädieren sie dafür, die Teilnehmenden situationsangemessen zu informieren, z. B. auch darüber, dass sie die Erhebung jederzeit abbrechen können, wenn sie sich unwohl fühlen. Neben ihrer Verantwortung für den Datenschutz sollten die Studienleiter:innen also auch für die körperliche Sicherheit der Teilnehmenden verantwortlich sein. Diese Verantwortung lässt sich als eine »Ethics of Care« begreifen, die sich fallspezifisch ausformen kann. Zu dieser gehöre auch, dass die Datenerhebungen so durchzuführen sind, dass Arbeitsabläufe in nicht-experimentellen Settings nicht gestört werden.

Literatur

Alirezabeigi, Samira/Masschelein, Jan/Decuypere, Mathias (2020): Investigating digital doings through breakdowns: a sociomaterial ethnography of a Bring Your Own Device school. In: Learning, Media and Technology 45, H. 2, S. 193–207. https://doi.org/10.1080/17439884.2020.1727501.

Allert, Heidrun/Asmussen, Michael (2017): Bildung als produktive Verwirklichung. In: Allert, Heidrun/Asmussen, Michael/Richter, Christoph (Hrsg.): Digitalität und Selbst. Bielefeld: transcript, S. 27–68.

Bareither, Christoph (2023): Ethnografische Affordanzanalyse. In: Stollfuß, Sven/Niebling, Laura/Raczkowski, Felix (Hrsg.): Handbuch Digitale Medien und Methoden. Wiesbaden: Springer Fachmedien, S. 1–21.

Bettinger, Patrick (2020): Materialität und digitale Medialität in der erziehungswissenschaftlichen Medienforschung. In: MedienPädagogik 15, S. 53–77. https://doi.org/10.21240/mpaed/jb15/2020.03.04.X.

Bettinger, Patrick (2023): Sozio-Medialität in der ethnografischen Forschung. Skizze eines medienpädagogischen Zugangs. In: Schiefner-Rohs, Mandy/Heymann, Carina/Neto Carvalho, Isabel (Hrsg.): Ethnographie und Videographie pädagogischer Praktiken. »Ein-Blicke« in Projekte der Schul- und Unterrichtsforschung in einer Kultur der Digitalität. Bad Heilbrunn: Klinkhardt, S. 68–83.

Bock, Bettina/Pirkko, Dresing (2021): Multimodale Lesestrategien. Eine empirische Rekonstruktion auf Basis von Laut-Denk- und Eye-Tracking-Daten. https://journals.ub.uni-koeln.de/index.php/midu/article/view/906, https://doi.org/10.18716/OJS/MIDU/2021.0.2 (10.10.2024).

Böhme, Jeanette (2006): Schule am Ende der Buchkultur. Medientheoretische Begründungen schulischer Bildungsarchitekturen. Bad Heilbrunn: Klinkhardt.

Bohnsack, Ralf (2009): Qualitative Bild- und Videointerpretation. Opladen/Farmington Hills: Barbara Budrich.

Bohnsack, Ralf (2021): Der Film als Erhebungsinstrument: Videographie und Dokumentarische Methode. In: Geimer, Alexander/Heinze, Carsten/Winter, Rainer (Hrsg.): Handbuch Filmsoziologie. Wiesbaden: Springer Fachmedien, S. 569–590.

Breidenstein, Georg (2010): Einen neuen Blick auf schulischen Unterricht entwickeln: Strategien der Befremdung. In: Heinzel, Friederike/Thole, Werner/Cloos, Peter/Köngeter, Stefan (Hrsg.): Auf unsicherem Terrain. Ethnographische Forschung im Kontext des Bildungs- und Sozialwesens. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 205–215.

Colemann, Gabriella (2010): Ethnographic Approaches to Digital Media. In: Annual Review of Anthropology, 39, S. 487–505.

Cortina, Kai/Müller, Katharina/Häusler, Janina/Stürmer, Kathleen/Seidel, Tina/Miller Kevin (2018): Feedback mit eigenen Augen. In: Beiträge zur Lehrerinnen- und Lehrerbildung 36, H. 2, S. 208–222.

Cramer, Florian (2015): What Is ›Post-digital‹? In: Berry, David M./Dieter, Michael (Hrsg.): Postdigital Aesthetics: Art, Computation and Design. London: Palgrave Macmillan UK, S. 12–26.

Dindar, Katja/Korkiakangas, Terhi/Laitila, Aarno/Kärnä, Eija (2017): An interactional »live eye tracking« study in autism spectrum disorder: combining qualitative and quantitative approaches in the study of gaze. In: Qualitative Research in Psychology 14, H. 3, S. 239–265. https://doi.org/10.1080/14780887.2017.1290174.

Dinkelaker, Jörg (2018): Selektion und Rekonstruktion. In: Moritz, Christine/Corsten, Moritz (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer, S. 153–165.

Dinkelaker, Jörg/Herrle, Matthias (2009): Erziehungswissenschaftliche Videographie. Wiesbaden: VS Verlag für Sozialwissenschaften.

Esau, Tatiana/Fletcher, Stefan (2018): Prozessorientierte Analyse von konstruktiven Problemlöseprozessen auf Basis von Eye-Tracking-Aufnahmen. In: Journal of Technical Education (JOTED) 6, H. 1. https://doi.org/10.48513/joted.v6i1.116.

Fawns, Tim/O'Shea, Clara (2018): Distributed learning and isolated testing: tensions in traditional assessment practices. www.academia.edu/93989955/Distributed_learning_and_isolated_testing_tensions_in_traditional_assessment_practices (10.10.2024).

Geertz, Clifford (1983): Dichte Beschreibung: Beiträge zum Verstehen kultureller Systeme. Frankfurt am Main: Suhrkamp-Taschenbuch Wissenschaft.

Herrle, Matthias/Hoffman, Markus/Proske, Matthias (2022): Unterrichtsgestaltung im Kontext digitalen Wandels: Untersuchungen zur soziomedialen Organisation Tablet-gestützter Gruppenarbeit. In: Zeitschrift für Erziehungswissenschaft 25, H. 6, S. 1389–1408. https://doi.org/10.1007/s11618-022-01099-8.

Hirschauer, Stefan (2001): Ethnographisches Schreiben und die Schweigsamkeit des Sozialen. In: ZfS, 30, S. 429–451.

Holler, Judith/Kendrick, Kobin H. (2015): Unaddressed participants' gaze in multi-person interaction. In: Frontiers in Psychology 6, H. 98. https://doi.org/10.3389/fpsyg.2015.00098.

Hopkins, Julian (2016): The Concept of Affordances in Digital Media. In: Friese, Heidrun/Rabane, Gala/Nolden, Marcus/Schreiter, Miriam (Hrsg.): Handbuch Soziale Praktiken und Digitale Alltagswelten. Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-08460-8_67-1.

Hörning, Karl H./Reuter, Julia (2004): Doing Culture: Neue Positionen zum Verhältnis von Kultur und sozialer Praxis. Bielefeld: transcript.

Kalthoff, Herbert/Cress Torsten/Röhl, Tobias Röhl (2016): Einleitung: Materialität in Kultur und Gesellschaft. In: Kalthoff, Herbert/Cress, Torsten/Röhl, Tobias (Hrsg.): Materialität. Herausforderungen für die Sozial- und Kulturwissenschaften. Paderborn: Wilhelm Fink, S. 11–41.

Keller, Lena/Cortina, Kai S./Müller, Katharina/Miller, Kevin F. (2022): Noticing and weighing alternatives in the reflection of regular classroom teaching: Evidence of expertise using mobile eye-tracking. In: Instructional Science 50, H. 2, S. 251–272. https://doi.org/10.1007/s11251-021-09570-5.

Kesselheim, Wolfgang/Hottinger, Christoph (2019): Besucherexperimente im Science Center: Welche Einsichten in die Herstellung von Wissen in Interaktion werden erst durch die Zusammenschau von Audio-, Video- und Eye-Tracking-Daten möglich? https://dh-abstracts.library.virginia.edu/works/10887 (10.10.2024).

Kleiner, Konrad/Rakoczi, Gergely/Krimm, Stefan (2012): Blickbewegungsverläufe (mobiles Eye-Tracking) als Element sportdidaktischer Lehr-Lern-Forschung. Purkersdorf: Hollinek.

Krämer, Sybille (2008): Medium, Bote, Übertragung: kleine Metaphysik der Medialität. Frankfurt am Main: Suhrkamp.
Krug, Maximilian/Heusser, Svenja (2018): Ethik im Feld: Forschungspraxis in audiovisuellen Studien. In: Forum qualitative Sozialforschung 19, H. 3. https://doi.org/10.17169/fqs-19.3.3103.
Luhmann, Niklas (1992): Die Wissenschaft der Gesellschaft. Frankfurt am Main: Suhrkamp.
Macgilchrist, Felicitas (2019): Digitale Bildungsmedien im Diskurs. In: Aus Politik und Zeitgeschichte 69, H. 27/28, S. 18–23.
Macgilchrist, Felicitas (2021): Theories of Postdigital Heterogeneity: Implications for Research on Education and Datafication. In: Postdigital Science and Education 3, H. 3, S. 660–667. https://doi.org/10.1007/s42438-021-00232-w.
Mayer, Christian W./Rausch, Andreas/Seifried, Jürgen (2023): Analysing Domain-Specific Problem-Solving Processes within Authentic Computer-Based Learning and Training Environments by Using Eye-Tracking: A Scoping Review. In: Empirical Research in Vocational Education and Training 15, S. 1–27. https://doi.org/10.1186/s40461-023-00140-2.
McGrenere, Joanna/Ho, Wayne (2000): Affordances: Clarifying and Evolving a Concept. In: Graphics Interface. https://graphicsinterface.org/wp-content/uploads/gi2000-24.pdf (10.12.2024).
McLuhan, Marshall (1992): Die magischen Kanäle: = »Understanding media«, Econ classics. Düsseldorf: Econ.
Mohn, Bina E. (2008): »Die Kunst des dichten Zeigens«. In: Binder, Beate/Neuland-Kitzerow, Dagmar/Noack, Karoline (Hrsg.): Kunst und Ethnographie. Berliner Blätter 46/2008. Münster: Lit, S. 61–72.
Mohn, Elisabeth/Amann Klaus (2005): Lernkörper. Kamera-ethnographische Studien zum Schülerjob. Göttingen: IWF.
Mohn, Elisabeth/Amann, Klaus (2006a): Lernkörper. Kamera-ethnographische Studien zum Schülerjob. Begleitpublikation zur DVD. Göttingen.
Mohn, Elisabeth/Amann, Klaus (2006b): Permanent Work on Gazes. Video Ethnography as an Alternative Methodology. In: Knoblauch, Hubert/Raab, Jürgen/Soeffner, Hans-Georg/Schnettler, Bernt (Hrsg.): Video Analysis. Methodology and Methods. Qualitative Data Analysis in Sociology. Frankfurt am Main: Peter Lang, S. 173–181.
Moritz, Christine/Corsten, Michael (Hrsg.) (2018): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer Fachmedien.
Neto Carvalho, Isabel (2017): Gymnasium und Ganztagsschule. Videographische Fallstudie zur Konstitution pädagogischer Ordnung. Wiesbaden: Springer VS.
Neto Carvalho, Isabel/Schiefner-Rohs, Mandy (2024): Digitale Artefakte als widerständige Akteure des Unterrichts. In: ZISU – Zeitschrift für interpretative Schul- und Unterrichtsforschung 13, H. 1, S. 109–125. https://doi.org/10.3224/zisu.v13i1.08.
Neto Carvalho, Isabel/Schiefner-Rohs, Mandy/Troxler, Carina (2022): Eye-Viewing als Verfahren der erziehungswissenschaftlichen Medienforschung. Fokussierte Videographie mit Eye-Tracking-Technologie am Beispiel der Erforschung von medienbezogenen Praktiken in Schule und Unterricht. In: ZQF 23, H. 2, S. 262–277. https://doi.org/10.3224/zqf.v23i2.09.
Passoth, Jan-Hendrik (2017): Hardware, Software, Runtime: das Politische der (zumindest) dreifachen Materialität des Digitalen. In: Behemoth – A Journal on Civilisation 10, H. 1, S. 57–73.
Rabenstein, Kerstin/Wagener-Böck, Nadine/Macgrilchrist, Felicitas/Bock, Annekatrin (2022): Interferenzen in digitalen Praktiken der Bereitstellung von unterrichtlichen Aufgaben: Ethnographische Beobachtungen in der Pandemie. In: Sozialer Sinn 23, H. 2, S. 297–315. https://doi.org/10.1515/sosi-2022-0016.
Reckwitz, Andreas (2003): Grundelemente einer Theorie sozialer Praktiken. In: Zeitschrift für Soziologie 32, H. 4, S. 282–301. https://doi.org/10.1515/zfsoz-2003-0401.
Reckwitz, Andreas (Hrsg.) (2010): Unscharfe Grenzen. Perspektiven der Kultursoziologie. 2. Auflage. Sozialtheorie. Bielefeld: transcript.
Röhl, Tobias (2013): Dinge des Wissens. Schulunterricht als sozio-materielle Praxis. Stuttgart: Lucius & Lucius.
Spiegler, Juliane/Herrle, Matthias/Hoffmann, Markus/Proske, Matthias (2023): Produktion von Leistung im digital-gestützten Unterricht. In: Schiefner-Rohs, Mandy/Heymann, Carina/Neto Carvalho, Isabel (Hrsg.): Ethnographie und Videographie pädagogischer Praktiken. »Ein-Blicke« in

Projekte der Schul- und Unterrichtsforschung in einer Kultur der Digitalität. Bad Heilbrunn: Klinkhardt, S. 120–140.
Sprenger, Priska (2021): Prozesse bei der strukturierenden Mengenwahrnehmung und strukturnutzenden Anzahlbestimmung von Kindern im Elementarbereich: Eine Eye-Tracking Studie. Wiesbaden: Springer Spektrum.
Strübing, Jörg (2014): Grounded Theory. Wiesbaden: Springer VS.
Stukenbrock, Anja (2018): Mobile dual eye-tracking in face-to-face interaction. In: Brône, Geert/Oben, Bert (Hrsg.): Advances in Interaction Studies 10, S. 265–302.
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie. Wiesbaden: Springer Fachmedien.
Wernet, Andreas (Hrsg.) (2009): Einführung in die Interpretationstechnik der objektiven Hermeneutik. Qualitative Sozialforschung. 3. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.
Wiedbusch, Megan/Lester, James/Azevedo Roger (2023) A Multi-Level Growth Modeling Approach to Measuring Learner Attention with Metacognitive Pedagogical Agents. In: Metacognition and Learning 18, H. 2, S. 465–494. https://doi.org/10.1007/s11409-023-09336-z.
Zillien, Nicole (2008): Die (Wieder-)Entdeckung der Medien. Das Affordanzkonzept in der Mediensoziologie. In: Sociologia Internationalis. Internationale Zeitschrift für Soziologie, Kommunikations- und Kulturforschung 2, S. 161–181.

Sehen und soziale Interaktion im Kunstmuseum

MET x EMKA als neuer Ansatz der Videoanalyse

Luise Reitstätter, Seda Pesen und Dirk vom Lehn

Zusammenfassung: Sehen und soziale Interaktion sind im Kunstmuseum konstitutiv, indem dort eine visuelle Auseinandersetzung mit Werken und ein Aufeinandertreffen mit anderen Personen stattfindet. Die Kombination von mobilem Eye-Tracking (MET) und ethnomethodologischer Konversationsanalyse (EMKA) ermöglicht es, diese ästhetischen und sozialen Erfahrungen methodisch zu erschließen. In der Studie »True to Life« in der Österreichischen Galerie Belvedere realisierten wir eine erste Anwendung von MET x EMKA als neue Form der Videoanalyse. Der Artikel ergründet über die Analyse von drei Situationen der Kunstbetrachtung, wie Praktiken der räumlichen Orientierung, des visuellen Erfassens von Werken, des Lesens von Labels sowie vokale und gestische Äußerungen multimodal verbunden sind. Während MET detaillierte Blickbewegungen aus der Egoperspektive von Besucher:innen offeriert, kann mit EMKA der Blick am Körper und im interaktionellen Geschehen im Raum verortet werden. Das Kunstmuseum fungiert als Ort der methodischen Exploration, um in Folge MET x EMKA auch für andere Settings fruchtbar zu machen.

Schlüsselwörter: mobiles Eye-Tracking, ethnomethodologische Konversationsanalyse, Sehen, soziale Interaktion, Kunstmuseum

1. Einleitung

»So, here I want to direct your attention to the artist's mother over there«. Mit diesen Worten beginnt John das Gespräch mit seiner Begleiterin Mary in der Österreichischen Galerie Belvedere. Er möchte ihre Aufmerksamkeit auf ein Exponat lenken, das er zuvor bereits allein betrachtet hat. Es handelt sich um das Porträt »Die Mutter des Künstlers« des österreichischen Malers Franz Rumpler. Nachdem John Mary an der gegenüberliegenden Wand im Ausstellungsraum abgeholt hat, huscht ihr Blick vorerst über viele andere Werke am Weg, bis sie – geleitet von Johns Bewegung im Raum, der Ausrichtung seines Körpers, seinen verbalen Hinweisen wie seinem Fingerzeig – das von ihm intendierte Werk erkennt. Sie stellt sich neben John vor das Porträt als dieser, angeregt durch die

Provenienzangabe »donation by the artist's widow«, scherzt, dass Rumplers Witwe die Schwiegermutter wohl loswerden wollte. Der Witz, den Mary mit einem amüsierten »Ohhh« quittiert, führt zu einem bleibenden Eindruck dieser Kunstrezeption. Im Interview nach dem Ausstellungsbesuch verweist Mary, gefragt nach dem Label, das den stärksten Eindruck bei ihr hinterlassen hat, auf das gemeinsam konstruierte Narrativ zu Schwiegermutter und -tochter. Die sachliche Werk- und Provenienzangabe, gepaart mit Johns Kommentar, zeigt sich in der gemeinsamen Sinnstiftung weitaus eindrücklicher als andere Labels mit längeren Werkbeschreibungen, die Mary aufmerksam allein gelesen hat.

Die Beschreibung dieser Kunstrezeption aus der Studie »True to Life« stellt ein Beispiel dar, wie Sehen und soziale Interaktion im Kunstmuseum zwar nicht immer zeitgleich stattfinden, aber doch häufig miteinander verwoben sind, sich ergänzen und bedingen. Sehen im Kunstmuseum findet sowohl als ästhetische Auseinandersetzung mit Objekten (wie etwa Gemälden oder Skulpturen) als auch als soziale Interaktion mit Personen (wie Begleiter:innen, Aufsichten und anderen Besucher:innen vor Ort) statt. Der Blick ist im Kunstmuseum als stark visuell konnotiertem aber auch sozial geprägtem Raum konstitutiv. Der Blick verortet sich dabei sowohl am Körper (über Haltung, Kopfbewegung und Blickrichtung) als auch im Raum (über Betrachtungsposition, Sichtfeld und Foki der visuellen Aufmerksamkeit) sowie am Gegenüber (sofern weitere Personen in visuelle Handlungen involviert sind). Gleichsam verbindet sich das Sehen im Kunstmuseum mit anderen Praktiken wie dem Gehen, Zuhören und Sprechen. Ausgehend von der Multimodalität und Komplexität von Praktiken der Kunstrezeption im Museum stellt sich die Frage, wie diese am besten empirisch zu erfassen und analytisch zu ergründen sind. Eine Möglichkeit, der wir in diesem Artikel nachgehen, ist die Kombination von mobilem Eye-Tracking (MET) und videobasierter ethnomethodologischer Konversationsanalyse (EMKA), kurz MET x EMKA.

Der Artikel resultiert aus einer ersten Anwendung von MET x EMKA im Rahmen der Studie »True to Life«, bei der Mitglieder des Labors für empirische Bildwissenschaft (Laboratory for Cognitive Research in Art History, CReA Lab) der Universität Wien und der Forschungsgruppe Work, Interaction and Technology (WIT) des King's College London zusammenarbeiteten. In seinem Aufbau stellt der Artikel theoretische Prämissen und empirische Untersuchungen zum Sehen und sozialer Interaktion im Kunstmuseum der beiden Methoden MET und EMKA voran, um in Folge ihre individuellen Stärken und blinden Flecken sowie das Potenzial ihrer Kombination zu beleuchten. Nach einer Erläuterung des Erhebungskontexts der Studie »True to Life« präsentiert der Artikel den Ausstellungsbesuch von John und Mary als Fallstudie mit dem Fokus auf drei konkrete Situationen der Kunstrezeption. Exemplarisch soll damit zum einen ein Einblick in Formen des Sehens und der sozialen Interaktion im Kontext des Kunstmuseums gegeben werden. Zum anderen soll die detaillierte Abhandlung der Fallstudie (von der Erhebung über die Auswertung bis möglicher Nachnutzung der

Forschungsdaten) methodologische Schlussfolgerungen von MET x EMKA als neuem Ansatz der Videoanalyse eröffnen sowie spätere Nutzungen abseits des Kunstmuseums ermöglichen.

2. Sehen und soziale Interaktion im Kunstmuseum

Besucher:innen des Kunstmuseums sind in ambivalenten Verhältnissen verortet: Während das traditionelle Ideal die stille, fokussierte Kunstbetrachtung vorsieht, lässt sich der Museums- und Ausstellungsbesuch als vorwiegend gemeinsam vollzogene Praktik gleichsam als Prozess der dynamischen sozialen Interaktion verorten (vgl. Reitstätter/Christidou 2024). So besuchen etwa 80 bis 95 Prozent der Besucher:innen Museen nicht etwa allein, sondern in Begleitung (Black 2012, S. 21; Kirchberg 2010, S. 178; Lindner 2016, S. 327). Während Besuchsmotivationen – mit dem Interesse, gemeinsame Zeit zu verbringen, oder auch Geselligkeit im öffentlichen Raum des Museums zu erfahren – häufig sozial einzuordnen sind (vgl. Debenedetti 2003; Jafari/Taheri/von Lehn 2013; Schönherr/Glaser 2023, S. 103 f.), bieten Museen auch die Möglichkeit, Zeit allein zu genießen und introspektive Erfahrungen zu machen (vgl. Art Fund 2019; Buchholz 2000; Høffding et al. 2020). Insbesondere im Kunstmuseum fungiert die solitäre Museumspraxis als Disposition für ästhetische Erfahrungen (vgl. Greenblatt 1991; Pelowski et al. 2014), verbunden mit einer Vertrautheit mit dem Ritual des Kunstmuseums (vgl. Duncan 1995) und Formen des »bürgerlichen Sehens« (vgl. Bennett 2006). Begleitung hilft hingegen, sich mit dem potenziell unbekannten Ort vertraut zu machen und den Ausstellungsbesuch sozial zu rahmen (vgl. Hood 1993; Tröndle 2019, S. 116), unterstützt durch einen Paradigmenwechsel hin zu stärker partizipativen Museen (vgl. Henning 2006; Kavanagh/McSweeney 2016; Simon 2010; Simon 2016). Sehen und soziale Interaktion zeigen sich dahingehend als Parameter der Kunstbetrachtung im Museum, die es theoretisch zu bestimmen wie empirisch zu untersuchen gilt.

Nachdem lange Zeit Experimente zum Blickverhalten nahezu ausschließlich im Labor und mit digitalen Reproduktionen und statischen Eye-Trackern durchgeführt wurden, bieten seit den frühen 2010er-Jahren mobile Eye-Tracking Systeme (vgl. Carvalho i. d. B.) die Möglichkeit, mit MET-Grundlagenforschung zum Sehen im Museum zu betreiben. Obgleich Milekic (2010) bereits damals postulierte, dass »eye tracking may prove to be the most powerful tool for museum studies«, waren erste MET-Studien im Museum aufgrund streng normierter Settings mit wenigen Werken und statischen Betrachter:innen, geringen Teilnehmer:innenzahlen oder ungenauen Blickdaten noch stark limitiert. Aktuelle MET-Studien profitieren von technischen und methodischen Weiterentwicklungen, um in natürlichen Ausstellungsbedingungen (ohne Vorgabe von Wegen oder Betrachtungszeiten) spezifischere Forschungsfragen (von ästhetischer

Erfahrung bis Usability) zu stellen sowie weitaus größere Samples (mit bis zu mehreren hundert Teilnehmer:innen) untersuchen zu können. So ermöglicht MET über kombinierte Videoaufnahmen der Augenbewegungen und des Sichtfeldes eine präzise Aufzeichnung dessen was, wie lange und in welcher Abfolge von Museumsbesucher:innen betrachtet wird. Bislang werden MET-Daten vor allem für quantitative Auswertungen des Blickverhaltens wie Betrachtungszeiten (etwa von Kunstwerken, Labels oder auch Personen), Heatmaps (Visualisierung von wenig bis häufig betrachteten Bereichen ähnlich einer Wärmebildkamera) oder Scanpaths (Visualisierung des Blickpfads) genutzt. Gleichsam wird für Methodenkombinationen mit Beobachtung und Befragung plädiert, um Formen des Sehens qualitativ kontextualisieren zu können (vgl. Reitstätter et al. 2025). MET-Daten ermöglichen es Forscher:innen jedoch nicht zu untersuchen, wie Sehen und Sehmuster in Interaktion mit anderen Besucher:innen, mit denen sie Ausstellungen gemeinsam explorieren oder die sich gleichzeitig in derselben Ausstellung aufhalten, entstehen.

Seit den 2000er-Jahren werden videobasierte EMKA in Museen durchgeführt, die es Forscher:innen erlauben die Interaktion von Besucher:innen mit vorhandenen Artefakten und vor allem die gemeinsame Rezeption von Besucher:innen multimodal zu verhandeln. Zuvor wurde die Methode videobasierte EMKA vor allem in den sogenannten Workplace Studies verwendet, um die Organisation von Interaktion in zumeist technisierten Arbeitsplätzen anhand von Video- und Audioaufnahmen zu analysieren (vgl. Knoblauch/Heath 1999). Im Fokus des Interesses steht, wie Akteur:innen durch ihre Interaktionen Kontexte hervorbringen, in die sie Aspekte der visuellen und materiellen Umgebung einbetten. Kontexte werden dabei fortlaufend durch Handlungen erneuert und gleichzeitig werden Handlungen durch den Kontext, in dem sie hervorgebracht werden, beeinflusst oder überformt (Heritage 1984, S. 242). EMKA segmentiert Handlungen in ihre Bestandteile und arbeitet heraus, wie diese als sequenzielle Ordnung temporal aufeinander folgen und sinnhaft aufeinander bezogen sind. Im Museum sind videobasierte EMKA daran interessiert, die soziale Organisation der Handlungen von Besucher:innen Moment für Moment offen zu legen. Das sozial produzierte Erlebnis gründet sich zum einen darauf, dass viele Besucher:innen Museen und Ausstellungen gemeinsam mit anderen erkunden und zum anderen dort häufig auf andere Besucher:innen treffen, mit denen sie ihre Erkundung abstimmen. So können Besucher:innen beispielsweise nur an ein Gemälde herantreten, wenn sich dort nicht schon andere Besucher:innen aufhalten. Ein gemeinsames Erlebnis eines Werks geschieht wiederum nicht allein dadurch, dass zwei Personen nebeneinander davorstehen, sondern wird erst durch die Interaktion miteinander unter möglicher Einbindung vorhandener Ressourcen wie Ausstellungstexte oder Audioguides erzielt (vgl. vom Lehn 2025).

Vergleicht man die Methoden MET und videobasierte EMKA wird deutlich, dass sie beide mit besonderen Stärken wie gleichzeitig blinden Flecken

in der Museumsforschung angewandt werden (vgl. Reitstätter et al. 2025; vom Lehn 2025). So findet bei MET über die Videodokumentation des Museumsbesuches aus der Egoperspektive buchstäblich eine methodische Integration der Besucher:innensicht statt. Zudem kann MET als mobile Methode (vgl. Büscher/Urry 2009) dem bewegten Sehen gerecht werden, wenn Objekte über wiederholte Blicke von verschiedenen Standpunkten Beachtung finden (vgl. Carbon 2017; Reitstätter/Galter/Bakondi 2022). In diesem visuell komplexen Aneignungsverhalten liefert MET umfangreiche und präzise Blickbewegungsdaten, wenngleich diese durch die Perspektive der Betrachter:innen körperlos bleiben, das Interaktionsverhalten im Raum nicht gesamtheitlich erfassen und erst kombiniert mit weiteren Methoden Prozesse der Sinnstiftung ergründet werden können. Die videobasierte EMKA kann wiederum mit ihrem analytischen und methodischen Instrumentarium Details der Hervorbringung und Gestaltung von Handlungen in ihrer Orientierung zu vorangegangenen und nachfolgenden Handlungen sichtbar machen. So erlaubt die Methode etwa nachzuvollziehen, wie Besucher:innen das Herantreten an und das Verlassen von Kunstwerken mit anderen organisieren und wie sie durch ihre vokalen und körperlichen Handlungen ihre Eindrücke füreinander nachvollziehbar machen (vgl. vom Lehn 2012; Heath/vom Lehn 2004). Videoanalysen mit EMKA sind jedoch visuell limitiert, indem Forscher:innen auf Kopfhaltung, Kopfbewegungen, teils sichtbare Augenbewegungen oder auch Gesprächsverweise vertrauen müssen, um Rückschlüsse zum Blickverhalten zu ziehen. Die Methodenkombination MET x EMKA eröffnet nun die Möglichkeit, detailgenau das singuläre Blickverhalten der Betrachtenden im größeren sozialen und räumlichen Setting des Ausstellungsraums zu verorten sowie die ebenfalls blickgeleitete soziale Interaktion von Besucher:innen als multimodales Zusammenspiel von körperlichen und vokalen Auseinandersetzungen zu untersuchen.

3. Erhebungskontext Studie »True to Life«

Die Studie »True to Life« ist Teil des übergeordneten FWF-DFG Forschungsprojekts »The Museum Gaze« (I 5342-G), das als transdisziplinäre Zusammenarbeit zwischen Kunsthistoriker:innen und Museumswissenschaftler:innen der Universität Wien, Informatiker:innen der Technischen Universität München und Kurator:innen der Österreichischen Galerie Belvedere Grundlagenforschung zur Kunstrezeption im Museum betreibt. Ziel der Studie »True to Life« – benannt nach dem englischen Titel der Sonderausstellung »Lebensnah« zum Thema Realismus, die sich über acht thematische Räume erstreckte – war es, die Auswirkung von kuratorischen Interventionen auf Betrachtungs- und Interpretationsmuster zu untersuchen. Während wir in der ersten Kondition (8. bis 16. Oktober 2022) die Ausstellung so untersuchten, wie sie von den Kurator:innen Franz Smola und

Kerstin Jesse konzipiert wurde, nahmen wir in der zweiten Kondition (24. bis 30. Oktober 2022) verschiedene kuratorische Interventionen an Kunstwerken und Labels vor. Konkret wurden Werke der ersten Kondition mit Werken populärer Künstler (Gustav Klimt, Vincent van Gogh) ausgetauscht, mit zeitgenössischer Kunst (Lisl Ponger) ersetzt, innerhalb des Raums neu gehängt oder entfernt. Labels wurden in ihrer Zeichenzahl verdoppelt, an allen Werken im Raum ergänzt oder nach unterschiedlichen Narrativen umgeschrieben.

Die Methodenkombination von MET x EMKA wurde im Rahmen der Zusammenarbeit des CReA Lab mit der Forschungsgruppe WIT vom King's College London in der Studie »True to Life« angewandt. Zusätzlich wurde im Anschluss an den Ausstellungsbesuch ein Fragebogen und ein stimulus-gestütztes Interview eingesetzt, um das Besuchs- und Interpretationsmuster einer ganzen Ausstellung ganzheitlich zu erfassen: Während MET detaillierte Blickbewegungsdaten von Besucher:innen während des gesamten Ausstellungserlebnisses von Raum 1 bis Raum 8 lieferte, gab EMKA Aufschluss über die körperlichen Erfahrungen und sozialen Interaktionen von Besucher:innen in Raum 2. Der quantitative Fragebogen erhob Daten zur Ausstellungserfahrung, dem persönlichen Kunstinteresse (basierend auf dem Vienna Art Interest and Knowledge Fragebogen VAIAK, vgl. Specker et al. 2018), dem Museumsbesuch, der Usability der Eye-Tracking-Brille sowie grundsätzliche soziodemografische Kontextinformationen. Das stimulusgestützte Interview fügte unseren Daten eine weitere audiovisuelle Ebene mit Einblick in Sinnstiftungsprozesse der Besucher:innen nach dem Ausstellungsbesuch hinzu. Konkret zeichneten wir mit einer Handykamera auf einem Stativ auf, wie Besucher:innen – gefragt nach dem Raum, Kunstwerk und Label mit der stärksten sowie schwächsten Resonanz – mit den Reproduktionen der Stimuli der Ausstellung verbal und gestisch interagierten.

Reguläre Museumsbesucher:innen über 18 Jahre wurden zur Teilnahme an der Studie eingeladen. Nach einer Einführung in das Verfahren und der Unterzeichnung einer Einverständniserklärung wurden die Teilnehmer:innen mit einer MET-Brille (Tobii Pro Glasses 3) und einem Aufnahmegerät in einer kleinen Hüfttasche ausgestattet. Anschließend besuchten sie die acht Räume der Ausstellung in ihrem eigenen Tempo und in der von ihnen gewählten Gruppenkonstellation. Wenn sich die Besucher:innen zu zweit für eine Studienteilnahme entschieden, wurden sie zusätzlich zur EMKA-Videoaufzeichnung in zwei der acht Räume mit einem Funkmikrofon für die Erhebung ihrer Gespräche ausgestattet. Nach dem Ausstellungsbesuch gaben sie die technischen Geräte ab, füllten den Fragebogen aus und nahmen am stimulus-gestützten Interview teil. Insgesamt nahmen 212 Besucher:innen (c1: n=108, c2: n=104) mit einem Durchschnittsalter von 35,72 Jahren (SD=15,76) an unserer Studie teil. 55,7 Prozent der Stichprobe identifizierten sich als weiblich, 43,4 Prozent als männlich und 0,9 Prozent als divers. Der primär internationale touristische Besuchskontext zeigte sich im Sample mit 37 Nationalitäten sowie dem hohen Anteil an Erstbesucher:innen

(91,0 Prozent) in einer Gruppenkonstellation (81,1 Prozent). Das bereits in früheren Studien als hochgebildet charakterisierte Kunstmuseumspublikum (vgl. Bourdieu/Darbel 1969; Donnat 1993) bestätigte sich auch in unserem Sample mit einer Vielzahl von Universitätsabschlüssen (71,2 Prozent) und einer hohen Kunstaffinität (4,68 von 7 Punkten) der Teilnehmer:innen, wenngleich sich das durchschnittliche Kunstengagement im Mittelfeld bewegte (3,44 von 7 Punkten).[1]

4. Datenanalyse Fallbeispiel »John und Mary«

Im Folgenden konzentrieren wir uns bei der Datenanalyse auf zwei der 212 Teilnehmer:innen, nämlich auf den Besuch von John (p094) und Mary (p093)[2] in der Ausstellung am 14. Oktober 2023 unter den von den Forscher:innen unveränderten Ausstellungsbedingungen (erste Kondition). Mary ist weiblich und 39 Jahre alt, John ist männlich und 44 Jahre alt. Beide Teilnehmer:innen haben einen Universitätsabschluss. Sie leben in Nottingham und besuchten das Museum im Rahmen einer Reise. Für Mary war es der erste Besuch in der Österreichischen Galerie Belvedere, John hatte das Museum bereits zwei- oder dreimal zuvor besucht. Beide gaben an, das Museum in Begleitung besucht zu haben. Während Mary jedoch ihre Begleitung John als »Kollegen« spezifizierte, stufte John Mary als »Freundin« ein. Marys Kunstaffinität ergab einen mittleren Wert von 3,6 – ihr eher niedriges Kunstengagement liegt bei 2,6; John hingegen weist eine höhere Kunstaffinität von 4,9 auf und ein ebenso gesteigertes Kunstengagement von 4,6 (bei 7 als jeweils höchstem möglichen Wert). Für die Analyse konzentrieren wir uns auf Raum 2, in dem zwei Videokameras in den jeweils diagonal gegenüber liegenden Ecken des Raums positioniert waren. Im Ausstellungsraum wurden unter dem Titel »Das echte Leben« zehn Gemälde ausgestellt, die über Porträts und Genrebilder Menschen in alltäglichen Situationen darstellen. Sieben dieser Gemälde waren mit Labels versehen, die kontextuelle Informationen zu den Kunstwerken boten.

Der Datenbestand für diese Fallstudie besteht aus a) zwei MET-Videosequenzen aus Raum 2 von John (4:01 Minuten) und Mary (2:54 Minuten), b) zwei

1 Die Kunstaffinität und das Kunstengagement wurden mit Fragen zum subjektiven Kunstinteresse aus dem VAIAK-Fragebogen (Part A) gemessen, der aus zwei Sets von Statements besteht, auf denen sich die Teilnehmer:innen mittels einer Likert-Skala verorten. Das erste Set beinhaltet Statements zu einer allgemeinen Kunstaffinität, das zweite Set beinhaltet Fragen zum Engagement innerhalb des Kunstfelds. Üblicherweise werden bei der Berechnung des Part A die Sets nicht separiert. Im Zuge unserer Studie haben wir diese Trennung jedoch vorgenommen, da sie Aufschluss über verschiedene Involvierungen im Kunstfeld geben.

2 Die Namen entsprechen nicht ihren realen, unbekannten Namen, sondern wurden den Teilnehmer:innen von den Forscher:innen zur Identifizierung und Personalisierung gegeben.

EMKA-Videos (7:29 Minuten und 7:36 Minuten) von den beiden in Raum 2 positionierten Kameras, c) zwei ausgefüllten Fragebögen und d) einem Video des gemeinsam geführten stimulus-gestützten Interview (21:32 Minuten). Um einen Einblick in die Möglichkeiten von MET x EMKA als neuem Ansatz der Videoanalyse zu geben, erörtern wir anhand von drei exemplarischen Situationen, wie John und Mary ihren Ausstellungsbesuch in Raum 2 gestalten. Wir beginnen mit Johns Kunstbetrachtungsverhalten in der Auseinandersetzung mit dem Werk »Zur Schule« (1883) von Edmund Krenn und der kontextbezogenen Raumsituation (MET + EMKA). Als zweite Situation analysieren wir, wie Mary sich in Form eines stark linearen Betrachtungsverhalten dem Ausstellungsraum inhaltlich nähert (EMKA + MET). Die dritte bereits eingangs beschriebene Situation offenbart, wie John und Mary sich das Werk »Die Mutter des Künstlers« (1878) von Franz Rumpler, ausgehend von Johns Hinweis auf dessen spezifische Provenienz, gemeinsam erschließen (MET x EMKA). Die Analyse des Seh- und Interaktionsverhaltens der beiden Teilnehmer:innen anhand von MET- und EMKA-Videodaten veranschaulicht sozial motivierte Muster der Kunstbetrachtung im Umgang mit den ausgestellten Werken und Labels. Methodisch erörtern die drei situativen Beispiele, wie die jeweiligen Kombinationen von MET und EMKA als Ergänzung oder Zusammenspiel funktionieren können.

a) Johns Kunstbetrachtung (MET + EMKA)

John betritt den Ausstellungsraum wesentlich früher als Mary. Nach einem kurzen, schweifenden Blick durch den Raum, der über die verschiedenen Objekte und kopräsenten Besucher:innen gleitet, orientiert er sich zu seiner Rechten und beginnt mit der Betrachtung des ersten Kunstwerks, nämlich »Zur Schule« (Abb. 1a). Johns Blick fokussiert sich auf ein Detail des Kunstwerks »Zur Schule«, das seine Entscheidung der spezifischen Gehrichtung (gegen den Uhrzeigersinn und ohne das Lesen des Raumtexts zu seiner Linken) begründet. Diese visuell-ästhetische Entscheidung ist charakteristisch für die folgenden Minuten seines Besuchs: Obwohl seine Blicke gelegentlich zwischen Werk und Label hin und her wechseln, konzentriert er sich stark auf das Erstere. Mit zunehmender Verweildauer bewegt er sich auf das Werk »Zur Schule« zu, um die Bestandteile der Komposition detaillierter zu betrachten. Sein Blick geht nur kurz über das Label, um ihn danach wieder auf das Werk zu richten. Das EMKA-Video ergänzt diese MET-Aufnahmen und zeigt, wie John seinen Körper im Raum bewegt, sich am Ausstellungsdisplay und anderen Besucher:innen orientiert. Mittels der EMKA-Aufnahme lässt sich feststellen, dass der Raum, als John ihn betrat, recht belebt war. Dies verleitete ihn dazu, andere Besucher:innen zu umgehen und sich seinen eigenen Betrachtungsraum vor den Werken zu schaffen (Abb. 1b). Sein Verhalten mutet dabei selbstverständlich wie selbstbewusst an.

Die Organisation seiner Handlungen sowie die Orientierung am Visuellen statt am Textuellen lassen Rückschlüsse über sein im Fragebogen bestätigtes habituell routiniertes Kunstbetrachtungsverhalten zu. Im weiteren Besuchsverlauf in Raum 2 wird er die Kunstwerke vor den Labels privilegieren, mit einer hohen durchschnittlichen Betrachtungszeit von 29,59 Sekunden pro Werk und nur 0,07 Sekunden pro Label.

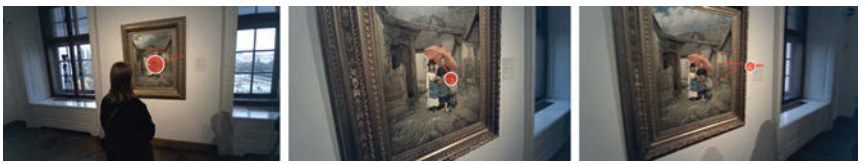


Abbildung 1a: MET Stills, die Johns Annäherung an das Werk »Zur Schule« und dessen genaue Betrachtung zeigen

Abbildung 1b: EMKA Stills, die zeigen, wie John den Raum 2 betritt und seinen Platz vor dem Werk »Zur Schule« einnimmt

Methodisch zeigt dieses Beispiel, wie Johns MET-Video zwar eine genaue Beobachtung seines Blickverhaltens zulässt, jedoch kaum Aufschluss über das Geschehen um ihn herum gibt. Vielmehr zeigen die MET-Videos mit der Darstellung seines Sichtfeldes und des Blickpunkts seinen entkörperlichten Blick. Das EMKA-Video ergänzt die MET-Aufnahmen, da es in der Außenperspektive zeigt, wie John seinen Körper im Raum bewegt und sich am Display und an anderen Besucher:innen orientiert. Da die Videokamera gegenüber vom Werk »Zur Schule« positioniert und John damit lediglich von hinten zu sehen ist, ist wiederum das MET-Video notwendig, um mehr über sein spezifisches Blickverhalten am Werk zu erfahren. Eine solche Ergänzung von MET mit EMKA ist insofern methodisch gewinnbringend als die Blickbewegungsdaten mit Informationen über das Geschehen im Raum und der genauen Betrachtungsposition als weitere Komponenten der Kunstbetrachtung augmentiert werden können.

b) Marys Kunstbetrachtung (EMKA + MET)

Abbildung 2a: EMKA Stills, die zeigen, wie Mary den Ausstellungsraum mit einer klaren Orientierung zum Raumtext betritt

Marys Betrachtungsverhalten steht in starkem Kontrast zu Johns. Nach einem kurzen Orientierungsblick, der weitaus schneller und weniger erkennend ist, findet Mary beim Betreten einen vornehmlich leeren Ausstellungsraum mit deutlich weniger Besucher:innen als bei John vor. Sie beginnt rasch mit der Lektüre des einführenden Raumtextes zu ihrer Linken. Mit dieser Gehrichtung befindet sich Mary im direkten Sichtfeld der EMKA-Aufnahme, sodass ihre Körperbewegungen gut nachvollzogen werden können. Nach einer ausführlichen Lektüre des Raumtextes setzt sie ihren Besuch entlang der Wand fort und orientiert sich an der linearen Hängung der Werke (Abb. 2a). Ihr Körper bleibt dabei in etwa in der gleichen Distanz zur Wand. Ihre Betrachtung erweist sich als strikt rhythmisch: Beginnend mit einem kurzen Blick auf das Werk liest sie folgend das Label gründlich, bevor sie das entsprechende Werk erneut kurz ansieht. Gleichförmig setzt sie dieses Betrachtungsverhalten für jedes Kunstwerk und jedes Label fort. Dies zeigt auch die Visualisierung der MET-Daten mittels eines Scanpaths, der die Fokussierung auf die Texte und auf die wesentlichen Bildbestandteile, nämlich die Gesichter der Dargestellten, nachweist (Abb. 2b). Weiter ist am Scanpath zu erkennen, dass Mary die Texte nicht nur überfliegt, sondern vollständig bis zum Ende liest. Durchschnittlich widmet Mary Kunstwerken 7,73 Sekunden und Labels 4,44 Sekunden ihrer visuellen Aufmerksamkeit in diesem Raum. Dass Mary sich stark am Textuellen orientiert, kann im Zusammenhang mit ihrem im Fragebogen erhobenen niedrigeren Kunstengagement stehen. So konnten wir in einer früheren MET-Studie nachweisen, dass weniger habituelle Besucher:innen tendenziell mehr lesen, da sie die textliche Informationen als Anleitung für ihre Kunstrezeption nutzen (Reitstätter/Galter/Bakondi 2022, S. 137).

Abbildung 2b: MET Scanpath, der über die Visualisierung der sequentiellen Blickpfade, die aus Fixationen (Punkte) und Sakkaden (Linien) bestehen, Marys rhythmisches Blickverhalten zeigt

Dieses Beispiel zeigt methodisch, wie EMKA-Videodaten – bei einem guten Blickwinkel und ohne weitere Betrachter:innen, die die Videoaufzeichnung verdecken – sehr aussagekräftige Daten zur Bewegung im Raum und Betrachtungspositionen liefern können. Damit gibt EMKA Aufschluss über Muster des »embodied seeing« in Relation zum Display (lineare Hängung der Werke an der Wand) und den angebotenen Ressourcen (Raumtext, Werke und Labels). Gleichsam können MET-Blickbewegungsdaten das im EMKA-Video ersichtliche lineare Betrachtungsmuster von Mary entlang der Wand durch Visualisierungen wie einem Scanpath gut ergänzen. Dieser veranschaulicht die Abfolge von Marys Betrachtung über Fixationen als Punkte und Sakkaden als Linien und fasst so das Sequentielle ihrer Betrachtung in einem aussagekräftigen Bild zusammen. Die Augmentierung der EMKA-Videodaten durch MET-Daten ermöglicht dahingehend, sequenzielle Kunstrezeptionsmuster im Stehen und Gehen sowohl körperlich im Raum verortet wie kondensiert als Blickmuster in Relation zu den angebotenen Ressourcen zu verstehen.

c) John und Marys gemeinsame Kunstbetrachtung (MET x EMKA)

Obwohl John zunächst auch dem linearen Aufbau der Ausstellung, aber im Unterschied zu Mary gegen den Uhrzeigersinn folgt, unterbricht er diesen Weg beim Werk »Die Mutter des Künstlers« (1878) von Franz Rumpler. Auf diesem Werk ist die Porträtierte aus einer leicht erhöhten Perspektive auf einem Stuhl sitzend in Tracht abgebildet. Ihre Hände sind auf dem Schoß gefaltet, ihr Blick ist auf die Betrachter:innen gerichtet. Das Label des Werks gibt Auskunft über grundlegende Angaben: Den Namen und die Lebensdaten des Künstlers, den Titel des Werks, das Entstehungsjahr, das Material sowie die Provenienz. Aus der Kombination der Lebensdaten des Künstlers und der Provenienz lässt sich schließen, dass das Kunstwerk von der Witwe des Künstlers ein Jahr nach dessen Tod dem Museum geschenkt wurde. Diese Erkenntnis verändert Johns Blickverhalten im Vergleich mit seinen vorangegangenen Werkrezeptionen im Raum. Während sich

John zuvor hauptsächlich auf das Kunstwerk (durchschnittlich 29,59 Sekunden) und weniger auf die zugehörigen Labels (durchschnittlich 0,07 Sekunden) konzentrierte, schaut er bei »Die Mutter des Künstlers« zunächst das Label, dann das Gemälde selbst und schließlich erneut das Label an, bevor er Mary aufsucht, um ein Gespräch mit ihr zu initiieren. Insgesamt betrachtete er das Kunstwerk 27,62 Sekunden und das Label 20,37 Sekunden lang. Die Betrachtungszeit für das Label erhöht sich damit – im Vergleich mit der durchschnittlichen Betrachtungszeit aller anderen bis dahin gelesenen Labels in Raum 2 – um das 291-fache.

Während im MET-Video Johns Betrachtungs- und Richtungswechsel sichtbar wird, zeigt das EMKA-Video, wie er sich Mary annähert. Nachdem er sich umgedreht und Mary erkannt hat, geht er langsam auf die gegenüberliegende Wand zu und schaut sich rasch die Werke dort an, bevor er, hinter ihr stehend, das Gespräch beginnt. Wie bei EMKA üblich, wird zunächst das Gespräch der Teilnehmer:innen transkribiert (Transkript 1.1), bevor das Gespräch in Beziehung zu den körperlichen und visuellen Aktionen untersucht wird (Transkript 1.2a; Transkript 1.2b).

Transcript 1.1. - John and Mary - (Belvedere)

```
1   J:  so:: (.6) he:r ehrm (.3) I want to direct your
2       attention to the ar(.)tist's mother (.3) overhe:re
3     (1.5)
4       rather ster:n portrai:t obviously she's got on her
5       very best headscar:f to be here (.3) donated by the
6       ar:tist's widow:
7   M:  Ohh::::::
```

Transkript 1.1

Die zu analysierende Situation beginnt mit der Zusammenkunft der beiden an der gegenüberliegenden Wand. Nachdem John »Die Mutter des Künstlers« wie beschrieben betrachtet und das dazugehörige Label gelesen hat, geht er zu Mary und bittet sie, sich das Werk anzusehen: »so:: (.6) he:r ehrm (.3) I want to direct your attention to the ar(.)tist's mother (.3) overhe:re« (Transkript 1.1., Zeile 1–2). Beide bewegen sich in Folge in die von John geleitete Richtung. Während seiner Äußerung ist der Kopf von John auf Mary gerichtet, die sich dem Exponat zu ihrer linken Seite widmet. Während er spricht, behält Mary ihre visuelle Aufmerksamkeit auf dem Gemälde zu ihrer Linken, lehnt sich aber zu John zurück und zeigt so eine doppelte Orientierung (Transkript 1.2a, Bild 1.1). Johns verbaler Hinweis auf »die Mutter des Künstlers« (Transkript 1.1, Zeile 2) wird begleitet von einem Schritt nach rechts in Marys Nähe. Gleichzeitig nickt er deutlich in Richtung des Porträts an der gegenüberliegenden Wand, um ihr vorzuschlagen, wohin sie gehen und was sie sich als Nächstes ansehen soll (Transkript 1.2a, Bild 1.2). Mary

folgt Johns Vorschlag sich zu dem Porträt zu begeben, indem sie ihre körperliche Orientierung auf dieses ausrichtet, während sie ihren Blick über andere Gemälde entlang der Ausstellungswand gleiten lässt (Transkript 1.2a, Bild 1.3). Als John »overhe:re« sagt (Transkript 1.2a, Zeile 2), scheint Mary ihren Blick so auszurichten, dass er mit dem von John übereinstimmt (Transkript 1.2a, Bild 1.4). Beide gehen nun näher an das Porträt heran, während sie in dessen Richtung schauen.

Bild 1.1

Bild 1.2

Bild 1.3.

Bild 1.4

John
Sprechen

so::------he:r ehrm---I want to direct your attention to the ar.tist's mother---overhe:re

Schritte
LR_____steht_____ L R L

Mar

Schritte
LR_____steht_____ L R L

Transkript 1.2a

Als sie in der Nähe des Porträts ankommen, werden Johns und Marys Schritte kürzer. So positionieren sie Schritt für Schritt ihre Körper nebeneinander im gemeinsamen Betrachtungsraum. Während sie sich aufstellen, beginnt John sein Gespräch erneut, indem er das Gemälde als »rather ster:n portrai:t« und die darin abgebildete Frau beschreibt, »obviously she's got on her very best headscar:f for thee:« (Transkript 1.1, Zeile 4–5). Während er spricht, widmen sich beide dem Kunstwerk (Transkript 1.2b, Bild 1.5–1.6) und machen fünf weitere Schritte bevor sie ihren Standpunkt vor dem Kunstwerk einnehmen. Johns Aussprache von »for thee:« geht mit einer sichtbaren Verschiebung seines Blicks einher; er wendet sich vom Porträt zu der Stelle, an der Mary steht (Transkript 1.2b, Bild 1.7), bevor er seinen Blick wieder auf das Exponat richtet. Diese Handlung kann als Versuch gewertet werden, Marys visuelle Orientierung und Reaktion auf das Werk zu überprüfen (vgl. Stukenbrock 2018; Stukenbrock/Dao 2019). Nach einer kurzen Pause von 0,3 Sekunden dreht sich John um und deutet mit dem rechten Zeigefinger auf das Label, das rechts neben dem Gemälde angebracht ist. Während er diese Zeigegeste ausführt, richtet er seinen Blick auf das Label aus und spricht mit dem Hinweis auf die Schenkung einen Teil des Inhalts aus (Transkript 1.2b, Bild 1.8).

John und Mary – Belvedere

Bild 1.5. Bild 1.6. Bild 1.7 Bild 1.8

John

Sprechen
rather ster:n portrai:t obviously she's got on her very best headscar:f for thee:----------donated

Gesten
x-----y

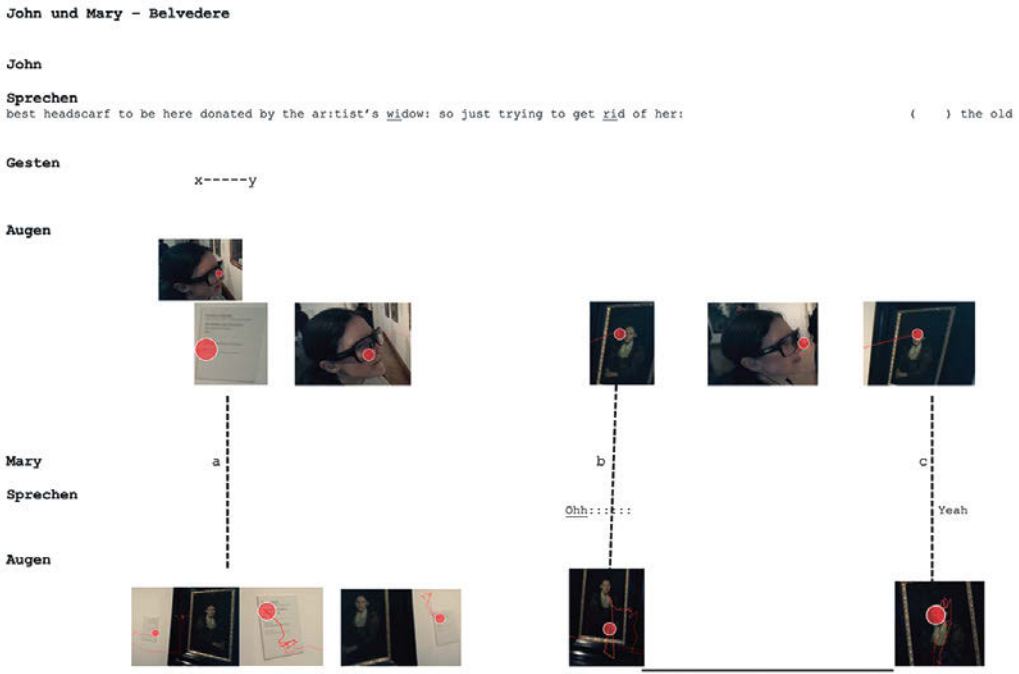
Schritte
L R L R LR_____LR

Mary

Schritte
L R L R LR_____LR

Transkript 1.2b

Die Analyse des EMKA-Videos zeigt auf, dass beide sich an den visuellen Merkmalen orientieren und diese hörbar beschreiben. Durch das laute Vorlesen einzelner Textkomponenten teilen sie Informationen, die ihnen helfen, die Ausstellung zu verstehen. Damit greift die Analyse ein Thema auf, das in der aktuellen Museums- und Besucher:innenforschung zentral ist, nämlich die Sinnstiftung als eine visuelle Praxis, die durch soziale Interaktion entsteht (vgl. Falk/Dierking 2013; Schmitt/Labour 2022). Obwohl es uns möglich ist, Veränderungen in der körperlichen und visuellen Orientierung von Teilnehmer:innen zu isolieren, können wir anhand dieser EMKA-Videoaufnahmen nicht abschließend beurteilen, wohin John und Mary genau schauen. Die Integration von MET-Daten bietet dahingehend die Möglichkeit, die visuellen Orientierungen der Teilnehmer:innen während einer Interaktionssequenz zu ermitteln. Das Transkript 1.2c visualisiert entsprechend, wie MET x EMKA als kombinierte Methode eine analytische Perspektive bieten kann, bei der das komplexe Feld der Interaktion Aktivitäten des Blicks inkludiert. Transkript 1.2c zeigt von oben nach unten zunächst Johns verbale Äußerungen, Gesten und Augenbewegungen, gefolgt von Marys verbalen Äußerungen und Augenbewegungen. Der rote Punkt visualisiert den Blick der Person; die gestrichelte Linie zwischen den Abbildungen zeigen Momente der synchronen visuellen Orientierung an. So wird beispielsweise deutlich, dass John während der humorvollen Erklärung der Provenienz mehrfach den Blick zwischen Mary und dem Label wechselt, um ihre Aufmerksamkeit auf Letzteres zu lenken *und* ihre Reaktion zu erkennen. Erst als Mary ihr Verstehen zu erkennen gibt (»Ohh:«) und letztlich zustimmt (»Yeah«), richten beide ihre Blicke zeitgleich auf das Porträt.



Transkript 1.2c[3]

3 Bei der Transkription des Gesprächs, der Gesten und der visuellen Handlungen folgen wir Heath, Hindmarsh und Luff (2010), integrieren aber auch Bilder aus den Videoaufzeichnungen, die mittels MET aus der Egoperspektive der Teilnehmer:innen aufgezeichnet wurden. Diese Aufnahmen der Szenenkamera werden in der MET-Forschung, die in erster Linie quantitative Auswertungen und Visualisierungen wie Heatmaps, Scanpaths usw. verwendet, weitgehend ignoriert.

Die hier kombinierte Anwendung von MET- und EMKA-Daten bietet also mehrere Vorteile. Die Hinführung zum situativen Geschehen mittels EMKA-Transkriptionen (Transkript 1.1, 1.2a, 1.2b) ermöglicht zum einen, einen detaillierten Eindruck der Interaktion zu erhalten und damit auch zu bestimmen, welche MET-Daten für die weitere Untersuchung relevant sein werden. Die analoge Setzung der ausgewählten MET-Stills zweier Teilnehmer:innen innerhalb eines Transkripts erlaubt zum anderen die Verortung der jeweiligen Blicke im interaktionellen Geschehen (Transkript 1.2c). Folglich können mit der kombinierten Anwendung von MET x EMKA soziale Interaktionen in Beziehung zum Blickverhalten gestellt und Momente des Austausches, der Vermittlung und der gemeinsamen Sinnstiftung im Museum genauer bestimmt werden. Diese gemeinsame Sinnstiftung beim Werk »Die Mutter des Künstlers« bestätigte sich auch beim stimulus-gestützten Interview mit den beiden Teilnehmer:innen im Anschluss an ihren Ausstellungsbesuch. Hier verwies Mary bei der Frage nach dem Label mit der größten Resonanz auf den Kommentar von John, der sie veranlasste, ein gemeinsames Narrativ zur Schwiegermutter und -tochter zu kreieren.

5. MET x EMKA als neuer Ansatz der Videoanalyse

Mit dem Fallbeispiel von John und Mary konnten wir aufzeigen, wie die Kombination von MET und EMKA als neue Form der Videoanalyse genutzt werden kann, um zu untersuchen, wie Sehen und soziale Interaktion im Kunstmuseum in und durch nahtlose Organisation von Handlungen vollzogen werden. Ziel war es, über einen detaillierten Einblick in die Datenaufbereitung inklusive Transkription und verschiedene Modi der Datenanalyse spezifische Situationen der Kunstrezeption als Kontext unserer Videoanalyse zu erörtern sowie methodische Schlussfolgerungen aus der konkreten Kombination von MET und EMKA zu ziehen. Dabei wurde deutlich, dass je nach Fokussierung des Erkenntnisinteresses eine unterschiedlich gewichtete Nutzung der beiden Methoden bzw. auch eine Kontextualisierung der Videodaten durch weitere Methoden (hier Fragebogen und stimulus-gestütztes Interviews) sinnvoll ist. Entsprechend lassen sich folgende Punkte, die bei den verschiedenen Zusammenführungen von MET und EMKA charakteristisch sind, konkludieren:

In der Betrachtung von Einzelpersonen und ihrer Verhaltensweisen im musealen Raum erscheint es ratsam, die Methoden in der Kombination unterschiedlich zu gewichten. Denn je nach Fall und Fragestellung kann es von Vorteil sein, die Analyse entweder mit MET- oder eben mit EMKA-Videodaten zu beginnen, wie die Analyse unserer beiden individuellen Kunstrezeptionen von John und Mary gezeigt hat. So sind die von MET-Brillen aufgezeichneten Blickbewegungen inhärent in die Produktion körperlicher Handlungen verwoben, die wiederum durch EMKA-Videos aufgezeichnet werden. Zugleich ermöglichen

EMKA-Aufnahmen, die situativen Bedingungen zu reflektieren und die Organisation der körperlichen Handlungen und Positionen in Relation zu anderen Besucher:innen zu untersuchen. Schlussfolgerungen bezüglich der Merkmale von Blickbewegungen, die bei der alleinigen Nutzung von EMKA nicht eindeutig sind, können wiederum durch die Zunahme von MET-Daten bestätigt, widerlegt oder differenzierter bestimmt werden. Beide Methoden ergänzen sich dahingehend in ihrer visuellen Qualität: Während MET-Videodaten sich durch ihre dynamischen Einstellungen aufgrund der raschen, kleinteiligen Kopfbewegungen auszeichnen, können EMKA-Aufnahmen mittels ihrer statischen Einstellung genauer Aufschluss über die Haltung des Körpers und die Verortung des Körpers im Raum geben. Je nach Untersuchungsschwerpunkt und schlicht der Sichtbarkeit der zu untersuchenden Aspekte am Videomaterial kann also eine der Methoden analytisch stärker in Betracht gezogen werden.

Wenn in der Analyse singulärer Teilnehmer:innen folglich eine Privilegierung einer der Videomethoden unter zeitgleicher Betrachtung der zweiten als sinnvoll erscheint, wird gerade die Zusammenführung beider Methoden in der Untersuchung der Interaktionen zwischen Teilnehmer:innen fruchtbar. Dieser methodische Zusammenschluss von MET x EMKA, in unserem Beispiel von John und Mary exemplifiziert, ermöglicht spezifische visuelle Praktiken zu identifizieren, die der Organisation von Interaktionen zugrunde liegen. Über die Kombination von MET- mit EMKA-Videodaten kann herausgearbeitet werden, wie das individuelle Betrachtungshandeln durch die Orientierung an den Handlungen des Interaktionspartners transformiert wird. Die regelmäßige Überprüfung der visuellen Orientierung eines:einer Partner:in stellt dabei eine gängige Praxis in der deiktischen Referenz dar (vgl. Stukenbrock 2018b). Mit MET x EMKA kann über das Zusammenspiel von Blick und sozialer Interaktion die sequentielle Anordnung referenzieller Praxis und gemeinsamer Handlungen erfasst werden. Dabei deuten unsere Daten darauf hin, dass es die unidirektionale Kontrolle der sich entwickelnden Orientierung eines:einer Partner:in als Reaktion auf Gespräche und körperliche Handlungen ist, welche das Fortschreiten der Interaktion am Exponat untermauert und die Grundlage für die Produktion einer gemeinsamen Erfahrung bildet.

Forschungspraktisch sind wiederum folgende Herausforderungen in der kombinierten Anwendung von MET x EMKA zu beachten: Hinsichtlich der Kosten sind die technisch fortschrittlichsten MET-Systeme sehr teuer. Die von uns verwendeten Tobii Pro Glasses etwa gehören zum höheren Preissegment. Ihre Anschaffung erfordert einen erheblichen finanziellen Aufwand, insbesondere da für die Analyse sozialer Interaktionen zwei Geräte notwendig sind. Im Gegensatz dazu bewegt sich das technisch aktuelle Videoequipment in einer deutlich niedrigeren Preisklasse. Die Leihe von Geräten oder die Kooperation mit Forschungspartner:innen, die solche bereits besitzen, ist dahingehend eine sinnvolle Alternative. Hinzu kommt bei MET die langwierige Datenaufbereitung, da die

automatisierte Objekterkennung bei dynamischen Aufzeichnungen wie jenen eines Museumsbesuchs aktuell nur bedingt greift und entsprechend zeitaufwändiger manueller Annotationen bedarf. In Kombination mit den EMKA-Daten ist es also allein vom Aufwand her notwendig, sich auf eine kleine Anzahl an Situationen zu beschränken. Des Weiteren erfordert der methodische Gleichschritt von MET x EMKA eine hohe theoretische Sensibilität. Da MET-Daten bisher vor allem in quantitativen und behaviouristischen Forschungskontexten genutzt wurden, müssen in der Verbindung mit der subjekt-orientierten, qualitativen EMKA noch entsprechende Terminologien für MET x EMKA gefunden wie geschärft werden. Zukünftige exemplarische Studien bzw. auch die Nachnutzung unserer bisher erhobenen Forschungsdaten sollen zeigen, welcher konkrete Nutzen im Vergleich zum Aufwand aus dem Mixed-Method-Ansatz von MET x EMKA zu ziehen ist.

Weitere Anwendungsgebiete von MET x EMKA, die wir im Kontext des Kunstmuseums aufgrund der möglichen Einsichten in das körperliche Sehen und die blickgelenkte soziale Interaktion untersuchen möchten, sind etwa der touristische Blick in Ausstellungen (im Gegensatz zum Besuchsverhalten des lokalen Publikums), Kunstbetrachtung und Kunstkommunikation (in ihrer gegenseitigen Verschränkung und möglichen Habitus-Steuerung) sowie das Daten im Museum (zur Exploration des romantischen Potenzials von Museen und der Agency von Kunstwerken). Derlei Anwendungsszenarien, die sich theoretisch sowohl aus der Kunstgeschichte als aus der Kultursoziologie speisen, fördern die inhaltliche Exploration von Museen und die Entwicklung methodischer Konventionen. Darüber hinaus sind Museen auch insofern eine produktive Forschungsumgebung, als hier neue Forschungstechniken und -technologien erprobt werden können, ohne Teilnehmer:innen zu gefährden. Dies steht im Gegensatz zu forschungsethisch sensiblen Orten wie beispielsweise Krankenhäusern oder komplexen technisierten Arbeitsplatzumgebungen, wo in der Vergangenheit Workplace Studies durchgeführt wurden (vgl. Luff/Hindmarsh/Heath 2000). Die Anwendung von MET x EMKA in Museen bietet daher schon jetzt die Möglichkeit methodische Entwicklungen anzustoßen, um in anderen Settings, an denen Sehen und soziale Interaktion ebenso intensiv zusammenspielt, diese neue Form der Videoanalyse in der nahen Zukunft einzusetzen.

Wenn auch die Konventionen der analytischen Praxis noch zu eruieren und neue Nutzungsszenarien noch zu erproben sind, sehen wir genau in dieser Offenheit das Potenzial der vorgeschlagenen Methodenkombination von MET x EMKA. Je nach Fragestellung und Aussagekraft der jeweiligen Videodaten kann die Gewichtung der einzelnen Methoden adaptiert werden, wie an den Beispielen der singulären Kunstbetrachtung von John und Mary deutlich wurde. Die Untersuchung von mehreren Teilnehmer:innen, wie im Beispiel der gemeinsamen Werk-Label-Exploration von John und Mary erläutert, eröffnet wiederum nochmals mehr die Möglichkeit, soziale und semantische Praktiken multimodal

zu untersuchen. Durch die Kombination von MET x EMKA findet wortwörtlich eine Perspektivenerweiterung statt. So ermöglicht der mehrdimensionale Blick, der sich einerseits aus der MET-Egoperspektive und andererseits aus der EMKA-Außenperspektive speist, ein vitales Verständnis der Handlungsabfolgen von Teilnehmer:innen in ihrer lokalen Umgebung zu erhalten. Bisherige Probleme, wie beispielsweise das durch EMKA-Daten nur zu erahnende Blickverhalten oder der entkörperte Blick beim MET, können so gelöst werden. Darüber hinaus können vorwiegend quantitative Auswertungen von Blickbewegungsdaten durch die interaktionelle Analyse kontextuell verankert und qualitativ eingebettet werden. Videoanalysen mit MET x EMKA eröffnen damit die Möglichkeit, Produkte des Sehens mit Praktiken des »embodied seeing« zu verschränken und über das interaktionale Geschehen kollektive Prozesse der Sinnstiftung weiter zu entschlüsseln.

Literatur

Art Fund (2019): Calm and Collected Museums and Galleries: The UK's Untapped Wellbeing Resource? www.bibli.artfund.org/m/300606adcae26d47/original/Insights_Calm_and_Collect.pdf (07.11.2024).

Bennett, Tony (2006): Civic Seeing: Museums and the Organization of Vision. In: Macdonald, Sharon (Hrsg.): A Companion to Museum Studies. Chichester: Wiley Blackwell, S. 263–281.

Black, Graham (2012): Transforming Museums in the Twenty-First Century. London, New York: Routledge.

Bourdieu, Pierre/Darbel, Alain (1969): L'amour de l'art. Paris: Les Éditions de Minuit.

Buchholz, Ester S. (2000): Echoes of Quietude: Alonetimes in Museums. In: Journal of Museum Education 25, H. 1–2, S. 3–8.

Büscher, Monika/Urry, John (2009): Mobile Methods and the Empirical. In: European Journal of Social Theory 12, H. 1, S. 99–116.

Carbon, Claus C. (2017): Art Perception in the Museum: How We Spend Time and Space in Art Exhibitions. In: i-Perception 8, H. 1, S. 1–15.

Debenedetti, Stéphane (2003): Investigating the Role of Companions in the Art Museum Experience. In: International Journal of Arts Management 5, H. 3, S. 52–63.

Donnat, Olivier (1993): Les publics des musées en France. In: Publics et Musées 3, H. 1, S. 29–46.

Duncan, Carol (1995): Civilizing Rituals: Inside Public Art Museums. London, New York: Routledge.

Falk, John/Dierking, Lynn D. (2013): The Museum Experience Revisited. Walnut Creek: Left Coast Press.

Greenblatt, Stephen (1991): Resonance and Wonder. In: Karp, Ivan/Lavine, Steven D. (Hrsg.): Exhibiting Cultures: The Poetics and Politics of Museum Display. Washington, D. C.: Smithsonian Institution Press, S. 42–57.

Heath, Christian/vom Lehn, Dirk (2004): Configuring Reception: (Dis-)Regarding the ›Spectator‹ in Museums and Galleries. In: Theory, Culture & Society 21, H. 6, S. 43–65.

Henning, Michelle (2006): Museums, Media and Cultural Theory. Maidenhead: Open University Press.

Heritage, John (1984): Garfinkel and Ethnomethodology. Cambridge, New York: Polity Press.

Høffding, Simon/Rung, Mette/Roald, Tone (2020): Participation and Receptivity in the Art Museum: A Phenomenological Exposition. In: Curator – The Museum Journal 63, H. 1, S. 69–81.

Hood, Marilyn G. (1993): Comfort and Caring: Two Essential Environmental Factors. In: Environment and Behavior 25, H. 6, S. 710–724.

Jafari, Aliakbar/Taheri, Babak/vom Lehn, Dirk (2013): Cultural Consumption, Interactive Sociality, and the Museum. In: Journal of Marketing Management 29, H. 15–16, S. 1729–1752.

Kavanagh, Jen/McSweeney, Katey (Hrsg.) (2016): Museum Participation: New Directions for Audience Collaboration. Edinburgh: MuseumsEtc.
Kirchberg, Volker (2010): Besucherforschung in Museen: Evaluation von Ausstellungen. In: Baur, Joachim (Hrsg.): Museumsanalyse: Methoden und Konturen eines neuen Forschungsfeldes. Bielefeld: transcript, S. 171–184.
Knoblauch, Hubert/Heath, Christian (1999): Technologie, Interaktion und Organisation: Die Workplace Studies. In: Schweizerische Zeitschrift für Soziologie 25, H. 2, S. 163–181.
Lindner, Bernd (2016): Soziodemographie des Museumspublikums. In: Walz, Markus (Hrsg.): Handbuch Museum: Geschichte – Aufgaben – Perspektiven. Stuttgart: J. B. Metzler, S. 323–329.
Luff, Paul/Hindmarsh, Jon/Heath, Christian (Hrsg.) (2000): Workplace Studies. Cambridge: Cambridge University Press.
Milekic, Slavko (2010): Gaze-Tracking and Museums: Current Research and Implications. In: Trant, Jennifer/Bearman, David (Hrsg.): Museums and the Web 2010: Proceedings. Toronto: Archives & Museum Informatics. http://www.archimuse.com/mw2010/papers/milekic/milekic.html (07.11.2024).
Pelowski, Matthew/Liu, Tao/Palacios, Victor/Akiba, Fuminori (2014): When a Body Meets a Body: An Exploration of the Negative Impact of Social Interactions on Museum Experiences of Art. In: International Journal of Education & the Arts 15, H. 14, S. 1–48.
Reitstätter, Luise/Christidou, Dimitra (2024): Alone Together? Solitary and Shared Visiting Practices of Pairs in the Art Museum. In: Museum Management and Curatorship, S. 1–21.
Reitstätter, Luise/Galter, Karolin/Bakondi, Flora (2022): Looking to Read: How Visitors Use Exhibit Labels in the Art Museum. In: Visitor Studies 25, H. 2, S. 127–150.
Reitstätter, Luise/Pesen, Seda/Rosenberg, Raphael/Kasneci, Enkelejda (2025): Mobiles Eye-Tracking: Blickverhalten in Ausstellungen untersuchen. In: Reitstätter, Luise/Schorr, Carla-Marinka (Hrsg.): Methoden der Ausstellungsanalyse. Bielefeld: transcript.
Schmitt, Daniel/Labour, Michel (2022): Making Sense of Visitors' Sense-Making Experiences: The REMIND Method. In: Museum Management and Curatorship 37, H. 3, S. 218–234.
Schönherr, Daniel/Glaser, Harald (2023): Kulturelle Beteiligung in Österreich: Besuch von Kulturveranstaltungen, Kultureinrichtungen und -stätten. SORA: Wien.
Simon, Nina (2010): The Participatory Museum. Santa Cruz, CA: Museum 2.0.
Simon, Nina (2016): The Art of Relevance. Santa Cruz, CA: Museum 2.0.
Specker, Eva et al. (2020): The Vienna Art Interest and Art Knowledge Questionnaire (VAIAK): A Unified and Validated Measure of Art Interest and Art Knowledge. In: Psychology of Aesthetics, Creativity, and the Arts 14, H. 2, S. 172–185.
Stukenbrock, Anja (2018a): Forward-Looking: Where Do We Go with Multimodal Projections? In: Deppermann, Arnulf/Streeck, Jürgen (Hrsg.): Time in Embodied Interaction Synchronicity and Sequentiality of Multimodal Resources. Amsterdam: John Benjamins, S. 31–68.
Stukenbrock, Anja (2018b): Blickpraktiken von SprecherInnen und AdressatInnen bei der Lokaldeixis: Mobile Eye Tracking-Analysen zur Herstellung von joint attention. In: Gesprächsforschung 19, S. 132–168.
Stukenbrock, Anja/Dao, Anh N. (2019): Joint Attention in Passing: What Dual Mobile Eye Tracking Reveals About Gaze in Coordinating Embodied Activities at a Market. In: Reber, Elisabeth/Gerhardt, Cornelia (Hrsg.): Embodied Activities in Face-to-Face and Mediated Settings. Frankfurt am Main, New York: Springer, S. 177–213.
Tröndle, Martin (Hrsg.) (2019): Nicht-Besucherforschung: Audience Development für Kultureinrichtungen. Wiesbaden: Springer VS.
vom Lehn, Dirk (2012): Configuring Standpoints: Aligning Perspectives in Art Exhibitions. In: Bulletin Suisse de Linguistique Appliquée 96, S. 69–90.
vom Lehn, Dirk (2025): Videobasierte Ethnomethodologische Konversationsanalyse: Soziale Interaktionen im Ausstellungsraum sequentiell erfassen. In: Reitstätter, Luise/Schorr, Carla-Marinka (Hrsg.): Methoden der Ausstellungsanalyse. Bielefeld: transcript.

Autovideographie

Zur Analyse digitaler Spielpraktiken mithilfe von Multi-Kamera-Videos[1]

Marcel Thiel-Woznica

Zusammenfassung: Im Beitrag wird die *Autovideographie* als Lösung für das methodologische Problem der Aufhebung der Differenz zwischen Innen- und Außenperspektive (Venus 2012) vorgestellt. Aufgehoben wird diese Differenz im Rahmen digitaler Spielpraktiken insofern, als sowohl die durch Interfaces hergestellte Außenperspektive (Beobachtung des Bildschirmgeschehens, Trennung zwischen Spieler:innenkörper und Spielwelt) als auch die Innenperspektive (Perspektivenübernahme, emobied make-believe) zusammenlaufen. Die Autovideographie wird als Methode vorgestellt, die durch die Kombination von autoethnografischen und videographischen Anteilen sowohl die Innenperspektive als auch die Außenperspektive der Praxis nachvollziehbar macht. Durch die simultane Aufnahme 1) des Bildschirmgeschehens, 2) der Frontalperspektive auf den Spielenden und 3) der Draufsicht-Perspektive auf das materielle Arrangement wird ein sog. Multi-Kamera-Video generiert. Am Beispiel der Spielpraxis des *Doppelsprungs* aus dem Action-Rollenspiel *Cyberpunk 2077* wird mithilfe der autoethnografischen Perspektive die Synchronisation zwischen Avatarkörper und Spieler:innen-Körper rekonstruiert.

Schlüsselwörter: digitale Spiele, Spielpraxis, Autoethnografie, Videographie, Methodologie

1. Einleitung

Der Artikel nimmt ein methodologisches Problem als Ausgangspunkt, um ein autovideographisches Forschungsdesign vorzustellen, das sich videographischer und autoethnografischer Methoden bedient, um eben jenes Problem methodologisch in den Griff zu kriegen. Nach einem kurzen Einblick in den Forschungsstand wird das methodologische Problem digitaler Spielpraktiken, d. h. die »Aufhebung der Differenz zwischen Innen- und Außenperspektive« (Venus 2012, S. 106) vorgestellt. Im Rahmen digitaler Spielpraktiken, so das Argument, kann

1 Danksagung: Gefördert wird das Projekt durch die Deutsche Forschungsgemeinschaft (DFG) – SFB 1482/1 – 2021 – 442261292.

die an die Spielwelt ausgerichtete Praxis einerseits beobachtet werden, während andererseits zugleich körperliche Inputs, z. B. das Drücken von Tasten, für den Vollzug notwendig sind. Digitale Spielpraktiken ermöglichen also die Beobachtung der vollzogenen Praxis und fordern somit gängige soziologische Theoriekonzepte heraus, innerhalb derer die Beobachtung der Praxis immer eine reflexive Zuwendung und somit eine Distanzierung zum Vollzug der Praxis voraussetzen (Schütz 1932; vgl. Schmidl 2022, S. 60).

Im Artikel wird am Beispiel der Spielpraxis des sog. Doppelsprungs im Spiel Cyberpunk 2077 gezeigt, wie durch den autoethnografischen Anteil des Forschungsdesigns Prozesse der Synchronisation zwischen Spieler:innen- und Avatarkörper rekonstruiert werden können. First-Person-Spielen im Allgemeinen liegt ein bestimmter Erfahrungsmodus zugrunde, der als vektionale Raumerfahrung beschrieben wird. Diese Spielerfahrung wird durch das Bildschirmgeschehen rekonstruiert und dann um autoethnografische Ausführungen erweitert, um zu zeigen, dass dem Doppelsprung als spezifische Spielpraxis eine räumliche und zeitliche Synchronisation von Avatar- und Spieler:innenkörper zugrunde liegt, die in seiner Durchführung zugleich einen dramaturgischen Höhepunkt erfährt. Der autoethnografische Erhebungskontext wird somit direkt mit der sequenziellen Analyse des Videos verschränkt und erlaubt so die Dekonstruktion der Hybridität zwischen Spieler:innen und Spielwelt (vgl. Wilde 2023).

2. Digitales Spielen als Gegenstand sozial- und kulturwissenschaftlicher Forschung

Seit einigen Jahrzehnten rücken auch digitale Spiele zunehmend in den Mittelpunkt der Sozial- und Kulturwissenschaften. Dass es sich bei digitalen Spielen schon lange nicht mehr nur um ein Nischenphänomen, sondern um eine der größten digitalen Branchen handelt, ist längst unumstritten (Statista 2024). Digitale Spiele sind Teil des gesellschaftlichen Alltags geworden.

Zu den frühesten Auseinandersetzungen mit Spielen im Allgemeinen zählen die beiden bereits zu Klassikern avancierten Werke »Homo Ludens« (Huizinga [1987] 2019) und »Die Spiele und die Menschen« (Caillois [1982] 2017) von Johan Huizinga und Roger Caillois. Während Huizingas weitestgehend historisch ausgerichtete Studie vor allem die kulturelle Relevanz des Spiels, d. h. die Verwobenheit zwischen gesellschaftlichen Institutionen (u. a. Krieg und Gerichte) und Spiel rekonstruiert und das spielerische des Menschen (»Homo Ludens«) zentral setzt, arbeitet Caillois neben den vier Spielformen (agôn, alea, mimicry, ilinx) und den zwei Spielweisen (paidia und ludus) u. a. am Beispiel des Glücksspiels heraus, dass die soziale Funktion des Spiels sich ändert, das Glücksspiel bspw. nicht mehr hilft, den Rätseln des Universums auf die Spur zu kommen, sondern weitestgehend zum Amüsement geworden ist.

In seiner bekannten Studie »Deep Play« (Geertz [1987] 2015, S. 202 ff.) zeigt Richard Geertz auf, wie im balinesischen Hahnenkampf sowohl soziale Beziehungen und Netzwerke als auch Machtstrukturen und Männlichkeit als dramatische Rituale vollzogen werden und damit auch, wie tief verwurzelt kulturelle Praktiken sein können. Neben der Forschung zu Spielen stellt sich derzeit auch die Frage nach der Digitalität von Spielen. So liefert bspw. Britta Neitzel eine kultursoziologische Erklärung zu digitalen Spielen, indem sie darauf verweist, dass durch, mit und in digitalen Spielen Konflikte unterschiedlicher Art bearbeitet werden (Neitzel 2008; Compagna/Derpmann 2018, S. 14), digitalen Spielen also die »Möglichkeit zur Reaktion auf die neuartige, technische Objektivität« (ebd.) zugrunde liegt. Neben soziologischen Annäherungen zu digitalen Spielen finden sich vor allem Studien, die bestimmte Aspekte digitalen Spielens in den Blick nehmen, seien es neue soziotechnische VR Spielpraktiken (Krell/Wettmann 2023), die Interaktion mit »Computergesteuerten Spielpartnern« (Harth 2014) oder mediatisierte Interaktionsordnungen beim Spielen von Pokémon GO (Woznica 2020) im Speziellen, die Untersuchung von Game Audiences (Gosling/Crawford 2011; Nash 2015) oder Game-Livestreaming via Twitch (Ochsner/Spöhrer/Waldrich 2020; Taylor 2018) im Allgemeinen. Zudem finden sich Konzepte von bspw. Play und Game auch in Sozialtheorien im Allgemeinen (vgl. Mead 1934). Welchen gesellschaftlichen Stellenwert digitale Spiele darüber hinaus haben, wird auch durch Begriffe wie Gamification und Serious Games deutlich, die auf die zunehmende Relevanz von Spielprinzipien in verschiedenen gesellschaftlichen Bereichen, bspw. Politik und Wirtschaft (Gamification) oder Kultur und Bildung (Serious Games) hinweisen (vgl. Stampfl 2012; Dicheva et al. 2015).

3. Die Aufhebung der Differenz zwischen Innen- und Außenperspektive als methodologisches Problem

Ausgangspunkt für den Beitrag ist ein methodologisches Problem, das sich aus den Spezifika digitalen Spielens ergibt: die Aufhebung der Differenz zwischen Innen- und Außenperspektive (Venus 2012). Im Folgenden wird es darum gehen, dieses Problem zunächst genauer zu erläutern, um anschließend darzulegen, inwiefern das autovideographische Forschungsdesign das Problem zu lösen versucht. Dabei soll vor allem die Verschränkung zwischen dem spezifischen Erhebungskontext (Autoethnographie) und dem Analysekontext konkreter herausgearbeitet werden.

Digitale Spielpraktiken lassen sich insofern immer auch als medial vermittelte Praktiken beschreiben, da ihnen Steuerungs- und Medienoperationen zugrunde liegen (vgl. Schemer-Reinhard 2018, S. 161): Bei der klassischen Computeranwendung außerhalb des digitalen Spielens wird i. d. R. mit Maus und Tastatur gearbeitet. Dabei werden bestimmte Inputs erzeugt, indem Körper(teile) bewegt,

Tasten und/oder Mauszeiger berührt werden. Erfolgreich ist ein solcher Input dann, wenn der (erwartete) Output beinahe simultan auf dem Bildschirm dargestellt wird. So resultiert bspw. das Drücken der Z-Taste in die Darstellung des Buchstabens Z, wenn in einem Textprogramm (bspw. Word) gearbeitet wird, der Doppelklick eines Ordners führt dazu, dass eben dieser Ordner geöffnet und Dateien angezeigt werden und während der linke Mauszeiger gedrückt gehalten wird, können Dateien verschoben werden. Der Computeranwendung liegt primär das Paradigma der »direkten Manipulation« (Shneiderman 1983) zugrunde.

Schemer-Reinhard spricht von einer »Doppelnatur des Interfaces als Werkzeug zur Steuerung und als Medium für Feedback« (Schemer-Reinhard 2018, S. 156): Mit dem Interface als Werkzeug kann der Mauszeiger bewegt werden, das Feedback wird auf dem Bildschirm im Sinne eines sich bewegenden Mauszeigers medial dargestellt. Entscheidend ist hierbei, dass »alles, was auf solch einem Bildschirm zur Darstellung kommt, […] trotz aller Virtualität doch eindeutig den Sinnzusammenhängen der realen Welt des Benutzers zugeordnet [ist]« (Schemer-Reinhard, S. 160). Die Spezifika des Computerspiels sieht Schemer-Reinhard im Unterschied zur klassischen Computernutzung in der auf die Spielwelt ausgerichteten praktischen Orientierung:

> »[Das Computerspiel] entwirft mit oft hohem Aufwand eine fiktionale Alternativwelt, innerhalb derer und in Bezug auf die gehandelt werden kann. Handlungen (vor allem Tastendrücken) des Spielers führen also auf der quasi-direkten Ebene des Interfaces zu unmittelbaren Effekten, die als virtuelle Handlungen wahrgenommen werden (Laufen, Springen, Ducken, Schießen …), das heißt sie beziehen sich auf der zeichenbeziehungsweise metaphernbasierten Ebene auf symbolische Entitäten, die Elemente einer Fiktion sind (Raumschiff, Auto, Fadenkreuz, Gangster, Soldat, Archäologin, Klempner …)« (ebd.).

Was damit einhergeht, beschreibt Jochen Venus als Aufhebung der »Differenz zwischen der Innen- und Außenperspektive des Handelns« (Venus 2012, S. 106) insofern beim »Computerspiel die Darstellung und [der] Nachvollzug eines Handlungstyps [ermöglicht wird], der zwischen direkter Tätigkeit, wie sie vom handelnden selbst vollzogen und erlebt wird, und der Tätigkeit einer anderen Person [steht], die von außen beobachtet werden kann« (ebd.).

Beide Autoren verweisen darauf, dass ähnlich wie bei der klassischen Computernutzung Inputs (direkte Tätigkeiten) in Outputs resultieren, diese beim Computerspiel allerdings als beobachtbare Tätigkeiten dargestellt werden und sich anders als in der klassischen Computernutzung primär an der Spielwelt ausrichten: Beim digitalen Spiel fungiert die Spielwelt als primärer Rahmen (vgl. Goffman [1974] 1980; vgl. Fine 2022), der permanent aufrechterhalten werden muss. Aus dem Drücken einer Taste kann so ein Springen oder Schießen, ein Bewegungsablauf oder eine andere spielspezifische Spielpraxis (Werfen, Tragen, Halten usw.) werden. In

seiner beeindruckenden Studie zeigt Rune Klevjer aus phänomenologischer Perspektive auf, wie durch prosthetische Avatare der visuelle Wahrnehmungsapparat in den »Space Screen« (Klevjer 2002, S. 13) verschoben, d. h. als Teil der Spielwelt konstituiert wird. Klevjer spricht hier von einer Rekonfiguration des Körpers als Teil der Spielwelt (vgl. Klevjer 2012). Die »Aufhebung der Differenz zwischen Innen- und Außenperspektive des Handelns« (Venus 2012, S. 106) ist also zum einen bedingt durch die Rekonfiguration des Körpers als Teil der Spielwelt (vgl. Klevjer 2012), zum anderen durch die Spielwelt, die zugleich das praktische Handeln orientiert (vgl. Schemer-Reinhard 2018). Die Grenze zwischen der beobachtbaren Tätigkeit und der von Spieler:innen ausgeführten direkten Tätigkeit verschwimmt nach Klevjer insofern, als der Wahrnehmungsapparat im Sinne eines »embodied make-believe« (Klevjer 2012, S. 23) während der Spielpraxis in der Spielwelt verortet wird. Das schließt unmittelbar an Venus an, der von der »Darstellung erlebten Handelns« (Venus 2012, S. 106) als Spezifika des digitalen Spiels spricht und dabei beispielhaft auf typische ›Ich-Beschreibungen‹ von Spieler:innen verweist, die während des Spielvollzugs üblicherweise gemacht werden (vgl. Venus 2012, S. 106 ff.). So werden Spielbeschreibungen üblicherweise aus der Ich-Perspektive formuliert: »Ich renne jetzt um die Ecke«, »Ich muss jetzt mit dem Charakter sprechen«, »Ich kaufe gerade eine Waffe« oder »Ich baue gerade Upgrades ein« und verweisen damit auf eine Relokalisierung des Selbst als Teil der Spielwelt.

Digitale Spielpraxen setzen außerdem einerseits menschliche Körper und technische Artefakte sowie deren funktionierende Verschaltung voraus und ermöglichen andererseits – im Modus des »embodied make-believe« (Klevjer 2012, S. 23) – Zuwendungen zu Erlebnissen (der Spielwelt). So können Spielpraktiken beobachtet werden, während sie vollzogen werden. Damit irritiert die Praxis digitalen Spielens den soziologischen Konsens, dass »die bewusste Zuwendung zu den Erlebnissen nie während des Erlebens selbst stattfinden kann, sondern immer nur durch einen rückschauenden Blick auf abgeschlossene Erlebnisse« (Schmidl 2022, S. 60, ursprünglich Schütz 1932, S. 55). Spieler:innen können während des Vollzugs die Spielpraxis beobachten, da die Spielpraxis in der Spielwelt lokalisiert ist, während die Spielwelt als primärer Rahmen aufrechterhalten wird (und vice versa). Die bereits genannte Aufhebung der »Differenz zwischen Innen- und Außenperspektive« (Venus 2012, S. 106) kann so, vor dem Hintergrund der Rekonfiguration des Körpers als Teil der Spielwelt, als Aufhebung der Differenz zwischen Beobachter:innen- und Teilnehmer:innenperspektive verstanden werden.

Digitales Spielen adäquat zu beschreiben, muss also einerseits im Sinne einer »Zoom-Out-Perspektive« (Alkemeyer 2017, S. 47 ff.) das Gefüge, in das Körper und Artefakte eingebettet werden, ernst nehmen, um so die spezifischen Arten und Weisen der Koordination von Wahrnehmung (vgl. Husel 2014, S. 48 ff.), seien es Spielperspektiven (z. B. First oder Third-Person-Perspektiven) oder Interfaceschnittstellen (z. B. Zeigefinger der rechten Hand – linke Maustaste), präzise zu beschreiben. Andererseits bedarf es einer »Zoom-In-Perspektive«

(Alkemeyer 2017, S. 48 ff.), um die praktische Arbeit des Arrangements, d. h. des Körpers, der Spieler:innen und der Technik zur Aufrechterhaltung des Geschehens sowie des Umgangs mit Problemen und Irritationen nachzuvollziehen.

4. Vorstellung des autovideographischen Forschungsdesigns

Der autoethnografische Anteil des Forschungsdesigns orientiert sich an dieser Zoom-In-Perspektive. Es zeichnet sich im Allgemeinen durch eine Kombination aus Autoethnografie und Videographie aus. Mithilfe zweier Kameras werde ich, der Spieler, wird mein Körper, die Tastatur und Maus, mein Schreibtisch, Bildschirme und viele mehr aus zwei Perspektiven aufgenommen: Zum einen aus der Vogelperspektive, zum anderen aus der Frontalperspektive. Zugleich wird das Bildschirmgeschehen aufgezeichnet. Mithilfe des Softwareprogramms Open-Broadcast-Software (OBS) wird ein Video generiert, das mich, den Spieler, meinen Körper, das Bildschirmgeschehen sowie meine Körperumgebung beim Spielen aufnimmt. Resultat ist ein sog. Multicam-Video, d. h. ein Video in dem ich, der Spieler, 1) aus der Frontalperspektive sowie 2) aus der Draufsicht-Perspektive zu sehen bin, ebenso wie 3) das Bildschirmgeschehen (vgl. Knoblauch/Wilke i. d. B.).

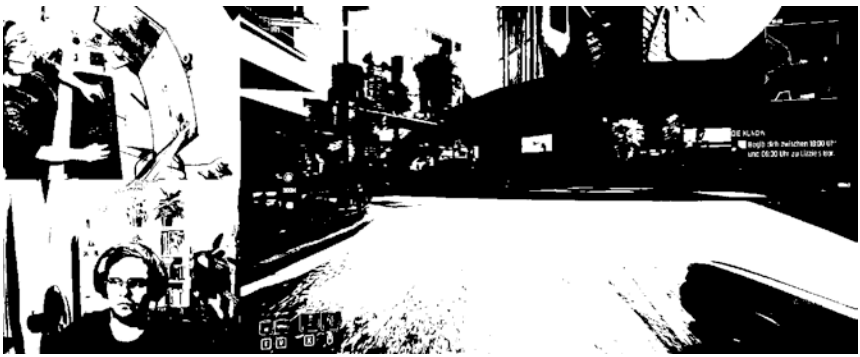


Abbildung 1: Screenshot des MultiCam-Videos

Dieses Multicam-Video dient dann als primäre Datengrundlage und wird um autoethnografische Notizen und Memos ergänzt. Während des Spielens wird das Spielgeschehen im Sinne der Thinking-Aloud-Methode (vgl. Eccles/Arsal 2017; vgl. Hoonhout 2022) nicht selten kommentiert, um so die »Schweigsamkeit des Sozialen« (Hirschauer 2001) ebenso wie die Flüchtigkeit kürzerer Spielzüge, Tasten- oder Mausinputs zu dokumentieren. Das Multicam-Video erlaubt also insofern die Befremdung der eigenen Kultur (Hirschauer/Amann 1997), als es den Körper des Spielers, meinen Körper (und seine Umgebung) zum Gegenstand der Analyse macht und dabei als Teil der Spielpraxis ernstnimmt. Vor allem hier setzt sich das Forschungsdesign von gängigen Analysen, vor allem im Rahmen der

Game Studies ab, da die Analyse nicht nur bei der Betrachtung des Bildschirmgeschehens, gegebenenfalls bei der Reflexion von Spielerfahrungen endet, sondern das materielle Arrangement des Spielgeschehens, das also, was ›vor‹ dem Bildschirm passiert, als Teil der Spielpraxis ernstnimmt. Während der Zoom-Out also stärker die Verschränkung zwischen materiellem Arrangement und Praxis in den Blick nimmt, erlaubt der autoethnografische Anteil des Forschungsdesigns ein Zoom-In, indem über subjektive Erfahrungen des Forschenden, über mich, der Versuch unternommen wird, »kulturelle Erfahrungen des Feldes [dem ich angehöre] zu verstehen« (Boll 2019, S. 33).

Ich halte an dieser Stelle fest: Der videographische Teil des Forschungsdesigns fokussiert das materielle Arrangement, die physischen und virtuellen Artefakte und Körper. Er erweitert die Außenperspektive, d. h. die Beobachter:innenperspektive der Spieler:innen um einen Zoom-Out und erlaubt es, den Körper des Spielers, meinen Körper, neben anderen Artefakten, zum Objekt der eigenen Anschauung zu machen. Der autoethnografische Teil des Forschungsdesigns fokussiert stärker die Innenseite und macht mich zum Forschungsobjekt, Erkenntnisinstrument und Wissensspeicher, indem auf Gedanken, Erlebnisse und Wahrnehmungen, ebenso wie kulturelles bzw. implizites Feldwissen zurückgegriffen und dieses so rekonstruiert werden kann (vgl. ebd.). Indem ich mich selbst zum Objekt der eigenen Anschauung mache, diese Anschauung zugleich um die autoethnografische Reflexion der ›Innenseite‹ ergänze, erlaubt das Forschungsdesign der Leiblichkeit des Spielgeschehens (vgl. Klevjer 2022) nachzugehen. Bspw. können so »Augen-Blicke, die für das bloße Auge unsichtbar sind« (Boll 2019, S. 33; vgl. Honer 2011), nachvollzogen werden.

Nach der Generierung des Multicam-Videos, wird das Video als Datengrundlage verwendet. Vor dem Hintergrund der Forschungsfrage wird das Videomaterial codiert, um zugleich relevante Sinneinheiten herauszuarbeiten, die in einem nächsten Schritt sequenzanalytisch untersucht werden. Entscheidend für den Artikel ist die nun kommende Übersetzungsleistung: Obwohl das Multicam-Video es erlaubt, die Temporalität sowohl des Bildschirmgeschehens als auch des darüber hinausgehenden materiellen und körperlichen Geschehens simultan, d. h. z. B. die Simultanität zwischen Inputs und Outputs, zu bewahren und zu verändern (Bergmann 1985; vgl. auch Hirschauer 2001, S. 441), ist eine »schriftliche Entfaltung von Ereignissen« (Hirschauer 2001, S. 441), wie sie das (auto-)ethnografische Schreiben ermöglicht, notwendig, um die Aufrechterhaltung der Differenz zwischen Innen- und Außenperspektive in ihrer (zeitlichen) Prozessualität nachvollziehbar zu machen. Diese »Dehnung der Zeit« (ebd.) leistet die Erstellung eines Videoverlaufsprotokolls, indem mikrosoziologisch beschrieben wird, wie Körper(teile) bewegt, Wahrnehmungen durch die materielle Infrastruktur koordiniert und Irritationen durch die Spielwelt erzeugt werden. Autoethnografische Ergänzungen werden in den Text kenntlich eingebaut, um so die Sequentialität des Geschehens auch um nicht beobachtbare Praktiken des Überlegens,

Lesens, Beobachtens (des Spielweltgeschehens) und Zuhörens zu verdichten. Nachdem das Videoverlaufsprotokoll verfasst wurde, wird es zugleich sequenzanalytisch untersucht, d. h. die Verknüpfung einzelner Sinneinheiten – (1) »Ich drücke« (2) »die Taste F« (3) »um zu springen« – wird in ihrer Vollzugswirklichkeit dekonstruiert, um so die relevanten Schnittstellen- und Körperkonfigurationen herauszuarbeiten und die Gelingensbedingungen für die Rekonfiguration des Körpers als Teil der Spielwelt adäquat zu beschreiben. Wie hierbei der autoethnografische Analysekontext und der Erhebungskontext verschränkt sind, soll beispielhaft an einer Spielpraktik verdeutlicht werden: Dem sog. Doppelsprung im Action-Roleplay-Game Cyberpunk 2077.

Das Fallbeispiel: Der Doppelsprung in Cyberpunk 2077

Bei Cyberpunk 2077 handelt es sich um ein digitales Spiel, welches im Jahr 2020 vom polnischen Entwicklerstudio CD Projekt Red veröffentlicht wurde. Das Spiel wird üblicherweise dem Genre des Action-Rollenspiels zugeordnet. In Cyberpunk 2077 steuern Spieler:innen einen Avatar, genannt V, der zugleich der Hauptcharakter einer Geschichte ist. V lebt in der fiktiven Stadt Night City, die an eine retrofuturistische Sci-Fi Metropole erinnert. Spieler:innen erleben die Spielwelt aus der sog. First-Person-Perspektive, d. h. sie schauen durch die Augen des Avatars. In First-Person-Spielen ist der Avatarkörper meist nicht beobachtbar für die Spieler:innen. Im Rahmen bestimmter Spielpraxen werden Teile des Avatarkörpers ab und zu in den Bildschirm eingeblendet, so bspw. wenn das Lauftempo (des Avatarkörpers) durch das Drücken einer Taste erhöht wird und aus dem Laufen ein Rennen wird. Auf dem Bildschirm wird dann ein Teil der Bewegung der Unterarme dargestellt, wodurch Teile des unsichtbaren Avatarkörpers beim Rennen beobachtet werden können. Eines der bekanntesten Beispiele für die Einblendung von Körperteilen im Kontext von First-Person-Spielen ist sicherlich die Einblendung von zwei Händen, die eine ›nach vorne ausgerichtete‹ Waffe halten, im Kontext von Ego-Shooter-Spielen.

Spieler:innen wird eine Welt aus dem Jahr 2077 präsentiert, in der Menschen ihre Körper auf ganz unterschiedliche Weisen modifizieren, bspw. mithilfe von Chips, Implantaten und/oder Prothesen. Im Laufe des Spiels können bestimmte Fähigkeiten so zum einen durch das Sammeln von Erfahrungspunkten, zum anderen durch den Einbau und Kauf von Technologien (in den Avatarkörper) freigeschaltet und verbessert werden. Im Folgenden soll es um eine Prothese gehen, die im Spiel als »Verstärkte Sehnen« beschrieben wird, und die neben anderen Vorteilen auch die Durchführung eines Doppelsprungs ermöglicht. »Drücke in der Luft [Symbol für Leertaste, Anmerkung M. T. W.], um einen *Doppelsprung* [Herv. i. O.] durchzuführen« lautet die entsprechende Beschreibung. Es handelt sich dabei um eine Prothese, die man bei einem sog. Ripperdoc erwerben kann,

also jenen Ärzten in der Welt von Cyberpunk 2077, die auf die technische Bearbeitung des Körpers, sein Enhancement, spezialisiert sind. Entscheidend für den Beitrag ist weniger diese narrative Hinführung zum Doppelsprung als vielmehr die Spielpraxis des Doppelsprungs.

Vektionale Raumerfahrung in digitalen First-Person-Spielen

Was bisher geschah: Während des Spiels habe ich mir die entsprechende Prothese gekauft und in den Avatarkörper einbauen lassen. Nachdem die Prothese eingebaut wurde, wollte ich testen, wie genau der Doppelsprung funktioniert. Die nun folgende Beschreibung des Bildschirmgeschehens soll als Ausgangspunkt dienen. Es handelt sich dabei um eine autoethnografische Videoverlaufsbeschreibung, bei der größere autoethnografische Anteile mit eckigen Klammern markiert sind.

Autoethnografische Videoverlaufsbeschreibung – der Doppelsprung: Auf dem Bildschirm wird ein Teil der Spielwelt in Form einer kleinen Straße und einer Mauer dargestellt. Die Mauer wird immer größer und nimmt irgendwann einen Großteil des Bildschirms ein. [Ich bewege meinen Avatar also auf die Wand zu, indem ich entsprechende Steuerungsinputs gebe.] Während sie größer wird, ist mal mehr vom oberen, mal mehr vom unteren Teil der Mauer zu sehen. [Vom oberen Teil der Mauer ist mehr zu sehen, wenn ich den Avatar steuere und den Input für Springen eingebe, d. h. die Leertaste drücke. Beim ›Landen‹ auf den Boden ist dann wieder der untere Teil der Wand zu sehen.] Durch das Drücken der Leertaste wird ein Sprung im dreidimensionalen Raum dargestellt, indem 1) Teile des Bodens am unteren Ende des Bildschirms ›verschwinden‹ und 2) das was vorher nicht zu sehen war, d. h. ›oberhalb‹ der Sichtreichweite des Avatars lag, nun von oben in den Bildschirm ›auftaucht‹. Diese Darstellung von Bewegung ist vergleichbar mit dem Blick in den Himmel: Wenn wir in den Himmel nach oben schauen und somit alles im unteren Ende unseres Sichtfeldes nach und nach verschwindet, kommt zugleich – durch die Bewegung – neues ›von oben‹ ins Sichtfeld. Während die Bewegung des Raumes in eine Richtung dargestellt wird, z. B. nach unten, wenn hochgesprungen wird, drücke ich erneut die Leertaste. Diesmal wird die Raumbewegung wenige Sekunden länger dargestellt. Der obere Teil der Mauer ragt nur bis zur Mitte des Bildschirms. [Damit wird gezeigt, dass der Sprung nun ›höher‹ ging, als der vorherige. Ich kann ›über die Mauer hinaussehen‹.] Der Doppelsprung hat funktioniert.

Dieser kurze Einblick verweist zunächst auf die Differenz zwischen kontrollierter Beschreibung und Feldbeschreibung, da das Bildschirmgeschehen zwar zum einen kontrolliert beschrieben, zum anderen auch um autoethnografische Beschreibungen erweitert wurde. Bei digitalen First-Person-Spielen wie Cyberpunk 2077 sehen Spieler:innen auf dem Bildschirm einen Raum (der Spielwelt), durch dessen Darstellung ein Effekt von Bewegung (trotz Stillstand) erzeugt wird. Der Begriff der Vektion beschreibt üblicherweise jene Wahrnehmung von

Bewegung, die Menschen bspw. dann haben, wenn sie in einem stehenden Zug sitzen, aus dem Fenster schauen und einen Zug sehen, der losfährt, sodass sie selbst (i. d. R. für einen kurzen Moment) den Eindruck bekommen, der Zug in dem sie sitzen, würde losfahren bzw. sich bewegen. First-Person-Spiele operieren entlang dieser Logik der Vektion, koppeln diese aber explizit, d. h. im Rahmen des beobachtbaren Bildschirmgeschehens, an die Darstellung von Raum. Die Rekonfiguration des Körpers als Teil der Spielwelt setzt also zugleich vektionale Raumerfahrungen voraus. Sie kann so als primäre Erfahrungsweise im Rahmen digitaler First-Person-Spiele verstanden werden, zu deren Aufrechterhaltung – und damit auch der Aufhebung der Differenz zwischen Innen- und Außenperspektive – bestimmte Schnittstellenkonfigurationen zwischen Körper/-teilen und materieller Infrastruktur beitragen, während sie die Körper der Spielenden entsprechend rekrutieren (Hirschauer 2004). Diese Rekrutierung läuft entlang von Schnittstellen und damit einhergehenden Konfigurationen zwischen Körper-Interface-Spielwelt, sodass bspw. aus der Bewegung der rechten Hand, mit der üblicherweise die Maus gesteuert wird, die Blickrichtung im Raum der Spielwelt gesteuert wird. Vereinfacht gesagt, wird aus der Bewegung der rechten Hand des Spielenden, meines Körpers, die Bewegung des Avatarkörpers, wodurch im Raum der Spielwelt in eine bestimmte Richtung geblickt wird. Die Steuerung des Avatars setzt dabei immer eine spezifische Anordnung, eine Ausrichtung des Spielenden, meines Körpers voraus. Es handelt sich dabei also um eine Spielpraxis, die verschiedene Teile des Arrangements auf je entsprechende Art und Weise rekrutiert, bspw. wird durch das Drücken der W-Taste mit dem Mittelfinger der linken Hand der Avatar üblicherweise ›nach vorne‹ gesteuert. Die Tastensteuerung definiert zum einen, wie Finger angeordnet werden, greift damit zugleich auf spezifisches kulturelles Wissen im Umgang mit Tastaturen und spezifischen Spielgenres zurück.

An dieser Stelle ist festzuhalten, dass der Doppelsprung im Modus der vektionalen Raumerfahrung vollzogen wird. Durch den autoethnografischen Erhebungskontext wird auf die spezifische Arbeit, auf das »Practical Accomplishment« (Reeves/Greiffenhagen/Laurier 2017) aufmerksam gemacht, das implizit oder explizit den Vollzug des Doppelsprungs als Spielpraxis ermöglicht und nicht etwa eine ›Bewegung auf dem Bildschirm‹ bleibt. Auch Goffmans Begriff der Modulation beschreibt diese Transformation einer Handlung (z. B. des Drückens einer Taste) in eine andere (z. B. des vorwärts Laufens) sehr präzise (vgl. Goffman [1974] 1980, S. 52 ff.) und verweist zudem auf jene Auseinandersetzung im Rahmen der Game Studies, die dem Spiel im Generellen immer auch die »Suspension of Disbelief« (vgl. Brown 2012) zuschreiben. Durch den autoethnografischen Erhebungskontext wird die Beschreibung des Bildschirmgeschehens um Feldbeschreibungen erweitert und reflektiert, sodass die vektionale Raumerfahrung als spezifische, d. h. für First-Person-Spiele primäre Erfahrungsweise herausgearbeitet werden konnte, die mit der »Rekonfiguration des Körpers als Teil der Spielwelt« einhergeht.

Im nächsten Schritt soll am Beispiel des Doppelsprungs konkret gezeigt werden, wie Spieler:innen durch die Synchronisation des visuellen Wahrnehmungsapparates zwischen Spieler:innen- und Avatarkörper zum einen entlang der Zeitdimension ein Gefühl für Zeit entwickeln und zum anderen entlang der Raumdimension ein Gefühl für Raum entwickeln – das Timing des zweiten Sprungs, d. h. des Drückens der Leertaste ist entscheidend für das Funktionieren des Doppelsprungs ebenso wie das Erfüllen der Voraussetzung, dabei ›in der Luft‹ zu sein.

Raumzeitliche Synchronisation von Körpern

Der Doppelsprung erlaubt mir einen »Sprung in der Luft« durchzuführen, d. h. ich bzw. Spieler:innen im Generellen erkennen, wann sich der Avatar ›in der Luft‹ befindet, obwohl eigentlich ›nur‹ die Bewegung von Raum auf dem Bildschirm dargestellt wird. Das erneute Drücken der Leertaste erlaubt es dann, höher zu springen als es vorher der Fall war und so gegebenenfalls Hindernisse wie bspw. hohe Mauern zu überwinden.

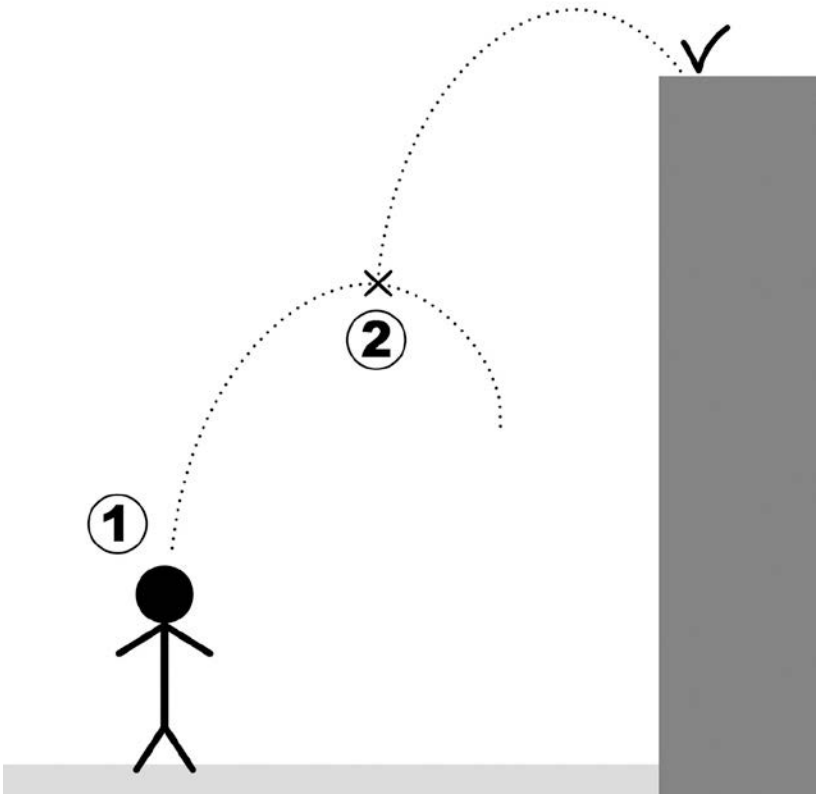


Abbildung 2: Die grafische Darstellung eines erfolgreichen Doppelsprungs

Abbildung 2 zeigt, wie der Doppelsprung funktionieren kann. Die »1« steht hier für das erste Drücken der Leertaste, um zu »springen«. Die »2« steht für das zweite Drücken der Leertaste, der Durchführung des »Doppelsprungs«. Dabei wird, so das zentrale Argument, der visuelle Wahrnehmungsapparat zwischen Spieler:innenkörper und (unsichtbarem) Avatarkörper synchronisiert (vgl. Krell/Wettmann 2023). Es ist hier vor allem das implizite Wissen über einen Avatarkörper, das eine solche Synchronisierung ermöglicht und zur Relokalisierung in der Spielwelt führt. Obwohl der Avatarkörper nicht immer zu sehen ist, ist doch klar – vor allem durch seine narrative Einbettung – dass es ihn nicht nur als virtuellen und meist unbeobachtbaren Körper gibt, sondern es darüber hinaus vor allem seine Augen sind, durch die die Spielwelt gesehen wird. Der Bildschirm fungiert hier als Interface zwischen zwei visuellen Wahrnehmungsapparaten, einem körperlich-fleischlichen und einem virtuellen, und gleicht diese einander an, ohne sie komplett zu hybridisieren: Der Bildschirm wird nicht zu den Augen der Spieler:innen. Er gleicht also einerseits an, trennt andererseits durch die Beobachterposition der Spieler:innen. Genau in dieser doppelten Hinsicht erfüllt der Bildschirm seine Funktion als Interface (vgl. Lipp/Dickel 2022; Schemer-Reinhard 2018). Zugleich ist diese Angleichung immer auch imaginative Praxis (Herbrik 2013; Herbrik/Röhl 2007): Es wird sich im Modus des »embodied make believe« (Klevjer 2012, S. 23) auf die »Suspension of Disbelief« (Brown 2012) eingelassen. Damit diese Aufhebung der Differenz zwischen Innen- und Außenperspektive aufrechterhalten bleibt, bedarf es also impliziten Avatarkörperwissens und einer spezifischen Rahmung, d. h. der Rahmung des Spiels, mit dem immer auch ein »So-tun-als-ob« (engl.: make-believe) einhergeht (Goffman [1974] 1980, S. 60 f.).

Beim Doppelsprung wird diese Synchronisation in räumlicher und zeitlicher Hinsicht aufrechterhalten. In räumlicher Hinsicht geht mit der Rekonfiguration des Körpers als Teil der Spielwelt eine Wahrnehmung von Höhen und Tiefen (des Raumes der Spielwelt) einher, das sich aus der vektionalen Raumerfahrung ergibt: Beim Rennen oder Laufen bspw. bleibt der sog. Point of View, d. h. die Perspektive, die auf dem Bildschirm dargestellt wird, in etwa gleich. Werden Treppen hoch oder heruntergelaufen oder wird gesprungen oder sich geduckt, verändert sich dieser Point of View und man kann je nachdem mehr oder weniger (weit) sehen. Die Veränderung dieses Points of View, d. h. die Abweichung seiner Standardansicht durch Springen oder Ducken führt dazu, dass Spieler:innen Höhen- und Tiefenunterschiede wahrnehmen, die beim Doppelsprung noch einmal auf andere Weise herausgefordert werden. Denn das Funktionieren des Doppelsprungs setzt ein »in der Luft sein« voraus. Damit setzt es in der Zeitdimension zugleich ein Timing voraus: Spieler:innen müssen in der Lage dazu sein, die Leertaste ein zweites Mal genau in jenem Zeitraum zu drücken, in dem der Avatarkörper nicht den Boden berührt. Misslingt dies, steht der Avatarkörper auf dem Boden und aus dem Sprung in der Luft, der es ermöglicht größere Höhen

zu überwinden, wird ein einfacher Sprung. Eine gesteigerte Dramaturgie erfährt das Ganze, wenn bspw. auf einer Mauer gelandet werden soll, was ein spezifisches Timing voraussetzt, da erst das Drücken der Leertaste im höchsten Punkt es ermöglicht, die maximale Sprunghöhe herauszuholen. Kommt es zum Fehltiming, misslingt die Landung auf der Mauer.

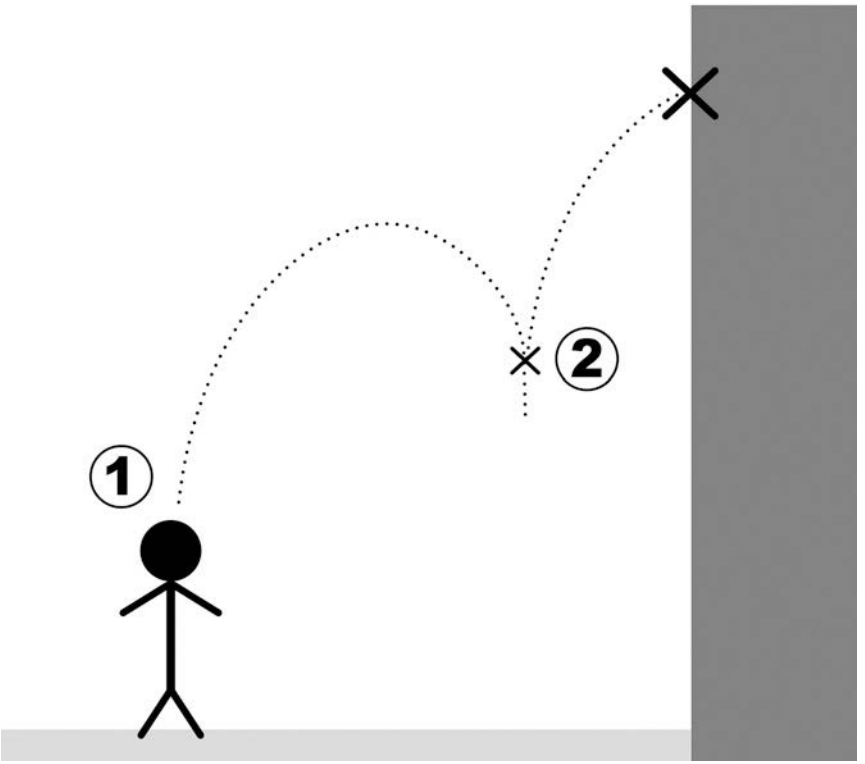


Abbildung 3: Die grafische Darstellung eines misslungenen Doppelsprungs auf eine Mauer

Vor dem Hintergrund der vektionalen Raumerfahrung setzt der Doppelsprung also zum einen die Synchronisation zwischen Spieler:innen- und Avatarkörper voraus, die durch Interfaces aufrechterhalten wird. Zum anderen erfordert die Spielpraxis selbst eine auf den Raum der Spielwelt ausgerichtete, d. h. auch körperlich verinnerlichte, Raum-Zeit-Wahrnehmung. Im Folgenden und letzten Abschnitt soll diskutiert und gezeigt werden, welchen Stellenwert die Verschränkung zwischen Analysekontext und Erhebungskontext (Autoethnografie) hat.

5. Fazit

Ausgangspunkt war das methodologische Problem der Aufhebung der Differenz zwischen Innen- und Außenperspektive. Die Hybridität des digitalen Spielgeschehens (vgl. Wilde 2023) kann mithilfe der Generierung und Analyse von Multi-Kamera-Videos aus einer Zoom-Out-Perspektive analysiert werden, während der autoethnografische Teil des Forschungsdesigns es erlaubt, das Spielgeschehen mithilfe des Feldwissens des spielenden Forschers, mir, nachvollziehbar zu machen und in seiner Prozessualität zu beschreiben. Der autoethnografische Erhebungskontext spielt also insofern eine Rolle in der Analyse, als die Videoverlaufsbeschreibungen um das ›Unbeobachtbare der Praxis‹ ergänzt werden, indem das Spielgeschehen durch die Beschreibung in seiner Zeitlichkeit entzerrt werden kann. Diese Entzerrung erlaubt es, die Relevanz verschiedener Interface-Konfigurationen hinsichtlich ihrer je spezifischen Koordination von Wahrnehmung zu rekonstruieren. Am Beispiel der First-Person-Spiels Cyberpunk 2077 konnte so gezeigt werden, dass mit der vektionalen Raumerfahrung auch ein spezifisches, implizites Avatarkörperwissen einhergeht, das zugleich narrativ eingebettet ist und im Modus des embodied make-believe aufrechterhalten wird. Die Koordination von Wahrnehmung wird dabei über Interfaces hergestellt, die Spieler:innenkörper und Spielwelt spezifisch miteinander konfigurieren: So wird der visuelle Wahrnehmungsapparat zwischen Spieler:innenkörper und Avatarkörper angeglichen, die avatariale Bewegungsfähigkeiten an die linke Hand (die Tasten WASD als Steuerungsinputs), Blickrichtungsbewegungen an die rechte Hand (die Maus als Steuerungsinput) gekoppelt. Beim Doppelsprung wird die Konfiguration insofern auf die Probe gestellt, als das Höhen- und Tiefenwahrnehmungen notwendig sind, um die Praxis erfolgreich zu vollziehen. Voraussetzung ist die räumliche Synchronisation von Körpern durch die oben beschriebene Konfiguration zum einen, das damit einhergehende und notwendige Timing zum anderen. Durch den autoethnografischen Teil des Forschungsdesigns ist es möglich, die zeitliche und räumliche Dimension beim Doppelsprung ebenso wie Synchronisationsprozesse zwischen Spieler:innenkörper und Avatarkörper nachvollziehbar zu beschreiben und mithilfe des Feldwissens das spezifische der Medienpraxis des Computerspiels ernst zu nehmen.

Literatur

Alkemeyer, Thomas (2017): Praktiken und Praxis: Gegenständlicher und gelebter Körper in den Vollzügen von Ordnungs- und Selbstbildung. In: Phänomenologische Forschungen, H. 2, S. 41–56.

Bergmann, Jörg ([1985] 2007): Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit – Aufzeichnungen als Daten der interpretativen Soziologie. In: Hausendorf, Heiko (Hrsg.): Gespräche als Prozess. Linguistische Aspekte der Zeitlichkeit verbaler Interaktion. Tübingen: Gunter Narr, S. 22–68.

Boll, Tobias (2019): Autopornografie: Eine Autoethnografie mediatisierter Körper. Berlin und Boston: De Gruyter Oldenbourg.
Brown, Douglas W. (2012): The Suspension of Disbelief in Videogames. Brunel: School of Arts PhD Theses.
Caillois, Roger ([1982] 2017): Die Spiele und die Menschen: Maske und Rausch. Berlin: Matthes & Seitz.
Compagna, Diego/Derpmann, Stefan (Hrsg.) (2018): Soziologische Perspektiven auf Digitale Spiele: Virtuelle Handlungsräume und neue Formen sozialer Wirklichkeit. Köln: Herbert von Halem.
Dicheva, Darina/Dichev, Christo/Agre, Gennady/Angelova, Galia (2015): Gamification in education: A systematic mapping study. In: Journal of Educational Technology & Society 18, H. 3, S. 75–88.
Eccles, David W./Arsal, Güler (2017): The think aloud method: What is it and how do I use it? In: Qualitative Research in Sport, Exercise and Health 9, H. 4, S. 514–531.
Fine, Gary A. (2002): Shared Fantasy: Role Playing Games as Social Worlds. Chicago: University of Chicago Press.
Geertz, Clifford ([1987] 2015): Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme. 13. Auflage. Frankfurt am Main: Suhrkamp.
Goffman, Erving (1980): Rahmen-Analyse: Ein Versuch über die Organisation von Alltagserfahrungen. Frankfurt am Main: Suhrkamp.
Gosling, Victoria K./Crawford, Garry (2011): Game Scenes: Theorizing Digital Game Audiences. In: Games and Culture 6, H. 2, S. 135–154. https://doi.org/10.1177/1555412010364979.
Harth, Jonathan (2014): Computergesteuerte Spielpartner. Wiesbaden: Springer.
Herbrik, Regine (2013): Das Imaginäre in der (Wissens-)Soziologie und seine kommunikative Konstruktion in der empirischen Praxis. In: Keller, Reiner/Reichertz, Jo/Knoblauch, Hubert (Hrsg.): Kommunikativer Konstruktivismus. Wiesbaden: Springer Fachmedien, S. 295–315.
Herbrik, Regine/Röhl, Tobias (2007): Visuelle Kommunikationsstrategien im Zusammenspiel: Gestik im Fantasy-Rollenspiel. In: Sozialer Sinn 8, H. 2, S. 237–266. https://doi.org/10.1515/sosi-2007-0205.
Hirschauer, Stefan (2001): Ethnografisches Schreiben und die Schweigsamkeit des Sozialen. In: Zeitschrift für Soziologie 30, H. 6, S. 429–451.
Hirschauer, Stefan (2004): Praktiken und ihre Körper. Über materielle Partizipanden des Tuns. In: Hörning, Karl/Reuter, Julia (Hrsg.): Doing Culture. Neue Positionen zum Verhältnis von Kultur und sozialer Praxis. Bielefeld: transcript, S. 73–91.
Hirschauer, Stefan/Amann, Klaus (Hrsg.) (1997): Die Befremdung der eigenen Kultur: Zur ethnographischen Herausforderung soziologischer Empirie. Frankfurt am Main: Suhrkamp.
Honer, Anne (2011): Kleine Leiblichkeiten: Erkundungen in Lebenswelten. Wiesbaden: VS Verlag für Sozialwissenschaften.
Hoonhout, Jettie (2022): Let the game tester do the talking: Think aloud and interviewing to learn about the game experience. In: Isbister, Katherine/Hodent, Celia (Hrsg.): Game Usability. Boca Raton: CRC Press, S. 115–125.
Huizinga, Johan/Flitner, Andreas ([1987] 2019): Homo ludens: Vom Ursprung der Kultur im Spiel. 26. Auflage. Reinbek bei Hamburg: Rowohlt.
Klevjer, Rune (2012): Enter the Avatar: The Phenomenology of Prosthetic Telepresence in Computer Games. In: Sageng, John R./Fossheim, Hallvard/Larsen, Tarjei M. (Hrsg.): The Philosophy of Computer Games. Dordrecht: Springer Netherlands, S. 17–38.
Klevjer, Rune (2022): What is the Avatar?: Fiction and Embodiment in Avatar-Based Singleplayer Computer Games. Bielefeld: transcript.
Krell, Felix/Wettmann, Nico (2023): Corporeal Interactions in VRChat: Situational Intensity and Body Synchronization. In: Symbolic Interaction 46, H. 2, S. 159–181. https://doi.org/10.1002/symb.629.
Lipp, Benjamin/Dickel, Sascha (2022): Interfacing the Human/Machine. In: Distinktion: Journal of Social Theory 24, H. 3, S. 425–443. https://doi.org/10.1080/1600910X.2021.2012709.
Mead, George H. ([1934] 2015): Mind, Self & Society. Chicago: University of Chicago Press.
Nash, Kate (2015): Simulation Games, Popular Factual Media and Civic Engagement: An Audience Study of Asylum Exit Australia. In: Media, Culture & Society, 37, H. 7, S. 959–971. https://doi.org/10.1177/0163443715584103.

Neitzel, Britta (2008): Das Computerspiel als Leitmedium des 21. Jahrhunderts. In: Kaminski, Winfred/Lorber, Martin (Hrsg.): Clash of Realities. Copaed, S. 61–75.
Ochsner, Beate/Spöhrer, Markus/Waldrich, Harald (2020): Twitching: Laborious play or: How gamestreaming changes modes of playing. Wiesbaden: Springer.
Reeves, Stuart/Greiffenhagen, Christian/Laurier, Eric (2017): Video Gaming as Practical Accomplishment: Ethnomethodology, Conversation Analysis, and Play. In: Topics in Cognitive Science 9, H. 2, S. 308–342. https://doi.org/10.1111/tops.12234.
Schemer-Reinhard, Timo (2018): Interface. In: Beil, Benjamin/Hensel, Thomas/Rauscher, Andreas (Hrsg.): Game Studies. Wiesbaden: Springer Fachmedien, S. 155–172. https://doi.org/10.1007/978-3-658-13498-3_8.
Schmidl, Alexander (2022): Relationen: Eine postphänomenologische Soziologie der Körper, Technologien und Wirklichkeiten. Weilerswist: Velbrück Wissenschaft. https://doi.org/10.5771/9783748936855.
Schütz, Alfred (1932): Der sinnhafte Aufbau der sozialen Welt. Wien: Springer.
Shneiderman, Ben (1983): Direct Manipulation: A Step Beyond Programming Languages. In: Computer 16, H. 8, S. 57–69.
Stampfl, Nora S. (2012): Die verspielte Gesellschaft. Gamification – oder Leben im Zeitalter des Computerspiels. Heidelberg: dpunkt.
Statista (2024): Digital Media – Video Games. www.statista.com/outlook/dmo/digital-media/video-games/worldwide (14.08.2024).
Taylor, T. L. (2018): Watch Me Play: Twitch and the Rise of Game Live Streaming. Princeton, Oxford: Princeton University Press.
Venus, Jochen (2012): Erlebtes Handeln in Computerspielen. In: GamesCoop (Hrsg.): Theorien des Computerspiels zur Einführung. Hamburg: Junius, S. 104–127.
Wilde, Poppy (2023): Posthuman Gaming: Avatars, Gamers, and Entangled Subjectivities. New York: Taylor & Francis.
Woznica, Marcel (2020): Videographie im Wandel: Zur Analyse des mediatisierten Alltagsraumes mittels mobiler Erhebungsinstrumente. In: Zeitschrift für Qualitative Forschung 21, H. 1, S. 53.

Nonverbale Interaktionsmuster in Mixed Reality

Videographische Analysen von Human-Agent-Interactions

Jonathan Harth

Zusammenfassung: Die jüngste Forschung zu Human-Agent-Interaktionen (HAI) zeigt eine verstärkte Fokussierung auf multimodale Interaktionen mit menschenähnlichen virtuellen Agenten. Während viele Studien sich auf die Wahrnehmung und Erfahrung der menschlichen Nutzer*innen konzentrieren, bleiben die komplexen, wechselseitigen Dynamiken der Interaktion oft unberücksichtigt. Dieser Beitrag zielt darauf ab, diese Forschungslücke zu schließen, indem er einen innovativen methodischen Ansatz präsentiert, der sowohl die Beziehungs- als auch die Inhaltsebene von Mensch-Agent-Interaktionen analysiert. Durch den Einsatz von Mixed-Reality-Darstellungen und gemischten Methoden ermöglicht dieser Ansatz eine detaillierte Erfassung und Analyse des koordinierten Verhaltens von menschlichen Nutzer*innen und virtuellen Agenten während ihrer Interaktion. Die Ergebnisse zeigen, dass diese Methode tiefe Einblicke in die Entwicklung (proto-)sozialer Beziehungsmuster zwischen Mensch und Agent ermöglicht und somit neue Perspektiven für die Gestaltung und Evaluation von HAI eröffnet.

Schlüsselwörter: Mixed Reality, Videographie, Human-Agent-Interaction, Konversationsanalyse, Ethnomethodologie

1. Einleitung

Im 21. Jahrhundert intensiviert sich die Verflechtung zwischen menschlichem Leben und künstlicher Intelligenz stetig. Eine besonders anschauliche Manifestation dieser Entwicklung zeigt sich im Aufstieg verkörperter Konversationsagenten (Embodied Conversational Agents, ECAs) – digitaler Entitäten, die konzipiert wurden, um menschliche Interaktionen zu simulieren (Cassel et al. 2000). Diese Agenten, die sich zu immer natürlicheren Schnittstellen der Kommunikation entwickeln, transformieren Anwendungsfelder insbesondere im geschäftlichen und administrativen Bereich (Mersch et al. 2023; Lugrin/Pelachaud/Traum 2021). ECAs versprechen natürliche multimodale Echtzeit-Gespräche, die

sowohl verbale wie auch nonverbale Kommunikationsformen integrieren (Lugrin/Pelachaud/Traum 2022).

Unternehmen wie Digital Human (2024) und SoulMachines (2024) nehmen eine Vorreiterrolle bei der Entwicklung digitaler Entitäten ein, die nicht nur zur Kommunikation fähig sind, sondern auch auf einer menschenähnlichen Ebene mit Nutzer*innen interagieren können. Während das Aufkommen großer Sprachmodelle wie GPT (Brown et al. 2020) signifikante Fortschritte in der verbalen Interaktion ermöglicht hat, bleibt die Komplexität der nonverbalen Kommunikation in Mensch-Agent-Interaktionen (Human-Agent Interactions, HAI) ein Bereich mit erheblichem Forschungspotenzial.

Nonverbale Aspekte von Kommunikation werden bereits seit längerem im Forschungsfeld der Mensch-Roboter-Interaktion (Human-Robot Interaction, HRI) untersucht und entwickelt (Bartneck et al. 2019). Hier ergeben sich ähnliche Herausforderungen im Bereich der HAI: Die Erfassung und Generierung nonverbaler Kommunikation eröffnet neue Perspektiven für die Konzeption, Entwicklung und Evaluation von Interaktionsprozessen zwischen Menschen und virtuellen Agenten (Deriu et al. 2021).

Zudem nimmt das Medium der virtuellen Realität (Virtual Reality, VR) eine zentrale, transformative Rolle in der Erforschung von ECA-Interaktionen ein und hat sich als unverzichtbares Instrument für diese Untersuchungen etabliert. VR ermöglicht Nutzer*innen direkte *face-to-face*-Interaktionen mit virtuellen Entitäten, was Bavelas et al. (1997) zufolge weiterhin als ›Goldstandard‹ der Interaktionsforschung gelten kann. In der VR treten Nutzer*innen nicht nur in Kontakt mit einem Agenten, sondern tauchen in dessen Welt ein und erleben Interaktionen, die der *face-to-face*-Kommunikation sehr nahekommen. Die immersive Eigenschaft der VR fördert nicht nur einen authentischeren kommunikativen Austausch, sondern eröffnet auch Möglichkeiten für eine detaillierte Analyse der subtilen gemeinsamen Verhaltensweisen und Dynamiken in HAI (Foster 2019).

Ungeachtet der Vorteile von VR sieht sich die sozialwissenschaftliche Forschung bei der Untersuchung von Mensch-Agent-Interaktionen weiterhin mit einigen Herausforderungen konfrontiert, insbesondere aufgrund der begrenzten Möglichkeiten zur externen Beobachtung dieser Interaktionen (Fitrianie et al. 2019). Vor diesem Hintergrund präsentiert dieser Artikel eine innovative Methodik der Mixed-Reality-Videographie (vgl. Knoblauch/Wilke i. d. B.), die darauf ausgerichtet ist, die kommunikativen Nuancen und Verhaltensdynamiken in Mensch-Agent-Interaktionen innerhalb eines Mixed-Reality-Kontextes zu erfassen und zu analysieren. Diese Methode ermöglicht die multiperspektivische Beobachtung und Dokumentation von Interaktionen und überwindet damit die Limitationen herkömmlicher VR-Studien, die auf die singuläre Perspektive des VR-Headset-Trägers beschränkt sind.

Basierend auf Daten aus einem laufenden Forschungsprojekt wird dieser Beitrag die methodischen Aspekte der Mixed-Reality-Videographie detailliert

erläutern und deren Materialien, Verfahren sowie spezifische Möglichkeiten vorstellen. Diese Präsentation beinhaltet eine exemplarische Analyse multimodaler Interaktionen, inklusive Transkriptionen und grundlegender Interpretationen.

2. Kommunikation mit verkörperten Konversationsagenten

Dieser Abschnitt beleuchtet drei zentrale Aspekte: erstens, die technologischen Fortschritte bei der Simulation menschlicher Kommunikation; zweitens, die bestehenden (technischen) Herausforderungen, insbesondere im Bereich der nonverbalen Kommunikation; und drittens, die Implikationen dieser Entwicklungen für die Mensch-Agent-Interaktion. Das Verständnis dieser Aspekte ist entscheidend für die Entwicklung geeigneter Forschungsmethoden zur Analyse von HAI, insbesondere im Kontext audiovisueller Datenerhebung und -auswertung.

Moderne ECAs haben durch die Integration von Gesichtsmikroausdrücken, Gesten und fortschrittlichen Dialogfähigkeiten bedeutende technologische Fortschritte erzielt, um menschliche Eigenschaften präziser zu simulieren (Mori/McDorman/Kageki 2012). Trotz dieser Fortschritte bleibt die Entwicklung realistischer, lebensechter verkörperter Agenten durch das als »Uncanny Valley« bekannte Phänomen herausgefordert (Burden/Savin-Baden 2020). Folglich nehmen Nutzer*innen diese virtuellen Agenten oft als eine ambivalente Mischung aus menschenähnlich und gleichzeitig befremdlich wahr (Doyle et al. 2019; Lowe et al. 2017).

Die jüngsten Fortschritte in der visuellen Darstellung, Text-zu-Sprache- und Sprache-zu-Text-Technologien sowie verbesserte Dialogfähigkeiten haben die Bedeutung nonverbaler Kommunikation als essenziellen Bestandteil komplexer HAI in den Vordergrund gerückt. Diese Entwicklung ist nachvollziehbar: Nutzer*innen präferieren nicht nur verkörperte, realistische und menschenähnliche Agenten-Visualisierungen gegenüber reinen Sprachschnittstellen; diese ECAs fördern zudem das intuitive Verständnis und intensivieren das Gefühl der Verbundenheit (Huang 2018; Pauw et al. 2022).

Ungeachtet der beachtlichen Fortschritte in der Verarbeitung natürlicher Sprache durch Konversationsagenten behalten Menschen einen signifikanten Vorsprung in der Verarbeitung multimodaler Informationen. Die menschliche Kommunikation ist vielschichtig und umfasst neben der Sprache auch Gesten, paraverbale Elemente, Mimik und eine Vielfalt expressiver Körperhaltungen. Im Kontrast dazu sind gegenwärtige virtuelle Agenten in ihrer nonverbalen Ausdrucksfähigkeit stark limitiert und verfügen i. d. R. nicht über die Kapazität, nonverbale Signale menschlicher Nutzer*innen adäquat zu interpretieren (van Pinxteren/Pluymaekers/Lemmink 2020).

Über die technischen Herausforderungen hinaus weist nonverbale Kommunikation eine inhärente Mehrdeutigkeit auf, die nicht nur ECAs vor erhebliche

Herausforderungen stellt. Watzlawick, Beavin und Jackson (1967) betonen bekanntlich die Ambiguität nonverbaler Ausdrücke: Tränen können sowohl Trauer als auch Freude symbolisieren, eine geballte Faust kann Aggression oder Selbstbeherrschung signalisieren, ein Lächeln kann Zuneigung oder Geringschätzung ausdrücken, und Zurückhaltung lässt sich als Taktgefühl oder Indifferenz deuten. Diese *inhärente* Mehrdeutigkeit veranlasste Watzlawick et al. (1967) zur Differenzierung zwischen digitalen und analogen Kommunikationsmodalitäten. Digitale Codes repräsentieren den konkreten Inhalt und die wörtliche Bedeutung des Gesagten, wohingegen analoge Codes die Art und Weise der Kommunikation vermitteln.

Diese Differenzierung impliziert, dass Individuen simultan kontradiktorische Botschaften übermitteln können – ein Umstand, der virtuelle Agenten vor erhebliche Herausforderungen bei der Interpretation nonverbaler Aktionen stellt (Weiss et al. 2015; Freigang/Klett/Kopp 2017). Selbst unter der Annahme, dass virtuelle Agenten technisch fähig wären, diese Informationen zu erfassen, stellt die präzise Deutung der komplexen und häufig ambivalenten nonverbalen Signale daher eine beträchtliche Herausforderung dar (Kontogiorgos/Pereira/Gustafson 2021). Diese Problematik wirft die zentrale Frage auf, wie virtuelle Agenten konzipiert werden müssten, um diese vielschichtigen nonverbalen Signale effektiv zu dekodieren und zu interpretieren. Während ein ECA leicht programmiert werden kann, um auf das Wort »Hallo« mit einer entsprechenden Begrüßung zu reagieren (digitale Kommunikation), ist es ungleich schwieriger, ihn so zu gestalten, dass er die subtilen Nuancen eines Lächelns korrekt interpretiert und darauf angemessen reagiert (analoge Kommunikation). Diese Komplexität unterstreicht die Notwendigkeit fortschrittlicher Forschungsmethoden, die in der Lage sind, die vielschichtigen Aspekte der HAI zu erfassen und zu analysieren.

Zudem konzentriert sich die Forschung zu HAI gegenwärtig häufig (gezwungenermaßen) ausschließlich auf die inhaltlichen Ebenen, wobei die relationale Beziehungsdimension weitgehend außer Acht gelassen oder gar vernachlässigt wird. Eine rein inhaltsbezogene Analyse erweist sich jedoch aus den bereits genannten Gründen häufig als unzureichend: Auch die Ethnomethodologie argumentiert, dass soziale Realität von Natur aus instabil ist; soziale Ordnung und Strukturen existieren lediglich als temporäre, störungsanfällige Konstrukte, die einer kontinuierlichen Erneuerung, Anpassung und Reparatur bedürfen (Hayashi/Raymon/Sidnell 1992).

Obwohl spezialisierte Algorithmen etwa Gesichtsausdrücke als potenzielle Emotionsindikatoren erfassen können, stellt die Interpretation komplexerer nonverbaler Signale wie Handbewegungen, Intonation oder Körperhaltung hinsichtlich ihrer möglichen Bedeutungen eine weitaus größere Herausforderung dar (Kucherenko et al. 2021, 2022). Angesichts des gegenwärtigen Stands der Technik zeigen sich menschliche Nutzer dem virtuellen Agenten somit bei der Entschlüsselung dieser impliziten Botschaften deutlich überlegen (Saberi/

DiPaola/Bernardet 2021). Bei auftretenden Missverständnissen obliegt es i. d. R. ausschließlich den menschlichen Nutzer*innen, korrigierend in den Kommunikationsprozess einzugreifen (Meadan/Halle 2004). Den Nutzer*innen stehen diverse Strategien zur kommunikativen Reparatur zur Verfügung, darunter Wiederholung, Elaboration oder Themenwechsel (Dingemanse et al. 2015).

Hiermit wird evident, dass der gegenwärtigen HAI eine entscheidende relationale Dimension fehlt. Es offenbart sich eine signifikante Diskrepanz zwischen den Fähigkeiten zur verbalen und nonverbalen Kommunikation in der Interaktion mit virtuellen Agenten. Diese Defizite treten besonders prägnant in Studien zutage, die das Uncanny-Valley-Phänomen in Mensch-Agent-Interaktionen untersuchen (Thaler/Schlögl/Groth 2020), da hier sowohl der situative Kontext als auch die spezifische Dynamik zwischen Nutzer*in und Agent nicht adäquat erfasst und verarbeitet werden können (Weiss et al. 2015).

3. Die Analyse von Mensch-Agent-Interaktionen

Die Erforschung von HAI erfordert innovative methodische Ansätze, die der Komplexität dieser neuartigen Interaktionsform gerecht werden. In diesem Abschnitt stellen wir eine Methodik vor, die auf etablierten sozialwissenschaftlichen Traditionen aufbaut und diese für den spezifischen Kontext der HAI adaptiert. Unser Ansatz integriert Elemente der Ethnomethodologie, Konversationsanalyse und Workplace Studies mit modernsten Mixed-Reality-Technologien.

Die Methodik gliedert sich in folgende Hauptschritte:

1. Theoretische Fundierung in Ethnographie und Konversationsanalyse
2. Einsatz von Mixed-Reality-Videographie zur Datenerhebung
3. Entwicklung eines spezialisierten Transkriptionssystems für Mixed-Reality-Daten
4. Detaillierte Analyse von Human-Agent-Interaction Sequenzen

Jeder dieser Schritte wird im Folgenden ausführlich erläutert, wobei wir besonders auf die Innovationen und Anpassungen eingehen, die für die Erfassung und Analyse audiovisueller Daten in Kontexten der HAI notwendig sind.

In der sich rasch entwickelnden Domäne von VR und HAI offenbart sich wie oben dargestellt eine signifikante methodologische Lücke, vor allem bezüglich der Erforschung relationaler Dynamiken zwischen Nutzer*innen und verkörperten Konversationsagenten (Fitrianie 2022; Alexandrovsky et al. 2021). Jüngste Meta-Analysen in diesem Forschungsfeld bekräftigen diese Problematik und decken eine auffällige Vernachlässigung der Beziehungsebene zwischen Nutzer*in und Agent auf (Mara/Appel/Gnambs 2022). Zur Überwindung dieser Forschungslücke bedarf es eines paradigmatischen Wandels im Forschungsansatz – weg von

einer primär psychologischen hin zu einer dezidiert soziologischen Betrachtungsweise der HAI.

In der Untersuchung der relationalen Dimension kommunikativer Interaktionen ist es von zentraler Bedeutung, die Daten auf Muster zu prüfen, die sowohl implizite als auch explizite Indikatoren für die Beziehungsqualität und übergeordnete strukturelle Dynamiken liefern können. Watzlawick et al. betonen, dass ein ganzheitliches Verständnis der relationalen Dimension in Interaktionen nur durch die simultane Berücksichtigung gestischer Ausdrücke, wie Körpersprache, und verbaler Kommunikation erreicht werden kann. Selbst der Versuch, nicht zu kommunizieren, stellt bekanntlich eine Form interaktiven Verhaltens dar und übermittelt eine implizite Botschaft. Folglich sollte die Interaktionsanalyse ein besonderes Augenmerk auf atypische, unerwartete oder deviante Interaktionsmuster richten, da diese die Komplexität nonverbaler Kommunikation verdeutlichen und zur Entschlüsselung der vielschichtigen, ko-konstruierten Beziehungen zwischen den Interaktionspartner*innen beitragen können.

Zur effektiven Erforschung dieses komplexen Feldes bedienen wir uns eines fundierten Repertoires etablierter sozialwissenschaftlicher und soziologischer Methodologien, darunter Konversationsanalyse, Ethnomethodologie und Workplace Studies. Diese Methodologien haben sich als essenziell für die Aufschlüsselung menschlicher Interaktionsdynamiken erwiesen und ermöglichen tiefgreifende Einblicke in sequenzielle Muster und interaktionale Praktiken in diversen Kontexten (Mondada 2012). Die Adaption dieser Ansätze für HAI erfordert eine sorgfältige Analyse der Art und Weise, wie menschliche Nutzer*innen und virtuelle Agenten Interaktionsregeln und -praktiken kontinuierlich anwenden, um ein geteiltes Verständnis der Interaktion zu entwickeln.

3.1 Ethnographie, Konversationsanalyse und Workplace Studies

Die hier implementierte Forschungsmethodik fußt auf den Grundsätzen der Wissenschafts- und Technikforschung, wobei ein besonderer Schwerpunkt auf den Workplace Studies liegt. Das methodologische Instrumentarium des hier vorgestellten Ansatzes entstammt dabei der Ethnographie, Ethnomethodologie und Konversationsanalyse. Diese Methoden erweisen sich als besonders geeignet, um die Entfaltung technologievermittelter Kommunikation in Echtzeit zu untersuchen und Diskrepanzen zwischen den interpretativen Schemata menschlicher Nutzer*innen und den vorprogrammierten Konfigurationen technischer Systeme aufzudecken (Heritage 1998).

Seit Suchmans wegweisenden Studien haben sich Ethnographie und Konversationsanalyse als fundamentale Methoden zur Erforschung von Mensch-Maschine-Interaktionen etabliert (Suchman 2007, 2012). Workplace Studies haben sich zu einer eigenständigen Forschungstradition entwickelt, die die kontextbezogene und

interaktive Technologienutzung empirisch erforscht. Die Konversationsanalyse hat ihre Leistungsfähigkeit im Feld der Mensch-Roboter- und Mensch-Agent-Interaktionen mehrfach unter Beweis gestellt (Breazeal 2003; Mayer/Muhle/Bock 2020).

Ein zentrales Axiom der Konversationsanalyse ist die Hervorhebung des kontextuellen Charakters von Interaktionen (Heritage 1998; Garfinkel 1967). Dieser Ansatz postuliert, dass jede Interaktion in einem spezifischen Kontext eingebettet ist (Mayer/Muhle/Bock 2020), wobei die Relation zwischen Interaktion und Kontext als »reflexiv« konzeptualisiert wird (Garfinkel 1967, 2002). Fundamental für diesen Prozess ist das von Heritage akzentuierte Konzept der *reflexiven Dimension* sozialer Handlungen (Heritage 1998).

Es ist jedoch wichtig zu betonen, dass unsere Studie sich von klassischen Workplace Studies unterscheidet: Statt Routineabläufe zu untersuchen, befassen wir uns mit der Interaktion mit einem neuartigen, »ungewohnten Artefakt« (Krummheuer 2010). Diese Unvertrautheit wird besonders in Laborstudien deutlich, wo die Teilnehmer*innen mit neuen Prototypen – wie virtuellen Gesprächspartner*innen in VR – konfrontiert werden und dadurch völlig neue Erfahrungen machen.

Die Konversationsanalyse umfasst die minutiöse Untersuchung audiovisueller Interaktionsaufzeichnungen, mit dem Ziel, sowohl kontextunabhängige als auch kontextsensitive Regeln und Muster zu extrahieren. Diese identifizierten Muster dienen den Teilnehmer*innen als Orientierungsrahmen zum Verständnis und zur Navigation sozialer Situationen (Sacks/Schegloff/Jefferson 1974; Goffman 1972). Das Ziel ist es, die interaktiven Muster und organisierenden Strukturen innerhalb von Interaktionen aufzudecken, wobei der Hauptfokus auf den Kommunikationsereignissen zwischen Menschen und Maschinen liegt, statt sich in subjektiven Interpretationen des technischen Artefakts zu verlieren (Bock/Mayer 2020). Ethnographie und Konversationsanalyse erweisen sich somit als besonders geeignet für die Analyse von HAI aus mehreren Gründen. Fassen wir nochmals die Vorteile zusammen:

a) Kontextsensitivität: Diese Methoden berücksichtigen den spezifischen Kontext der Interaktion, was angesichts der Neuartigkeit und Unvertrautheit von HAI-Situationen besonders wichtig ist.
b) Fokus auf emergente Prozesse: Sie ermöglichen die Untersuchung, wie Bedeutung und Verständnis in der Interaktion zwischen Mensch und Agent gemeinsam konstruiert werden, statt von vorgefassten Kategorien auszugehen.
c) Multimodale Analyse: Insbesondere die Konversationsanalyse bietet Werkzeuge zur detaillierten Untersuchung verbaler und nonverbaler Kommunikationselemente, die in HAI eine zentrale Rolle spielen.
d) Offenheit für unerwartete Phänomene: Diese Ansätze erlauben es, neuartige Interaktionsmuster zu entdecken, die in der noch jungen Forschung zu HAI besonders relevant sein können.

3.2 Mixed-Reality-Videographie

Aufbauend auf dem dargelegten methodologischen Fundament präsentieren wir einen innovativen rekonstruktiven Ansatz für die Erhebung und Auswertung audiovisueller Daten in der Mixed-Reality-Nutzung. Der Ansatz charakterisiert sich durch ein außerordentlich hohes Maß an Flexibilität und Offenheit. Diese Offenheit ermöglicht die Erfassung diverser Datentypen: Einerseits können wir weiterhin die subjektiven Einschätzungen der Nutzer*innen bezüglich des Agenten mittels standardisierter Fragebögen und qualitativer Interviews erheben. Darüber hinaus eröffnet dieser Ansatz die Möglichkeit, die ko-konstruierte Interaktion zwischen Nutzer*in und Agent in situ zu observieren und zu analysieren. Über die Analyse rein verbaler Äußerungen hinaus ermöglicht diese Methode die systematische Auswertung nonverbaler Kommunikationselemente. Dieser umfassende Ansatz verspricht tiefgreifende Einsichten in diverse Interaktionsmodi, sowohl auf inhaltlicher als auch auf relationaler Ebene, die potenziell generalisierbare Interaktionsmuster mit virtuellen Entitäten offenlegen könnten.

Zur exemplarischen Veranschaulichung der methodischen Verfahrensweisen greifen wir auf Daten eines aktuellen Forschungsprojekts zurück. Der in dieser Studie zur Analyse multimodalen Verhaltens eingesetzte ECA entstammt dem Forschungsprojekt »Ai.vatar – Der virtuelle intelligente Assistent« (EFRE). Der Agent operiert mit gesprochener Sprache als Ein- und Ausgabemedium und verwendet Google DialogFlow zur Verarbeitung natürlicher Sprache. Die Steuerung erfolgt über ein eigens entwickeltes Dialog-Management-System. Das System ermöglicht den Nutzer*innen eine direkte Interaktion mit dem Agenten via Sprachsteuerung über das integrierte Mikrofon des VR-Headsets (Valve Index). Die visuelle Repräsentation des Agenten basiert auf photogrammetrischen Aufnahmen und umfasst einen vollständig geriggten sowie animierten Körper, realisiert mittels Unreal Engine 4.27. Die technische Implementierung des Projekts erfolgte in Kooperation mit den beiden Unternehmen HHVision und IOX.

Abbildung 1: Nutzung von Green Screens für die Mixed-Reality-Darstellung

Im Gegensatz zu herkömmlichen Videographie-Ansätzen, die oft nur eine Perspektive erfassen, fusioniert unsere Methode mehrere Datenströme (vgl. Thiel-Woznica i. d. B.):

1. Externe Kamera-Aufnahmen der Nutzer*innen im physischen Raum
2. In-VR-Aufnahmen aus der Perspektive der Nutzer*innen
3. Renderings des virtuellen Agenten und der VR-Umgebung

Diese Datenströme werden in Echtzeit zu einer kohärenten Mixed-Reality-Darstellung zusammengeführt (siehe Abb. 1). Die technische Umsetzung erfolgt durch die Integration des LIV Mixed Reality Software Developer Kit (SDK) in unsere Anwendungsarchitektur (LIV 2024). Ein zentraler Vorteil dieser Methode ist die Möglichkeit, sowohl die physischen Aktionen des Nutzers als auch deren Auswirkungen in der virtuellen Umgebung simultan zu erfassen. Dies erlaubt uns, subtile Aspekte der Interaktion zu analysieren, die bei herkömmlichen Methoden oft verloren gehen, wie etwa die Synchronisation von Körperbewegungen und virtuellen Reaktionen. Die größte Herausforderung bei diesem Ansatz liegt in der präzisen Synchronisation der verschiedenen Perspektiven. Wir lösen dies durch eine sorgfältige Kalibrierung und kontinuierliche Anpassung der Kamerapositionierung relativ zum VR-Tracking-System.

3.3 Transkription von Mixed-Reality-Videodaten

Die ethnomethodologisch fundierte Konversationsanalyse zeichnet sich durch ihre Unabhängigkeit von einem starren Set präskriptiver Methoden aus. Vielmehr orientiert sich die Analyse an Garfinkels Prinzip der »unique adequacy of methods« (Garfinkel 2002), welches eine auf den spezifischen Untersuchungsgegenstand zugeschnittene analytische Herangehensweise fordert. Dennoch lassen sich gewisse generelle Verfahrensschritte in der Konversationsanalyse identifizieren: Initial erfolgt die Erstellung audiovisueller Aufzeichnungen der zu untersuchenden Interaktion für analytische Zwecke. Diese Aufzeichnungen fixieren die temporale Struktur der Interaktionssituation und eröffnen die Möglichkeit einer Analyse der Interaktionsprozesse mittels detaillierter Transkriptionen. Sacks postuliert für den Transkriptionsprozess das Prinzip »order at all points« (Sacks 1984), welches impliziert, dass Interaktionssequenzen nicht arbiträr, sondern Manifestationen einer von den Teilnehmer*innen ko-konstruierten, latenten sozialen Ordnung sind.

Die Untersuchung multimodaler Interaktionen erfordert die Nutzung audiovisueller Dokumente, um ein breites Spektrum kommunikativer Elemente zu erfassen – von Sprecher*innenwechseln und Unterbrechungen bis hin zu diversen Ausdrucksmodalitäten wie Blickkontakt, Gestik und Mimik. Diese Analyse

basiert primär auf Videoaufzeichnungen zur Rekonstruktion von Interaktionsstrukturen, mit besonderem Augenmerk auf der praktischen Orchestrierung des Äußerungsaustauschs zwischen Nutzer*innen und Agenten (Tuma/Schnettler/Knoblauch 2013; Schmitt 2015). Zentrale Untersuchungsaspekte umfassen häufig das Turn-Taking, nonverbale Verhaltensweisen sowie Strategien zur Bewältigung potenzieller Interaktionsstörungen.

Im Gegensatz zu rein auditiven Daten erfordern audiovisuelle Daten spezifische Transkriptionsanforderungen (Knoblauch/Tuma 2011). Eine umfassende Dokumentation sowohl verbaler als auch nonverbaler Ausdrucksformen und Handlungen, die im Verlauf einer Interaktion auftreten, ist essenziell. Das Spektrum zu erfassender Verhaltensweisen erstreckt sich dabei von paralinguistischen Phänomenen wie Lachen und Räuspern über Pausen und unvollendete Äußerungen bis hin zu Handlungssynchronisationen, Unterbrechungsmustern und prosodischen Besonderheiten, einschließlich etwaiger Interaktionen mit anwesenden Experimentatoren. Bartneck et al. (2019) identifizieren zudem ein breites Spektrum nonverbaler Signale, die für die HRI relevant sind: Diese reichen von Blickverhalten und Okulomotorik über Gestik, Mimik und taktile Interaktionen bis hin zu Körperhaltung, Bewegungsdynamik und Interaktionsrhythmus sowie weiteren sozialen Mechanismen der *face-to-face*-Interaktion, einschließlich Reparaturstrategien bei Kommunikationsstörungen (Sidnell 2010). Darüber hinaus schlägt Suchman (2012) vor, dass die Einbeziehung der Perspektiven der beteiligten Akteure den Transkriptionsprozess für Mensch-Maschine-Interaktionen verbessern kann. Suchman plädiert für ein tabellarisches Transkriptionsformat, welches die wechselseitige Transparenz und Zugänglichkeit der Handlungen von Nutzer*in und Maschine explizit berücksichtigt und visualisiert (vgl. Tab. 1).

Tabelle 1: Suchmans Perspektivierung von Mensch-Maschine-Interaktionsprozessen

User	**User**	**Machine**	**Machine**
Situational interpretations of the human users that are *not accessible* to the machine.	Actions of the human users that *are accessible* to the machine.	Actions of the machine that *are accessible* to the human users.	Machine processes that are *not accessible* to the human users.

Die Anwendung von Suchmans analytischem Rahmenwerk auf die Interaktionen in unserem Forschungsprojekt offenbart schnell eine markante Asymmetrie in der reziproken Wahrnehmungsfähigkeit: Die menschlichen Nutzer*innen verfügen über die Fähigkeit, die Kommunikation des Agenten sowohl auf verbaler als auch auf nonverbaler Ebene zu dekodieren, wodurch sie nicht nur den Inhalt der Aussagen erfassen, sondern auch deren Vermittlungsmodus interpretieren können. Der Agent hingegen ist in seiner Wahrnehmung auf die rein verbale Ebene der Kommunikation limitiert. Bedingt durch die technischen Limitationen von

DialogFlow erfolgt die Interpretation des kommunikativen Austauschs seitens des Agenten ausschließlich im Rahmen eines digitalen, informationsverarbeitenden Paradigmas. Folglich mangelt es dem Agenten an der Kapazität, subtile sprachliche Nuancen wie Stimmlagen zu erfassen. Darüber hinaus ist der gegenwärtige Prototyp nicht befähigt, körpersprachliche Signale, Haltungen oder mimische Ausdrücke der Nutzer*innen zu registrieren und zu interpretieren. Im Kontext des aktuellen technologischen Entwicklungsstands verfügen menschliche Nutzer*innen daher derzeit noch über ein signifikant umfangreicheres Repertoire an Wahrnehmungsfähigkeiten, welches die beschränkten Kapazitäten des Agenten deutlich übersteigt.

Auf einer stärker erkenntnistheoretischen Ebene können wir weitere Ähnlichkeiten und Diskrepanzen erkennen: Analog zur Intransparenz menschlicher kognitiver Prozesse bleibt die interne Funktionsweise des Agenten für die Nutzer*in verborgen, während umgekehrt die kognitive Architektur der Nutzer*in für den Agenten ebenfalls unzugänglich ist. Auf dieser fundamentalen Ebene manifestiert sich somit eine bemerkenswerte Symmetrie in der gegenseitigen Unzugänglichkeit interner Prozesse. Hinsichtlich der Undurchschaubarkeit der jeweiligen internen Strukturen herrscht eine symmetrische ›Blindheit‹: Auch wir Menschen verwenden verbale und nonverbale Hinweise, um auf die internen Abläufe des anderen zu schließen, gerade weil sie unzugänglich sind (Luhmann 1984). Bezüglich der Inferenzmöglichkeiten auf diese internen Organisationen offenbart sich jedoch eine markante Asymmetrie in der Vielfalt und Tiefe der verfügbaren Interpretationsansätze. Die Nutzer*innen müssen sich den potenziell fundamentalen Differenzen in der Funktionsweise und ›Denkstruktur‹ des Agenten bewusst sein und diese in ihre Interaktionsstrategien einbeziehen. Die Fremdartigkeit des Agenten ist somit eine stärkere Fremdartigkeit als in interkulturellen Situationen und nähert sich eher einem Paradigma des interspezifischen Verstehens an (Schetsche 2004).

Bei der Transkription audiovisueller Daten steht man oft vor der Herausforderung, ein Gleichgewicht zwischen Lesbarkeit und Detailgenauigkeit zu finden (Knoblauch/Schnettler/Tuma 2013). Die Detailtiefe der Beobachtungen und Notationen richtet sich folglich nach der spezifischen Forschungsfrage der jeweiligen Studie. Bei der Erstellung der Transkripte ist es zudem wesentlich, neben den verbalen Äußerungen auch alle bedeutsamen körperlichen Handlungen und nonverbalen Signale der Teilnehmer*in zu dokumentieren. Während verbale Äußerungen nach gängigen Konventionen notiert werden, kennzeichnet man nonverbale Aktivitäten durch Klammern, um sie als Interpretationen der Beobachter*in kenntlich zu machen. In manchen Fällen kann es sinnvoll sein, die Texttranskription durch ausgewählte Bildschirmaufnahmen aus dem Mixed-Reality-Video zu ergänzen, um visuelle Aspekte der Interaktion zu verdeutlichen. Konsequenterweise erfolgt eine chronologisch strukturierte Transkription sämtlicher Aktivitäten beider Interaktionspartner*innen, um die sequenzielle

Natur der Interaktion adäquat zu erfassen. Der Vorteil dieses Ansatzes ist die Möglichkeit, simultane Handlungen angemessen zu dokumentieren, ohne sie in ein sequenzielles Format zwängen zu müssen. Dieser Ansatz erlaubt es, die Ereignisse diachron von links nach rechts und synchron von oben nach unten zu lesen. Letztlich ist das Hauptziel der Transkription, eine solide Grundlage für die anschließende Analyse und Interpretation zu schaffen.

Im nächsten Abschnitt präsentieren wir ein Beispiel eines solchen Transkripts, um die beschriebene Vorgehensweise zu veranschaulichen. Die hierzu verwendeten audiovisuellen Daten entstammen einem technischen Pretest unserer Studie.

3.4 Mensch-Agent-Interaktion (Sequenz 1)

Tabelle 2: Interaktionssequenz 1

Abbildung 2: Mixed-Reality-Ansicht (Sequenz 1)

Segment	**01**	**02**	**03**
Zeitmarke	*00:01:12*	*00:02:01*	*00:03:20*
User Sprache	*»Hi«*		
User Blick	Agent Kopf	User Hände	Agent
User Körperhaltung	Agent		
User Gestik	steht still	Handbewegungen	
User Bewegung			
Agent Sprache			*»Hallo, mein Name ist Florian. Wie heißt du?«*
Agent Blick	User	Agent Hände	User
Agent Körperhaltung	User		
Agent Gestik	steht still	Handbewegungen	steht still
Agent Bewegung			

Die ersten Interaktionssequenzen (Tab. 2) zeigen, wie Nutzer und Agent gegenseitig ihre Anwesenheit wahrnehmen und anerkennen. Der Nutzer initiiert die Interaktion, indem er Blickkontakt mit dem Agenten aufnimmt und ihn direkt anschaut, bevor er die verbale Kommunikation mit einem schlichten »Hi« eröffnet. Diese

Verhaltensweisen entsprechen den üblichen Mustern in persönlichen Begegnungen, bei denen Blickkontakt als erste Bestätigung der gegenseitigen Wahrnehmung dient. Interessanterweise imitiert der Nutzer das Verhalten des Agenten, etwa beim gegenseitigen Betrachten der Hände (Abb. 2). Dies deutet anfänglich auf eine Situation hin, in der beide Akteure scheinbar ähnlich mit Unsicherheiten umgehen.

Wenn der Nutzer seinen Blick auf den sprechenden Agenten richtet, zeigt dies, dass er die üblichen Gesprächskonventionen wie Aufmerksamkeit und Höflichkeit einhält – ähnlich wie in Gesprächen zwischen Menschen. Auch der Agent scheint die Anwesenheit des Nutzers zunächst durch Blickkontakt wahrzunehmen. Anders als der Nutzer beginnt der Agent nicht mit einer verbalen Begrüßung. Dies erklärt sich aus unserem Hintergrundwissen: Der Agent ist so programmiert, dass er auf Eingaben reagiert, statt sie selbst zu initiieren. Die erste Reaktion des Agenten auf das »Hi« – ein Blick auf die eigenen Hände – lässt Raum für verschiedene Deutungen: Es könnte Unsicherheit, Verlegenheit oder sogar Desinteresse signalisieren. Die ausbleibende verbale Antwort des Agenten und seine Fokussierung auf die eigenen Hände könnten als eine Art nonverbale Botschaft interpretiert werden.

Letztendlich gibt der Agent doch eine verbale Antwort und vermeidet so den möglichen Eindruck von Unhöflichkeit. Eine alternative Erklärung für die verzögerte Reaktion des Agenten könnte in unterschiedlichen Verarbeitungsgeschwindigkeiten liegen. Möglicherweise benötigt der Agent mehr Zeit, um eine passende Antwort zu generieren. Diese Deutung lässt sich zum jetzigen Zeitpunkt allerdings nicht endgültig verifizieren. Um diese Hypothese der unterschiedlichen Verarbeitungszeiten zu stützen, wären weitere Beispiele verzögerter Antworten nötig, vergleichbar mit einer »Loading … Please wait«-Situation.

Die anfängliche Interaktion zwischen Nutzer und Agent weist Merkmale auf, die für Erstkontakte zwischen Fremden charakteristisch sind: Zurückhaltung, vorsichtiges Abtasten und rituelle Wahrnehmungsprozesse. In diesem Stadium sind die Beziehungsdynamiken zwischen den beiden Akteuren noch nicht klar definiert. Auffällig ist die verzögerte Reaktion des Agenten auf die Begrüßung sowie sein Blick auf die eigenen Hände. Dies könnte auf einen ungewöhnlichen Interaktionskontext hindeuten. Interessanterweise gleicht der Nutzer diese Unregelmäßigkeit aus, indem er das Verhalten des Agenten spiegelt und ebenfalls auf seine Hände schaut. Durch diese symmetrische und synchronisierende Geste mindert der Nutzer die anfängliche Diskrepanz und hilft, einen normaleren Interaktionsfluss herzustellen.

Die Mixed-Reality-Videographie ermöglichte es uns, sowohl die physischen Aktionen des Nutzers als auch die virtuellen Reaktionen des Agenten simultan zu erfassen. Dies war besonders wertvoll bei der Analyse der Blickbewegungen und Körperorientierungen, die in traditionellen Videostudien oft schwer zu erfassen sind. Eine Limitation unserer Methode zeigte sich in der Schwierigkeit, subtile Gesichtsausdrücke des Nutzers aufgrund der VR-Brille vollständig zu erfassen.

3.5 Mensch-Agent-Interaktion (Sequenz 2)

Tabelle 3: Interaktionssequenz 2

Abbildung 3: Mixed-Reality-Ansicht (Sequenz 2)

Segment	**04**	**05**	**06**
Zeitmarke	*00:04:15*	*00:05:30*	*00:07:22*
User Sprache	*»Mein Name ist John Doe.«*		
User Blick	Agent Kopf	Agent Kopf	Agent
User Körperhaltung	Agent		
User Gestik		Tätscheln, tasten des Kopfes des Agenten	
User Bewegung	Lehnt sich zu Agent		
Agent Sprache			*»Schön dich kennenzulernen, John Doe.«*
Agent Blick	User	User	User
Agent Körperhaltung	User		
Agent Gestik	Hände am Körper		
Agent Bewegung	steht still		

In seiner Antwort an den Agenten nennt der Nutzer nicht seinen tatsächlichen Namen, sondern das Pseudonym »John Doe« (Tab. 3). »John Doe« ist im US-amerikanischen Sprachraum ein gängiger Platzhalter-Name, der oft zur Wahrung der Anonymität eingesetzt wird. Diese Wahl eines Pseudonyms bringt eine gewisse Unaufrichtigkeit in den inhaltlichen Aspekt der Interaktion. Die Diskrepanz zur Realität wird durch die nächste Aktion des Nutzers noch verstärkt: Er tätschelt oder stupst den Kopf des virtuellen Agenten an (Abb. 3). In anderen typischen Szenarien, in denen sich Fremde kennenlernen, würde eine solche Geste als klarer Verstoß gegen konventionelle soziale Normen betrachtet werden. Allerdings ist zu berücksichtigen, dass diese Begegnung zwischen Nutzer und Agent im Rahmen des Forschungsprojekts nicht die erste ihrer Art ist. Die Berührung des Agenten-Kopfes kann als Versuch interpretiert werden, die Reaktionsfähigkeit des Agenten zu testen oder zu provozieren. Diese Aktion lässt sich auf zweierlei Weise deuten: als Ausloten der Fähigkeiten und Grenzen des

Agenten und als Indikator für ein Ungleichgewicht in der Beziehung zwischen Nutzer und Agent. Unter ethischen Gesichtspunkten könnte man dieses Verhalten als Grenzüberschreitung bewerten.

Aus der Perspektive des Agenten gestaltet sich die Situation grundlegend anders. Die von externen Beobachtern erkannte mögliche Grenzverletzung im Verhalten des Nutzers liegt außerhalb der Wahrnehmungsfähigkeit des Agenten. Diese Unfähigkeit zur Wahrnehmung und Reaktion auf solche Handlungen entspricht genau dem Zweck des Tests: Sie offenbart die Limitationen des Agenten. Gemäß Suchmans Modell erkennen wir, dass der Agent technisch nicht in der Lage ist, die Handbewegungen des Nutzers als relevante Eingabe für sein Antwortverhalten zu verarbeiten. Folglich existieren diese Handlungen in der Wahrnehmungswelt des Agenten nicht und werden daher nicht als Regelverletzungen registriert.

Ohne eine Kenntnis der genauen Fähigkeiten des Agenten lässt sich jedoch nicht eindeutig feststellen, ob dieser die provokative Geste des Nutzers nicht wahrnahm oder bewusst ignorierte. Wenn Letzteres zutreffen würde, müsste der Nutzer über das Motiv hinter einer solchen Nicht-Reaktivität nachdenken. Die verbale Reaktion des Agenten (»Schön, dich kennenzulernen, John Doe.«) klärt diese Ambiguität nicht auf. Sie könnte sowohl auf begrenzte Fähigkeiten hindeuten als auch eine bewusst reservierte Haltung widerspiegeln, wie man sie etwa von britischen Wachsoldat*innen oder Butler*innen kennt. Davon unbeeindruckt fällt auf, dass die verbale Interaktion flüssiger verläuft als in der ersten Sequenz.

Die ersten drei Interaktionssequenzen lassen somit ein klares Muster erkennen. Es zeigt sich eine deutliche Asymmetrie und Hierarchie in der Interaktion. Diese manifestiert sich sowohl auf der Beziehungsebene, vorwiegend durch nonverbale Gesten, als auch auf der Informationsebene in der verbalen Kommunikation. Der Nutzer geht in seinen Aktionen über ein simples Testen der Agentenfähigkeiten hinaus. Er demonstriert aktiv die begrenzten Möglichkeiten des Agenten im Bereich der multimodalen Kommunikation. Diese Beobachtungen werfen Fragen zu den Motiven des Nutzers auf: Zielt sein Verhalten darauf ab, die Schwächen des Agenten herabwürdigend aufzuzeigen, oder handelt es sich um einen konstruktiven Versuch, die Systemgrenzen auszuloten, vergleichbar mit den praktischen Aspekten eines Turing-Tests (Turing 1950; Humphreys 2009)? So offenbart sich die Vielschichtigkeit der Nutzerhandlungen, die auf komplexe und möglicherweise mehrfach motivierte Absichten hindeuten.

Die Mixed-Reality-Videographie erwies sich hier als besonders wertvoll, da sie es ermöglichte, die physische Interaktion des Nutzers mit dem virtuellen Agenten präzise zu erfassen. Traditionelle Videomethoden hätten Schwierigkeiten gehabt, diese Interaktion zwischen realer und virtueller Welt adäquat darzustellen. Eine Herausforderung unserer Methode zeigte sich in der Interpretation

der Motivation der Nutzer*innen für die ungewöhnliche physische Interaktion. Hier wird eine Ergänzung durch nachträgliche Interviews mit den Nutzer*innen hilfreich, um ihre Gedanken und Absichten besser zu verstehen.

3.6 Mensch-Agent-Interaktion (Sequenz 3)

Tabelle 4: Interaktionssequenz 3

Abbildung 4: Mixed-Reality-Ansicht (Sequenz 3)

Segment	**07**	**08**	**09**
Zeitmarke	*00:08:11*	*00:09:04*	*00:10:10*
User Sprache			
User Blick	Agent	Agent	Agent
User Körperhaltung	Agent		
User Gestik	Hände am Körper		
User Bewegung	Bewegt sich um Agenten herum	Bewegt sich um Agenten herum	Bewegt sich um Agenten herum
Agent Sprache		*»Was kann ich für dich tun?«*	
Agent Blick	User	User	User
Agent Körperhaltung	User		
Agent Gestik	steht still		
Agent Bewegung	steht still		

In den drei Segmenten der dritten Sequenz (Tab. 4) verzichtet der Nutzer auf jegliche weitere verbale Kommunikation. Der Nutzer beschränkt sich stattdessen darauf, sich um den Agenten zu bewegen und dabei konstanten Blickkontakt mit dessen Gesicht zu halten (Abb. 4). Diese Verhaltensweise geht über ein einfaches Passieren des Agenten hinaus; sie suggeriert eine eingehende Beobachtung, gepaart mit Skepsis und prüfender Haltung.

Der Agent passt seinen Blick und seine Körperhaltung den Bewegungen des Nutzers an, was auf eine fortlaufende gegenseitige Wahrnehmung hindeutet. Die Sequenz schließt mit der Frage des Agenten, wie er dem Nutzer behilflich sein kann. Dies unterstreicht erneut seine dienende, untergeordnete Position in der gemeinsamen Interaktion.

Betrachtet man diese drei exemplarischen Sequenzen aus interaktionaler Sicht, zeigt sich durchgängig eine vom Nutzer dominierte Testdynamik. Die

Asymmetrie in Status und Beziehung bleibt konstant sichtbar und findet ihren Höhepunkt in der dienstleistungsorientierten Frage des Agenten. Diese Frage zementiert die dienende, reaktive Rolle des Agenten und bestätigt die Annahme, dass diese Interaktionsdynamik stabil bleibt. Damit scheint sich das Beziehungsmuster zwischen den beiden Akteuren zu verfestigen und illustriert Goffmans Konzept der »Interaktionsordnung«, wobei hier eine neue, spezifische Ordnung für die Mensch-Agent-Interaktion ausgehandelt wird.

Die Bewegungen des Nutzers und die Anpassungen des Agenten zeigen eine verbesserte Synchronizität im Vergleich zu früheren Sequenzen. Es entsteht eine spürbare Spannung durch die anhaltende Stille, die schließlich vom Agenten durchbrochen wird. Die Aufrechterhaltung des Blickkontakts durch beide Akteure signalisiert zumindest eine fortlaufende gegenseitige Aufmerksamkeit und ein Engagement in der Interaktion. Die (vermeintliche) Anpassungsfähigkeit des Agenten in Bezug auf Blick und Körperhaltung demonstriert (vermeintlich) fortgeschrittene nonverbale Interaktionsfähigkeiten.

Die Mixed-Reality-Videographie erwies sich in dieser Sequenz als besonders wertvoll:

a) Sie ermöglichte die präzise Erfassung der räumlichen Bewegungen und Blickrichtungen beider Akteure, was für die Analyse der nonverbalen Kommunikation entscheidend war.
b) Die simultane Aufzeichnung der realen Nutzerbewegungen und der virtuellen Agenten-Reaktionen erlaubte eine detaillierte Untersuchung ihrer Synchronizität.

Eine Herausforderung unserer Methode zeigte sich in der Interpretation der Stille. Hier wären ergänzende physiologische Messungen (z. B. Hautleitwert, Herzrate) denkbar, um die emotionale Verfassung der Nutzer*innen während dieser intensiven nonverbalen Phase besser zu verstehen.

4. Schlussfolgerungen und Ausblick

Dieser Beitrag stellt einen innovativen methodischen Ansatz zur Untersuchung von Mensch-Agent-Interaktionen (HAI) im Kontext von Mixed-Reality-Technologien vor. Durch die Integration von Mixed-Reality-Videographie mit etablierten qualitativen Forschungsmethoden können wir tiefere Einblicke in die komplexen Dynamiken der HAI gewinnen. Im Folgenden fassen wir die Haupterkenntnisse zusammen, diskutieren die theoretischen und praktischen Implikationen unserer Arbeit und geben einen Ausblick auf zukünftige Forschungsrichtungen.

Zentrale Erkenntnisse dieses Beitrags sind:

1. Multimodale Interaktionsanalyse: Unsere Methode ermöglicht die simultane Erfassung und Analyse verbaler und nonverbaler Kommunikationselemente in HAI. Dies hat sich als entscheidend für das Verständnis der vielschichtigen Beziehungsdynamiken zwischen Mensch und Agent erwiesen.
2. Emergente Interaktionsmuster: Die Studie offenbarte, wie Nutzer*innen und Agenten gemeinsam neue Interaktionsnormen aushandeln, die von typischen zwischenmenschlichen Interaktionsmustern abweichen.
3. Asymmetrien in der Kommunikation: Wir konnten signifikante Unterschiede in den kommunikativen Fähigkeiten von Mensch und Agent identifizieren, insbesondere im Bereich der nonverbalen Kommunikation.

Unser Ansatz zielt darauf ab, mittels Mixed-Reality-Videographie Daten über die ko-konstruierten Verhaltensweisen von menschlichen Nutzer*innen und virtuellen Agenten in einer virtuellen Umgebung zu erheben. Diese Methodik lässt sich flexibel auf verschiedene Forschungsfragen anwenden und kann durch Daten aus der Erste-Person-Perspektive, wie Interviews und Fragebögen, sowie durch Dritte-Person-Daten externer Sensoren ergänzt werden.

Ein wesentlicher Vorteil dieser Methodik für Mixed-Reality-Videographie ist ihre Fähigkeit, sowohl verbale als auch nonverbale Interaktionen in der virtuellen Realität zu erfassen. Mixed-Reality-Methoden erlauben es, das koordinierte Verhalten von Nutzer*innen und virtuellen Agenten in einer geteilten virtuellen Umgebung zu erfassen. Dies ermöglicht ein tieferes Verständnis der Entwicklung (proto-)sozialer Interaktionen, als es durch rein verbale Transkriptionen möglich wäre. Unsere Methode ermöglicht dabei die Auswertung vielfältiger Daten: Zum einen lassen sich die verbalen Interaktionen mit dem Agenten in detaillierte Gesprächsprotokolle für eingehende Konversationsanalysen überführen. Zum anderen bietet der videographische Ansatz die Möglichkeit, Mimik, Gestik und weitere Formen nonverbaler Kommunikation sowohl der Nutzer*innen als auch des Agenten zu analysieren, was Aufschluss über individuelle Verhaltensmuster gibt.

Der Einsatz von Mixed-Reality-Videographie unterstreicht die theoretische Auffassung von Interaktion als ko-konstruktivem Prozess. Diese Perspektive gewinnt, wie Kopp und Krämer (2021) jüngst dargelegt haben, zunehmend an Bedeutung für die Entwicklung und Evaluation virtueller Agenten. Ihre Argumentation betont, dass für eine hohe »Interaktionsintelligenz« koordinative Mechanismen entscheidend sind. Dazu zählen die partnerspezifische Anpassung multimodaler Äußerungen, schnelles Turn-Taking, informatives Feedback und die gemeinsame Bewältigung von Missverständnissen.

Unser Beitrag unterstreicht daher die Notwendigkeit, das gängige Verständnis von Mensch-Technologie-Interaktion zu erweitern oder gar neu zu konzipieren,

indem der Fokus auf die Analyse relationaler Muster gelegt wird. Eine verstärkt relationale Perspektive betrachtet diese Muster als Schnittstellen, an denen sich eine soziale Koordination vollzieht. Diese integriert organisch-physische, maschinell-elektronische und symbolisch-kulturelle Elemente. Obwohl die Forschung zur Mensch-Computer-Interaktion (HCI) in technischer Hinsicht weit fortgeschritten ist, bleibt das soziologische Verständnis verkörperter und situierter Interaktionen in der gegenwärtigen Forschungslandschaft noch unterentwickelt (Harrison/Sengers/Tatar 2011). Die Anwendung eines relationalen Ansatzes in Theorie und Methodik transformiert die herkömmlichen ontologischen Sichtweisen auf Mensch und Technologie in ein Geflecht interdependenter Identitäten (Ihde 1990). Die relationale Soziologie hebt hervor, dass sich Identitäten von Menschen und Technologien in spezifischen Beziehungskontexten herausbilden, wodurch komplexe Netzwerke entstehen, die über einfache dyadische Strukturen hinausgehen. Diese relationale Sichtweise konzentriert sich darauf, jene Beziehungsmuster zu identifizieren, die zur Formung von Identitäten im Interaktionsprozess beitragen. Statt von vordefinierten Beziehungen oder Definitionen auszugehen, strebt die relationale Soziologie danach, die Konstitution dieser Beziehungen empirisch zu ergründen (Emirbayer 2009).

Diese Betrachtungsweise der Interaktion wirft zudem die grundlegende Frage auf, was eine ›gute‹ Interaktion ausmacht. Der Begriff der »Interaktionsqualität« erweist sich nach wie vor als vielschichtig und schwer fassbar (Deriu et al. 2021). Unser methodischer Ansatz ermöglicht es jedoch, spezifische Interaktionsaspekte zu erfassen, die Hinweise auf den Erfolg der Interaktion geben können. Beispielsweise lässt sich untersuchen, ob Mensch und Agent über den gesamten Interaktionsverlauf hinweg einen kohärenten Bezugsrahmen aufrechterhalten. Es stellt sich dann die Frage, ob die Kommunikation direkt zwischen den Akteuren stattfindet oder ob sie eher übereinander sprechen. Weist die Interaktion eine Symmetrie auf, oder zeigen sich Asymmetrien, die an Beziehungsmuster wie Eltern-Kind, Betreuer*innen und Personen mit besonderen Bedürfnissen oder Besitzer*in-Haustier erinnern? Solche Betrachtungen ermöglichen eine differenzierte Einschätzung der Interaktionsqualität und -effektivität, indem sie über oberflächliche Beobachtungen hinaus tieferliegende relationale Dynamiken erfassen.

Dieser Ansatz schafft nicht nur eine neutrale Grundlage für einen fundierten kommunikationstheoretischen Rahmen zur Analyse von Mensch-Nicht-Mensch-Interaktionen, sondern bezieht auch den historischen Kontext diverser sozialer Interaktionsformen mit ein. Die menschliche Geschichte zeigt eine Vielfalt an Interaktionsformen – mit Angehörigen fremder Kulturen, Menschen mit kognitiven Einschränkungen, Kleinkindern und Säuglingen, Tieren sowie in spirituellen Kontexten mit vorgestellten Geistern oder Gottheiten. Für jede dieser Interaktionsformen haben sich spezifische soziale Beziehungsmuster herausgebildet. Die Integration künstlich erzeugter nicht-menschlicher Entitäten

in dieses Interaktionsspektrum stellt eine folgerichtige Weiterentwicklung dar (Harth 2017).

Bei der Anwendung dieser Methodik zur Datenauswertung sind allerdings zwei zentrale Limitationen zu beachten. Zum einen sind verkörperte Konversationsagenten (ECAs) noch nicht Teil der alltäglichen Lebenswelt. ECAs unterliegen einer kontinuierlichen technologischen und funktionalen Evolution, wobei sich gesellschaftliche Konventionen für den Umgang mit diesen KI-gestützten Entitäten erst allmählich herausbilden. Folglich tragen Interaktionen mit ECAs und virtuellen Agenten nach wie vor einen explorativen und experimentellen Charakter. Die in unserer Studie beobachteten Interaktionen mit dem Prototyp-Agenten sind demnach inhärent von ihrer Neuartigkeit gekennzeichnet. Zum anderen trägt auch der Untersuchungskontext zur Neuartigkeit bei: Der Besuch eines Universitätslabors und die Teilnahme an einem Experiment verleihen der Interaktion einen außergewöhnlichen und singulären Charakter. Für manche Teilnehmer*in mag dies zudem die erste Erfahrung mit virtueller Realität darstellen, was die Situation zusätzlich mit Neuheit auflädt. Folglich sind sowohl der Interaktionspartner als auch das Untersuchungssetting von einem hohen Maß an Neuartigkeit, Unvertrautheit und Ambiguität geprägt. Die Teilnehmer*innen sehen sich mit der doppelten Herausforderung konfrontiert, eine angemessene Interaktionsform mit dem virtuellen Agenten zu finden und gleichzeitig ihr Verhalten vor der Kamera zu regulieren.

Für die weitere Entwicklung dieses Forschungsfeldes ist der technologische Fortschritt bei der Gestaltung virtueller Agenten von zentraler Bedeutung. Eine authentisch bidirektionale Interaktion, die sowohl verbale als auch nonverbale Kommunikation umfasst, setzt voraus, dass virtuelle Agenten die Fähigkeit entwickeln, nonverbale Signale der Nutzer*innen zu erfassen und zu interpretieren. Während die verbale Kommunikation gegenwärtig bereits relativ ausgeglichen verläuft, bleibt die nonverbale Ebene asymmetrisch: Lediglich die menschlichen Nutzer*innen können die potenziellen nonverbalen Signale des Agenten deuten. Die für Menschen selbstverständliche Fähigkeit zur multimodalen Kommunikation stellt für virtuelle Agenten und humanoide Roboter nach wie vor eine erhebliche technologische Hürde dar (Magnenat Thalmann et al. 2016).

Die aktuelle technologische Forschung und Entwicklung zielen jedoch darauf ab, diese Herausforderungen schrittweise zu bewältigen. Gegenwärtige Forschungsansätze fokussieren verstärkt auf die Entwicklung künstlicher Agenten, die ihre Umgebung wahrnehmen und beeinflussen, natürliche Sprache verarbeiten und generieren sowie angemessen auf menschliche Anfragen und Anweisungen reagieren können (DeepMind Interactive Agents Group 2021). Diese Bestrebungen zur Schaffung kontextsensitiver und multimodal kommunizierender Agenten werden zweifellos unsere (proto-)sozialen Interaktionen mit ihnen maßgeblich beeinflussen und verändern (Moon et al. 2020). In Verbindung mit den Fortschritten bei domänenübergreifenden Algorithmen des maschinellen

Lernens, die eine Anpassung an neuartige, unbekannte Umgebungen ermöglichen, eröffnen sich vielversprechende Perspektiven für die zukünftige Forschung im Bereich der Mensch-Agent-Interaktion. Diese antizipierten technologischen Fortschritte erfordern ihrerseits die Entwicklung und Anwendung fundierter und belastbarer Forschungsmethoden.

Die vorgestellte Methode der Mixed-Reality-Videographie hat sich hierbei als leistungsfähiges Werkzeug für die Analyse von HAI erwiesen. Sie ermöglicht die detaillierte Erfassung komplexer Interaktionsprozesse, die mit traditionellen Methoden schwer zu beobachten wären. Diese Methode trägt somit wesentlich zur Weiterentwicklung qualitativer Forschungsansätze im Bereich der audiovisuellen Datenerhebung und -analyse bei.

Literatur

Alexandrovsky, Dmitry/Putze, Susanne/Schwind, Valentin/Mekler, Elisa D./Smeddinck, Jan D./Kahl, Denise/Krüger, Antonio/Malaka, Rainer (2021): Evaluating User Experiences in Mixed Reality. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21), 87, S. 1–5. https://doi.org/10.1145/3411763.3441337.

Bartneck, Christoph/Belpaeme, Tony/Eyssel, Friederike/Kanda, Takayuki/Keijsers, Merel/Sabanovic, Selma (2019): Human-Robot Interaction. An Introduction. Cambridge: Cambridge University Press.

Bavelas, Janet/Hutchinson, Sarah/Kenwood, Christine/Hunt Matheson, Deborah (1997): Using face-to-face dialogue as a standard for other communication systems. In: Canadian Journal of Communication, 22(1), 5. https://doi.org/10.22230/cjc.1997v22n1a973.

Bock, Indra/Mayer, Henning (2020): Humanoide Roboter und virtuelle Agenten als Kommunikationsteilnehmer? Konversationsanalytische Studien der Mensch-Maschine-Interaktion. In: Ahner, Helen/Metzger, Max/Nolte, Mathis (Hrsg.): Von Menschen und Maschinen: Interdisziplinäre Perspektiven auf das Verhältnis von Gesellschaft und Technik in Vergangenheit, Gegenwart und Zukunft. Proceedings der 3. Tagung des Nachwuchsnetzwerks »INSIST«, 5.–7. Oktober 2018, Karlsruhe (INSIST-Proceedings, 3), S. 159–182.

Breazeal, Cynthia (2003): Toward sociable robots. In: Robotics and Autonomous Systems, 42(3–4), S. 167–175.

Brown, Tom/Mann, Benjamin/Ryder, Nick/Subbiah, Melanie/Kaplan, Jared/Dhariwal, Prafulla/Neelakantan, Arvind/Shyam, Pranav/Sastry, Girish/Askill, Amanda/Agarwal, Sandhini/Herbert-Voss, Ariel/Krueger, Gretchen/Henighan, Tom/Child, Rewon/Ramesh, Aditya/Ziegler, Daniel M./Wu, Jeffrey/Winter, Clemens/Hesse, Christopher/Chen, Mark/Sigler, Eric/Litwin, Mateusz/Gray, Scott/Chess, Benjamin/Clark, Jack/Berner, Christopher/McCandlish, Sam/Radford, Alec/Sutskever, Ilya/Amodei, Dario (2020): Language models are few-shot learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20), Article 159, S. 1877–1901.

Burden, David/Savin-Baden, Maggi (2020): Virtual Humans. Today and Tomorrow. Boca Raton: CRC Press.

Cassel, Justine/Sullivan, Joseph/Prevost, Scott/Churchill, Elizabeth (2000): Embodied Conversational Agents. Cambridge, Massachusetts, London: MIT Press.

DeepMind Interactive Agents Group (2021): Imitating Interactive Intelligence. In: arXiv:2012.05672v2 (28.08.2024).

Deriu, Jan/Rodrigo, Alvaro/Otegi, Arantxa/Echegoyen, Guillermo/Rosset, Sophie/Agirre, Eneko/Cieliebak, Mark (2021): Survey on evaluation methods for dialogue systems. In: Artificial Intelligence Review 54, S. 755–810. https://doi.org/10.1007/s10462-020-09866-x.

Digital Humans Inc (2024): Website. https://digitalhumans.com (28.08.2024).

Dingemanse, Mark/Roberts, Seán G./Baranova, Julija/Blythe, Joe/Drew, Paul/Floyd, Simeon/Gisladottir, Rosa S./Kendrick, Kobin H./Levinson, Stephen C./Manrique, Elizabeth/Rossi, Giovanni/Enfield, N. J. (2015): Universal Principles in the Repair of Communication Problems. In: PLOS ONE, e0136100. https://doi.org/10.1371/journal.pone.0136100.

Doyle, Philip R./Edwards, Justin/Dumbleton, Odile/Clark, Leigh/Cowan, Benjamin R. (2019): Mapping Perceptions of Humanness in Intelligent Personal Assistant Interaction. In: Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services. Article No.: 5, S. 1–12. https://doi.org/10.1145/3338286.334011.

Emirbayer, Mustafa (2009): Manifesto for a Relational Sociology. In: American Journal of Sociology, 103(2), S. 281–317.

Fitrianie, Siska/Bruijnes, Merijn/Li, Fengxiang/Abdulrahman, Amal/Brinkman, Willem-Paul (2022): The artificial-social-agent questionnaire: establishing the long and short questionnaire versions. In: Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents (IVA '22), 18, S. 1–8. https://doi.org/10.1145/3514197.3549612.

Fitrianie, Siska/Bruijnes, Merijn/Richards, Deborah/Abdulrahman, Amal/Brinkman, Willem-Paul (2019): What are We Measuring Anyway? A Literature Survey of Questionnaires Used in Studies Reported in the Intelligent Virtual Agent Conferences. In: Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents, S. 159–161.

Foster, Mary E. (2019): Face-to-Face Conversation: Why Embodiment Matters for Conversational User Interfaces. In: Proceedings of the 1st International Conference on Conversational User Interfaces (CUI '19), Article 13, S. 1–3.

Freigang, Farina/Klett, Sören/Kopp, Stefan (2017): Pragmatic Multimodality: Effects of Nonverbal Cues of Focus and Certainty in a Virtual Human. In: Beskow, John (Hrsg.): Intelligent Virtual Agents. 17th International Conference, IVA 2017, Stockholm, Sweden, August 27–30, S. 142–155.

Garfinkel, Harold (1967): Studies in Ethnomethodology. Englewood Cliffs: Prentice-Hall.

Garfinkel, Harold (2002): Ethnomethodology's Program. Working out Durkheim's Aphorism. Lanham: Rowman & Littlefield.

Goffman, Erving (1972): Encounters. Two Studies in the Sociology of Interaction. Hamrondsworth: Penguin.

Harrison, Steve/Sengers, Phoebe/Tatar, Deborah (2011): Making epistemological trouble: Third-paradigm HCI as successor science. In: Interacting With Computers, 23, S. 385–392.

Harth, Jonathan (2017): Empathy with non-player characters? An empirical approach to the foundations of human/non-human-relationships. In: Journal of Virtual Worlds Research, 10(2). https://journals.tdl.org/jvwr/index.php/jvwr/article/view/7272 (10.12.2024).

Hayashi, Makoto/Raymond, Geoffrey/Sidnell, Jack (1992): Conversational Repair and Human Understanding. Cambridge: Cambridge University Press.

Heritage, John (1998): Conversation Analysis and Institutional Talk: Analyzing Distinctive Turn-Taking Systems. In: Proceedings of the 6th International Congresss of IADA (International Association for Dialog Analysis), Tübingen, S. 3–17.

Huang, Hung-Hsuan (2018): Embodied conversational agents. In: Norman, Kent/Kirakowski, Jurek (Hrsg.): The Wiley handbook of human computer interaction. Hoboken: Wiley Blackwell, S. 601–614.

Humphrys, Mark (2009): How My Program Passed the Turing Test. In: Epstein, Robert/Roberts, Gary/Beber, Grace (Hrsg.): Parsing the Turing Test. Dordrecht: Springer, S. 237–260. https://doi.org/10.1007/978-1-4020-6710-5_15.

Ihde, Don (1990): Technology and the Lifeworld. Indiana: Indiana University Press.

Knoblauch, Hubert/Tuma, René (2011): Videography: an interpretive approach to video-recorded micro-social interaction. In: Margolis, Eric/Pauwels, Luc (Hrsg.): The Sage Handbook of Visual Methods. Thousand Oaks: Sage, S. 414–430.

Kontogiorgos, Dimosthenis/Pereira, Andre/Gustafson, Joakim (2021): Grounding behaviours with conversational interfaces: effects of embodiment and failures. In: Journal of Multimodal User Interfaces, 15, S. 239–254. https://doi.org/10.1007/s12193-021-00366-y.

Kopp, Stefan/Krämer, Nicole (2021): Revisiting Human-Agent Communication: The Importance of Joint Co-construction and Understanding Mental States. In: Frontiers of Psychology, 12:580955. https://doi.org/10.3389/fpsyg.2021.580955.

Krummheuer, Antonia (2010): Interaktion mit virtuellen Agenten? Zur Aneignung eines ungewohnten Artefakts. Stuttgart: Lucius & Lucius.
Kucherenko, Taras/Nagy, Rajmund//Neff, Michael/Kjellström, Hedvig/Henter, Gustav Eje (2022): Multimodal Analysis of the Predictability of Hand-gesture Properties. In: Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS '22), S. 770–779.
Kucherenko, Taras/Nagy, Rajmund/Jonell, Patrik/Neff, Michael/Kjellström, Hedvig/Henter, Gustav Eje (2021): Speech2Properties2Gestures: Gesture-Property Prediction as a Tool for Generating Representational Gestures from Speech. In: Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents (IVA '21), S. 145–147. https://doi.org/10.1145/3472306.3478333.
LIV (2024): Website. https://liv.tv (28.08.2024).
Lowe, Ryan/Noseworthy, Michael/Serban, Iulian V./Angelard-Gontier, Nicolas/Bengio, Yoshua/Pineau, Joelle (2017): Towards an Automatic Turing Test: Learning to Evaluate Dialogue Responses. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Volume 1, S. 1116–1126.
Lugrin, Birgit/Pelachaud, Catherine/Traum, David (2021): The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition. New York: ACM Books.
Lugrin, Birgit/Pelachaud, Catherine/Traum, David (2022): The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application. New York: ACM Books.
Luhmann, Niklas (1984): Soziale Systeme. Grundriß einer allgemeinen Theorie sozialer Systeme. Frankfurt am Main: Suhrkamp.
Magnenat Thalmann, Nadia/Yuan, Junsong/Thalmann, Daniel/You, Bum-Jae (2016): Context Aware Human-Robot and Human-Agent Interaction. Singapore: Springer.
Mara, Martina/Appel, Markus/Gnambs, Timo (2022): Human-Like Robots and the Uncanny Valley. A Meta-Analysis of User Responses Based on the Godspeed Scales. In: Zeitschrift für Psychologie, 230(1), S. 33–46. https://doi.org/10.1027/2151-2604/a000486.
Mayer, Henning/Muhle, Florian/Bock, Indra (2020): Whiteboxing MAX. Zur äußeren und inneren Interaktionsarchitektur eines virtuellen Agenten. In: Geitz, Echkard/Vater, Christian/Zimmer-Merkle, Silke (Hrsg.): Black Boxes – Versiegelungskontexte und Öffnungsversuche. Berlin, S. 295–322.
Meadan, Hedda/Halle, James W. (2004): Communication Repair and Response Classes. In: The Behavior Analyst Today, 5(3).
Mersch, Dieter/Rey, Anton/Grunwald, Thomas/Sternagel, Jörg/Kegel, Lorena/Loertscher, Miriam Laura (2023): Actor & Avatar. A scientific and artistic catalog. Bielefeld: transcript.
Mondada, Lorenza (2012): Conversation Analysis and Institutional Interaction. In: The Encyclopedia of Applied Linguistics. https://doi.org/10.1002/9781405198431.wbeal0202.
Moon, Seungwhan/Kottur, Satwik/Crook, Paul A./De, Ankita/Poddar, Shivani/Levin, Theodore/Whitney, David/Difranco, Daniel/Beirami, Ahmad/Cho, Eunjoon/Subba, Rajen/Geramifard, Alborz (2020): Situated and Interactive Multimodal Conversations. In: Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, S. 1103–1121. https://doi.org/10.18653/v1/2020.coling-main.96.
Mori, Masahiro/MacDorman, Karl/Kageki, Norri (2012): The Uncanny Valley. In: IEEE Robotics & Automation Magazine, 19, S. 98–100.
Pauw, Lisanne/Sauter, Disa/van Kleef, Gerben/Lucas, Gale M./Gratch, Jonathan/Fischer, Agneta H. (2022): The avatar will see you now: Support from a virtual human provides socio-emotional benefits. In: Computers in Human Behavior, 136, 107368. https://doi.org/10.1016/j.chb.2022.107368.
Saberi, Maryam/DiPaola, Steve/Bernardet, Ulysses (2021): Expressing Personality Through Nonverbal Behaviour in Real-Time Interaction. In: Frontiers of Psychology, 12:660895. https://doi.org/10.3389/fpsyg.2021.660895.
Sacks, Harvey (1984): Notes on Methodology. In: Atkinson, J. Maxwell/Heritage, John (Hrsg.): Structures of Social Actions. Studies in Conversation Analysis, Cambridge: Cambridge University Press, S. 21–27.
Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A Simplest Systematics for the Organization of Turn-Taking for Conversation. In: Language, 50(4), S. 696–735.

Schetsche, Michael (Hrsg.) (2004): Der maximal Fremde. Begegnungen mit dem Nichtmenschlichen und die Grenzen des Verstehens. Würzburg: Ergon.
Schmitt, Reinhold (2015): Positionspapier: Multimodale Interaktionsanalyse. In: Dausendschön-Gay, Ulrich/Gülich, Elisabeth/Krafft, Ulrich (Hrsg.): Ko-Konstruktionen in der Interaktion: Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen. Bielefeld: transcript, S. 43–51.
Sidnell, Jack (2010): Conversation analysis. An introduction. Chichester: Wiley.
SoulMachines (2024): Website. https://www.soulmachines.com (28.08.2024).
Suchman, Lucy (1987): Plans and situated actions: The problem of human-machine communication. Cambridge: Cambridge University Press.
Suchman, Lucy (2007): Human-Machine Reconfigurations. Plans and Situated Actions. Cambridge: Cambridge University Press.
Thaler, Markus/Schlögl, Stephan/Groth, Aleksander (2020): Agent vs. Avatar: Comparing Embodied Conversational Agents Concerning Characteristics of the Uncanny Valley. In: IEEE International Conference on Human-Machine Systems (ICHMS).
Tuma, René/Schnettler, Bernt/Knoblauch, Hubert (2013): Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: VS Verlag für Sozialwissenschaften.
Turing, Alan M. (1950): Computing Machinery and Intelligence. In: Mind, 59, S. 433–460.
Van Pinxteren, Beavin M./Pluymaekers, Mark/Lemmink, Jos (2020): Human-like communication in conversational agents: a literature review and research agenda. In: Journal of Service Management, 31(2), S. 203–225. https://doi.org/10.1108/JOSM-06-2019-0175.
Watzlawick, Paul/Beavin, Janet B./Jackson, Don (1967): Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes. New York: W. W. Norton & Company.
Weiss, Benjamin/Wechsung, Ina/Kühnel, Christine/Möller, Sebastian (2015): Evaluating embodied conversational agents in multimodal interfaces. In: Computational Cognitive Science, 1, 6. https://doi.org/10.1186/s40469-015-0006-9.

IV. Nutzungsformen: Forschungsvideos im wissenschaftlichen Film und Forschungsdatennachnutzung

Wie Videoanalyse und soziologischer Film voneinander lernen können

Versuch einer theoretischen Grundlegung visualisierter Kommunikation in der Soziologie

Katharina Miko-Schefzig

Zusammenfassung: Dieser Artikel untersucht, wie Videoanalyse und soziologischer Film voneinander lernen können und welche Potenziale in der Kombination beider Ansätze liegen. Dabei steht die Frage im Mittelpunkt, wie visuelle Methoden zur Wissensproduktion in der Soziologie genutzt werden können. Durch die Gegenüberstellung von Videoanalyse und soziologischem Film wird deutlich, dass beide Ansätze nicht nur zur Datenerhebung und -analyse beitragen, sondern auch die soziologische Kommunikation erweitern können. Der soziologische Film wird dabei als eigenständige kommunikative Gattung vorgeschlagen, die visuelle Mittel zur Wissensproduktion einsetzt. Diese Form des Films schafft Möglichkeiten für Reflexivität und eröffnet neue Wege der Analyse sozialer Phänomene, die über die Grenzen sprachlicher Darstellungen hinausgehen. Zudem wird aufgezeigt, wie filmische Mittel zur Datenproduktion und zur reflexiven Auseinandersetzung mit sozialen Situationen genutzt werden können.

Schlüsselwörter: soziologischer Film, Videoanalyse, visuelle Wissensproduktion, Reflexivität, Kamera

1. Einleitung

Die Forschung zu Visualität hat in den letzten zwei Dekaden in den Sozialwissenschaften allgemein, und in der Soziologie im Speziellen, erheblich an Bedeutung gewonnen. Die behandelten Themen reichen von der theoretischen Fundierung von Bildern und Visuellem über die Erhebung visueller Daten bis hin zur Analyse von Bildmaterial und deren theoretischen und empirischen Implikationen für die Disziplin (Breckner 2012; Boxenbaum et al. 2018; Höllerer/Jancsary/Grafstrom 2018; Knoblauch et al. 2008; Linstead 2018; Raab 2008; Schnettler/Pötzsch 2007; Tuma 2018). Dieser Beitrag beleuchtet einen weniger häufig diskutierten Aspekt: die visuelle Wissens*produktion* und visuelle Kommunikation innerhalb der Soziologie selbst und ihr Verhältnis zu etablierten

Formen der Videoanalyse[1] (z. B. Knoblauch/Tuma/Schnettler 2015; Schindler/Liegl 2013).

Anhand von empirischem Material wird diskutiert, wie die beiden Ansätze der visuellen Soziologie – nämlich der *soziologische Film* (Miko-Schefzig et al. 2022) und die *Videographie* (Knoblauch et al. 2008; Tuma 2018) – miteinander verknüpft werden können und was durch eine solche Kontrastierung über die beiden Ansätze gelernt werden kann.[2] Was lässt sich mit der Videoanalyse zeigen und was mit dem soziologischen Film? Wo liegen die Grenzen der beiden Ansätze und welche Effekte ergeben sich aus dieser Art der Gegenüberstellung? Lassen sich neue Synergien entwickeln oder befördert diese Gegenüberstellung eine »Entmystifizierung« der einst durch den *visual turn* (Breckner/Raab 2016) geweckten Hoffnungen? Oder können umgekehrt die Grenzen der Verfahren produktiv genutzt werden?

Nur wenige Studien (Miko-Schefzig/Learmonth/McMurray 2022; Wood/Salovaara/Marti 2018; Toraldo et al. 2018; Pink 2015; Mohn 2002) haben untersucht, wie die Soziologie visuelle Kommunikation, meist im Zusammenhang mit dem sozialwissenschaftlich informierten Filmemachen (Elliker/Reichle 2024) nutzen könnte. Dieses vergleichsweise geringe Interesse verwundert vor allem vor dem Hintergrund der langjährigen Produktion von Videomaterial, das zu Analysezwecken verwendet wird. Bei dieser Produktion werden implizit visuelle Wissensbestände einbezogen, deren Reflexion jedoch unterbeleuchtet bleibt. Dies betrifft nicht unbedingt die Verwendung in der Analyse (hier hat die Videographie massiv zur theoretischen und methodischen Reflexion beigetragen), sondern vor allem die Aprioris der audiovisuellen Enkulturation unserer eigenen Wissensbestände, die sich in die Aufnahme einschreiben und meines Erachtens in ihrer Wirksamkeit unterbeleuchtet bleiben. Dabei darf nicht vergessen werden, dass mit PowerPoint-Präsentationen selbstverständlich bereits visuelle Wissenschafts*kommunikation* als kommunikative Gattung (Knoblauch 1995; Tuma 2018) etabliert und soziologisch analysiert wurde. Knoblauch (2007) hat diese in ihrer Einbettung in der »sozialen Ökologie« diskutiert, die nicht nur einen visuellen, sondern auch einen performativ-räumlichen Aspekt anzeigt. Wenn ich an anderer Stelle über die Nutzung des soziologischen Films als »Sinnes-Arrangements« (Miko-Schefzig 2023) gesprochen habe, meinte ich genau dieses performativ-räumliche Potenzial.

Insbesondere unter der Annahme, dass die Zunahme visueller Vergesellschaftung zu gesellschaftlichen Transformationsprozessen (Miko 2013) führt,

1 Teile der Argumentation sind bereits in Miko-Schefzig 2013; 2020; 2022; 2023 ausformuliert. Das Verhältnis zur Videographie ist hier erstmals ausformuliert.

2 An anderer Stelle habe ich den soziologischen Film in seinem Verhältnis zur performativen Sozialforschung besprochen (Miko-Schefzig 2023; für die allgemeine Debatte in der Zeitschrift für qualitative Forschung, siehe Mey 2023). In diesem Artikel stelle ich andere, nämliche methodologische Aspekte in den Vordergrund.

stellt sich die Frage nach der Angemessenheit rein sprachbasierter soziologischer Konstruktion und Kommunikation für die Herausforderungen eines zunehmend visuellen Modus der gesellschaftlichen Repräsentation. Dabei folge ich grundsätzlich Eisewicht (2016) in der Einschätzung, dass die visuelle Soziologie, trotz ihrer Beschäftigung mit dem Visuellen, der Sprache dominant verbunden bleibt:

»Dieses Verstehen wird dann oftmals wieder in (wissenschafts-)sprachlich formulierte Ergebnisse übersetzt. Es geht also weniger darum, der Sprache zu ›entkommen‹, als vielmehr darum, andere signifikante Symbole als Sprache bzw. Versprachlichtes zu berücksichtigen und deren Bedeutung für die ›gesellschaftliche Konstruktion der Wirklichkeit‹ (Berger/Luckmann 1969) zu erforschen« (Eisewicht 2016, S. 115 f.). Der soziologische Film adressiert diesen Befund und versucht explizit, sprachlichen Ausdruck zu transzendieren.

Dabei geht es nicht nur darum, »elusive knowledges« (Toraldo et al. 2018) in der Forschung zu erfassen, sondern als Soziologinnen und Soziologen solches Wissen zu produzieren und für sich in Anspruch zu nehmen, dass dieses der Repräsentation sozialer Wirklichkeit in der Wissenschaft näherkommt. Die Kompetenz, gesprochene und geschriebene Sprache zu verwenden, wird dabei interessanterweise als Apriori für kommunikative Aktivitäten innerhalb der wissenschaftlichen Gemeinschaft vorausgesetzt, sei es bei der Datenerhebung, der Analyse und der Berichterstattung. Weniger wird vorausgesetzt, dass wir filmsprachliche Mittel verwenden können, obwohl unsere eigene visuelle Enkulturation in die soziologische Wissensproduktion kontinuierlich einfließt. In diesem Beitrag wird daher untersucht, was wir von visuellen Sprachsystemen (Schnitt, Kamera, Bildauswahl), in denen wir sozialisiert sind (im vorliegenden Beispiel speziell in der Sprache des Films), für die Soziologie lernen können. Es wird erörtert, wie »thick«, im Sinne von Geertz' (1973) »thick descriptions«, bestehende soziologische Analysen und Theorien des Visuellen sein können, da die Kompetenz zur Anwendung visueller, z. B. filmischer Ausdrucksmittel bisher nur in wenigen Fällen dokumentiert ist (Kaczmarek 2008; Kurt 2010; Miko-Schefzig/Learmonth/McMurray 2022). Dabei möchte ich an einem Beispiel explizit machen, was ich mit der filmischen Enkulturation meine: Die Kamera agiert heute nicht mehr als statische Beobachterin aus der Distanz, wie es bis in die 1960er Jahre der Fall war, als sie aufgrund ihrer Größe und der technischen Limitierung, Ton und Bild getrennt aufzuzeichnen, unbeweglich war. Heutzutage ist die Dynamik der Kamera visuell tief in unserer Wahrnehmung verankert, auch wenn wir uns der technischen Entwicklung dahinter nicht immer bewusst sind. In den frühen 1960er Jahren war das anders. Ein Video aus Kennedys Wahlkampf illustriert dies gut: Zum ersten Mal war es damals möglich, Ton und Bild gleichzeitig aufzuzeichnen, was zu einer bewegteren und dynamischeren Bildästhetik führte. Aus heutiger Sicht sind diese Bilder kaum bemerkenswert, aber damals war die Ästhetik der bewegten Kamera revolutionär. Kennedys Wahlerfolg wurde

teilweise auf den geschickten Einsatz dieser Technik und die damit transportierten Attribute – dynamisch, sympathisch, populär – zurückgeführt (siehe[3]: https://www.youtube.com/watch?v=wLkT7z0C5Bc).

Jeder Akt des Sehens ist somit in einen Kontext eingebettet. Soeffner und Raab (2004) haben diese soziale Dimension des Sehens mit dem Konzept der Seh- und Schnitttechnik gefasst, womit kontinuierlich bestimmte Aspekte hervorgehoben und andere ausgeblendet werden. Dieser Schnittprozess wird im soziologischen Film selbst zu einem Analyseinstrument, wie von Miko-Schefzig, Learmonth und McMurray (2022) beschrieben (vgl. Gruber/Rizzolli i. d. B.). Durch die bewusste Auswahl und Bearbeitung von Filmmaterial können neue Erkenntnisse gewonnen werden, die über die reine Beobachtung hinausgehen. Die Verbindung von soziologischer Film- und Videographie bietet daher eine erweiterte Perspektive auf visuelle Daten und eröffnet neue Wege der Wissensgenerierung. Ob sie sich als eigene kommunikative Gattung in der Soziologie durchsetzt, ist noch unentschieden.

2. Tacit Visual Knowledges: Filmsprachliche Sozialisation als Basis wissenschaftlicher visueller Filmproduktion

In der soziologischen Organisationsforschung ist der Beitrag der visuellen Wissensproduktion und -kommunikation zur empirischen und theoretischen Fundierung soziologischer Analyse etabliert (Linstead 2018; Wood/Salovaara/Marti 2018; Miko-Schefzig/Learmonth/McMurray 2022). Besonderes Augenmerk liegt dabei auf soziologischen Filmen, die auf qualitativer, insbesondere ethnografischer Forschung basieren. Da, wie bei der Videographie, der Anfang solcher Filme im Forschen mit der Kamera im Feld liegt, lohnt sich die Parallelschaltung beider Ansätze. Dabei möchte ich hier nicht erörtern, warum Filme einen erkenntnistheoretischen Nutzen haben. Dies wurde in der jüngeren Literatur allgemein und von mir an anderer Stelle bereits diskutiert (z. B. Höllerer/Jancsary/Grafstrom 2018; Miko 2013; Miko-Schefzig 2020; Miko-Schefzig/Learmonth/McMurray 2022). In meiner theoretischen Herleitung beziehe ich mich auf (i) die journalistischen Wurzeln der Chicagoer Schule der Soziologie, (ii) die Visuelle Wissenssoziologie sowie (iii) die Visuelle Anthropologie. Auch in diesem Artikel werde ich zum Teil darauf zurückgreifen. Mehr noch aber möchte ich den logozentrischen Blick der Soziologie selbst thematisieren und die Möglichkeiten, die der soziologische Film hier anzubieten hat.

3 In dieser Dokumentation folgt die Kamera John F. Kennedy in die American Serb Hall in Milwaukee. Die Szene fängt die intensive Bewunderung ein, die Kennedy zu diesem Zeitpunkt seiner Kampagne entgegengebracht wird. Die Kamera begleitet ihn dabei aus der Vogelperspektive, was das Gefühl vermittelt, selbst mitten in der Menge zu stehen.

2.1 Der Umgang mit audiovisuellem Material: Über Nutzungsszenarien und Analysekontexte

Der soziologische, methodisch-interpretative Umgang mit audiovisuellem Material, der über die gängige Dichotomie von Erhebung und Auswertung hinausgeht, ist mein zentrales Anliegen. Soziologischer Film kann dichtes *Zeigen* (Mohn 2013; Miko-Schefzig 2013) ermöglichen. Aber wie kann das gelingen? Und wie kann dieses »dichte Zeigen« in ein Verhältnis zur Videographie gesetzt werden?

Zunächst – und das ist vielleicht schon eine erste Markierung im Verhältnis zur Videographie – ist soziologischer Film kein Abbilden einer vermeintlich abbildbaren Wirklichkeit »da draußen«, sondern ein *Verbildlichen soziologischer Analyse*. Wenn Tuma (2018) über videographische Daten als »›mimetische‹ Abbildung des Geschehens« (ebd., S. 424) schreibt, so ist dem sicher für die sequenzielle Analyse zuzustimmen. Auch der soziologische Film ist selbstverständlich beobachterabhängig, ähnlich wie es der ethnografische Schreibprozess und eben die Videoaufnahmen der Videographie sind. In der Überfülle des Visuellen – darauf hat Mohn (2013, 2022) hingewiesen – herrscht jedoch zunächst Bildlosigkeit. Was ist also der soziologische Beitrag, damit aus Visuellem soziologische Verbildlichung wird? Hier hat die Videographie eine klare Antwort: »Im Unterschied zu Videoaufzeichnungen aus Laborsituationen beziehen sich diese Videos auf soziale Situationen, die von den Handelnden selbst nach ihrer Maßgabe hergestellt werden« (Knoblauch/Vollmer 2018, S. 122). Auf diese Situationen fokussiert auch der soziologische Film. Weitere Analyseschwerpunkte sind jedoch:

1. die Seh- und Schnitttechniken (Soeffner/Raab 2004) der Personen in den Interaktionen, um deren Perspektiven auf die soziale Situation auf visueller Ebene rekonstruierbar zu machen (*Analysekontext*).
2. die Wissensbestände, die analytisch fokussiert wurden, im wörtlichen Sinne sichtbar zu machen, und in einer zunehmend visualisierten und mediatisierten Gesellschaft visuelles Wissen auch visuell zu analysieren (*Analysekontext*).
3. das soziologische Argument mit visuellen Mitteln innerhalb der scientific community kommunikativ verfügbar zu machen und visuell zu verhandeln (*Nutzungsszenario*).
4. Zusätzlich kann der soziologische Film dazu genützt werden, um reflexiv das eigene soziologische Tun zu fokussieren (Knoblauch 2021) (*Analysekontext und Nutzungsszenario*). Slutskaya, Game und Simpson (2018, S. 346) verstehen unter ihrem Konzept kollaborativer ethnografischer Dokumentarfilme ein »untersuchendes« und »reflektierendes« Werkzeug (im Original »investigative«, »reflective«). Im frühen Film *Passing girl: riverside* etwa, unterzieht

der Forschende im Film sein eigenes videographisches Material einer reflexiven Analyse (siehe[4]: https://youtu.be/Pna42AF4gys).

Wenn man die vier Punkte kondensiert, bedeutet dies: Zur Methode wird der soziologische Film dann, wenn er durch Schnitt und Gestaltung plausibel macht, dass ein soziologisches Argument vorgebracht wurde. Mein Argument ist, dass filmische Inszenierung – ein kommunikatives Mittel – ermöglicht, mit den Sehtechniken des Feldes eine dichte Beschreibung zu erzeugen und diese sichtbar zu machen. Gleichzeitig kann dieses Vorgehen auch für einen reflexiven Umgang mit dem soziologischen Material selbst (Knoblauch 2021) genutzt werden.

Um dies zu erreichen, muss man sich zunächst über die filmsprachlichen Aprioris Gedanken machen, mit denen argumentiert wird. Ähnlich wie manche bildanalytischen Verfahren kunsthistorische Bezüge aufweisen (etwa Breckner 2012), kann der soziologische Film auf die Wissensbestände der Filmwissenschaft und der Montagetechnik (Miko 2013; Monaco 1997) zurückgreifen. Durch filmsprachliche Gestaltung wird das soziologische Argument – im wörtlichen Sinne – ersichtlich. Es gibt dabei zunächst keine Unterscheidung zwischen allgemeiner Filmsprache und wissenschaftlicher filmsprachlicher Mittel, die speziell auf den wissenschaftlichen Diskurs zugeschnitten sind.

2.2 Sehen ist Schneiden: Seh- und Schnitttechniken als Basis soziologischer Filme

Jedes Sehen und jedes Schneiden beinhaltet visuelle Wissensbestände. Soziologischer Schnitt (Miko 2013) schafft also zweierlei: Zum einen Reflexion über die soziologische Praxis visuellen Wissens und die Lernmöglichkeiten aus fachfremden visuellen Disziplinen (Schnitttechnik, Dokumentarfilm etc.). Zum anderen kann er gegenstandsbezogen die Sehfertigkeiten und die Schnitttechniken des Feldes dramaturgisch nachvollziehbar und ethnografisch informiert (Miko 2013) visuell zur Disposition stellen. Das heißt, das Ziel soziologisch-ethnografischer Studien wäre dann nicht nur soziologisch-empirisch praxisgeschultes *Sehen* (Schindler/Liegl 2013), sondern auch praxisgeschultes

4 Bei einem Straßenfest in Westafrika bemerkt ein Mädchen erfreut, dass eine Videokamera auf sie gerichtet ist. Ihre anfängliche Begeisterung erlischt jedoch rasch, als sie erkennt, dass der Kameramann bereits das Interesse an ihr verloren hat. Diese Szene wird dazu genützt, um die (Macht-)Beziehung zwischen Forschenden/Filmenden und Forschungssubjekt filmisch zu reflektieren. Im Film folgt die Kamera dem Forscher bis in den Raum, im dem er den Film schneidet. Dadurch wird filmisch visuelle Reflexivität hergestellt.

Schneiden, das soziologische visuelle Kommunikation möglich macht. Die Visuelle Wissenssoziologie hat dabei nicht nur die Methoden zur Analyse von Bildmaterial erweitert, sondern auch herausgestellt, dass Schnitt in der menschlichen Wahrnehmung ein grundlegender und natürlicher Bestandteil des Alltagslebens (und somit auch der Wissenschaft) ist.

> »Wir nehmen die Welt um uns herum nicht als ›solche‹ wahr, sondern ›schneiden‹ sie im Sehen für uns zurecht – nach Perspektive, Standpunkt, Interesse, Sehgewohnheiten und Erinnerungsbildern, die in die gegenwärtige Wahrnehmung eingelagert sind. Unsere Wahrnehmung formt die ›Welt um uns herum‹ zu einer ›Welt für uns‹, die mit Schnittmustern unseres Sehens überzogen und als ein Zusammenhang ausgestattet ist, der allen Details Ort und Sinn kraft unseres positionierenden, deutenden und erinnernden Sehens gibt« (Soeffner/Raab 2004, S. 266).

In seinem Werk zur Visuellen Wissenssoziologie geht Raab (2008) noch weiter und präzisiert nicht nur den Begriff der *Sehtechniken*, sondern entwickelt auch das Konzept der *Schnitttechniken* weiter. Anhand von Beispielen semiprofessioneller Videofilmer:innen, wie etwa Amateurvereinen, zeigt er, wie deren visuelle Wahrnehmungsmuster sich auch in ihren Schnitttechniken widerspiegeln und zur Herausbildung sozialer Normen beitragen. Zusammengefasst lässt sich sagen, dass visuelle Ausdrucksformen kulturell geprägt sind. Sie vermitteln uns nicht nur etwas über soziale Beziehungen, sondern können auch dazu genutzt werden, diese sichtbar zu machen. Letzteres macht sich der soziologische Film zunutze.

Der Kritik an hermeneutischen Verfahren für audiovisuelles Material, dass nämlich die komplexe soziale Situation bei der Analyse sequenziert wird (was, so die Debatte, de facto dem Sozialen nicht gerecht wird), und die prinzipielle Kritik an der Selektivität des Kameraauges bei den Aufnahmen, kann vonseiten des soziologischen Films vielfältig durch Schnittentscheidungen (siehe dazu weiter unten) adressiert werden. Sowohl soziologischer Film als auch Videographie problematisieren diese Selektivität per se jedoch nicht. Der soziologische Film unterscheidet sich von anderen Visualisierungen im Wissenschaftssystem (etwa Powerpoint-Präsentationen, Knoblauch 2007) dadurch, dass mit visuellen Mitteln genuin selbständig zu kommunizieren versucht wird. Er ist nicht bloß Veranschaulichung oder illustrative Beigabe, sondern weist ein hohes Maß an Eigenständigkeit auf (Schnettler/Pötzsch 2007).

Vorsicht ist geboten, wo der soziologische Film in der Form der Dokumentation »ursprüngliche Bedeutung unterdrückt« (Goffman 1980 nach Schindler/Liegl 2013). Hier liegt der Kernpunkt für die Qualitätssicherung im Bereich des soziologischen Films. Er muss beobachterzentriert sein und trotzdem nicht ursprüngliche Bedeutungen unterdrücken. Er nimmt die Seh- und Schnitttechniken auf und ist nicht nur »Verdopplung des Feldes« (Eisewicht/Grenz 2018),

sondern bietet soziologische Einordnung an. Was wie die Quadratur des Kreises wirkt, kann durch die Nutzung von filmsprachlichen Mitteln adressiert werden. Filmsprachliche Mittel (vom Einsatz der Kamera bis zu stilistischen und technischen Schnittentscheidungen) sind mannigfaltig und bieten eine breite Palette an kommunikativen Möglichkeiten (siehe Abschnitt 2.3).

In der Soziologie zeigt sich eine Nähe zwischen soziologischem Film und ethnografischer Forschung (Miko 2013). Ein Grund dafür liegt im Fokus auf »natural situations« (Mohn 2013), ein anderer in der langen Tradition zwischen ethnografischer Forschung und Film in der Kultur- und Sozialanthropologie. Der ethnografische Film ist als kommunikative Gattung in dieser Disziplin etabliert. Dabei teilt der soziologische Film mit der Kameraethnografie eine gewisse Distanz zur Konzeption der natürlichen Situation. Grund dafür ist der Fokus beider Ansätze auf den Schnitt, der per se in die »Natürlichkeit« hineinschneidet. Kameraethnografie (Mohn 2013, 2022) hebt dabei das methodische Potenzial der Kamera hervor und zeigt durch ausgewählte Schnittfolgen, was in den Situationen nicht (sic!) offen*sichtlich* ist. Kameraethnografische Arbeit möchte »Blickschneisen« legen, die erst durch den Schnitt wahrnehmbar werden.

Dieses Anliegen teilt der soziologische Film, geht im Anspruch aber noch einen Schritt weiter. Der soziologische Film will innerhalb der Praxis der Soziologie selbst, also im wissenschaftlichen Diskurs, eine kommunikative Rolle spielen: Er hebt darauf ab, etwas zu zeigen oder zu formulieren, was vorher noch nicht in dieser Form soziologisch gezeigt oder formuliert wurde. Vielleicht kann man es als Stufenprozess konzeptualisieren: Über den eigenen Blick (Mohn 2013) entsteht mittels der videographischen Kamera (Knoblauch/Vollmer 2018) ein »praxisgeschulter Blick« (Schindler/Liegl 2013), der zu einem »gestalteten Blick« (Miko 2013; Miko-Schefzig/Learmonth/McMurray 2022) im soziologischen Diskurs wird.

Die Nachbardisziplin Kultur- und Sozialanthropologie hat für einen feldrelevanten Blick sehr früh das Konzept der »participatory camera« (MacDougall 1995) eingeführt. Die Kamera wird als Artefakt (Froschauer/Lueger 2020b) verstanden, das Teil der Forschung ist und als solches explizit gemacht werden kann. Ziel ist es, ihren analytischen Beitrag zu explizieren und darüber hinaus die Kamera als Apparat des Blickes (Mohn 2013) zu nützen. Mengis, Nicolini und Gorli (2018) beschreiben diesen Aspekt so: »Taken together, we can show that the combinations of camera angle and movement used to collect data (which we call video recording apparatuses) constitute a configuring device, which has a performative effect on the phenomenon of interest and does not simply record it« (ebd., S. 306).

Ein Beispiel dafür möchte ich der Lehrforschung entnehmen: Eine Gruppe Studierender forschte zu Graffitysprayer:innen und übergab die Kamera an die Sprayer:innen selbst, um deren Sehtechnik (während einer strafrechtlich relevanten Handlung) in den empirischen Prozess (und später in den soziologischen

Film[5]: https://www.youtube.com/watch?v=-l5c4MPSirg) einzubringen. Die Kamera half also dabei, einen verborgenen Prozess sichtbar zu machen: Wie organisieren Sprayer:innen die Produktion von – in ihrem Verständnis – »Kunst«, obwohl doch das Gesetz diese Produktion verbietet? Kurt (2010) weist darauf hin, dass, solange man die Kamera nicht aus der Hand gibt, man Gefahr läuft, systematisch an den eigenen Sehgewohnheiten haften zu bleiben. Das Weitergeben der Kamera macht also die Perspektiven und Sehtechniken des Feldes mit der Kamera soziologisch erforsch- und erzählbar. Letzteres, nämlich erzählbar, wird es durch ein dramaturgisches Konzept.

2.3 Die Relevanz filmsprachlicher Mittel für eine soziologische Dramaturgie

Soziologischer Film ohne Analyse ergibt wenig Sinn, aber soziologischer Film, der audiovisuelles Material methodisch einsetzt und dieses dann noch mit filmsprachlichen Mitteln in den soziologischen Diskurs einbringt, arbeitet sowohl auf der analytischen als auch auf der soziologisch-kommunikativen Ebene (Analysekontext und Nutzungsszenario).

Es gibt beim Schnitt nicht eine Storyline (Glaser/Strauss 1967), die erzählt werden muss oder kann, sondern es gibt eine soziologisch fundierte, auf eine Forschungsfrage bezogene, empirisch fundierte Geschichte, die etwas sichtbar macht, das noch nicht im soziologischen Diskurs beinhaltet war. Schnitt ist in diesem Sinne Theoriearbeit, während das Drehen das Erarbeiten von Sehmöglichkeiten darstellt, das darauf angewiesen ist, ethnografische Beziehungen im Feld aufgebaut zu haben – ein Aspekt, den der soziologische Film mit der Videographie teilt. Hassard et al. (2018, S. 1404) betonen den theoretisierenden Beitrag von Schnitt, wenn sie festhalten, dass »editing is core to the work of organisational scholars because, expressive visuals, including selected close-ups, pans, and transitions that add to the affective intensity of the videography are where traditional theorising is deterritorialised and this becomes the theoretical act of expressive videography itself« (Hietanen/Rokka 2018, S. 328).

Um die im letzten Abschnitt vorgeschlagene Reihenfolge nochmals aufzugreifen: Videographie und Interaktionsanalyse können die Basis dafür sein, das soziologische Argument überhaupt erzählen zu können. Umgekehrt kann das Experimentieren mit narrativen Strukturen durch Schnittfolgen, Interaktionsanalysen

5 Dieser Film stammt von einer Studierendengruppe, die die Kamera an Graffiti-Sprayer:innen übergeben hat. Dabei wird deutlich, dass die Kamera aktiv in den Sprayprozess eingebunden ist. Die Aufnahmen wirken angespannt, bewegen sich hin und her, was auch daran liegt, dass die Sprayer:innen ständig darauf achten müssen, ob sie bei ihrem Tun beobachtet werden. Eine Kameraaufnahme zeigt das verdeckte Gesicht eines Sprayers im Spiegelbild der U-Bahn. Hier wird visuell nochmals kommuniziert: Dies ist keine legale Handlung.

mit neuen Impulsen bereichern. Beiden Ansätzen geht es im Aufnahmeverfahren darum, Beobachtbarkeit zu erzeugen (Mohn 2013; Tuma 2018). Im Schneiden jedoch differenziert und fokussiert der Ansatz des soziologischen Films andere Ebenen. Filmsprachliche Mittel – und darin liegt ihre Stärke – »verleihen der Interpretationsfigur des Sozialwissenschaftlers Gestalt« (Kurt 2010, S. 200). Somit ist Filmschnitt kommunizierte Interpretation und Deutung von visuellen Daten.

Mohn (2022) sieht dabei im Schneiden diverse methodische Deutungsleistungen: etwa das Aufzeigen von Vielfältigkeit von Perspektiven oder den Entwurf möglicher Zusammenhänge, ein »zerlegendes Differenzieren« und »montierendes Theoretisieren« (Mohn 2013). Das Kameraauge schneidet dabei Teile aus der sozialen Situation (Schindler/Liegl 2013) heraus (das wird für die Videographie auch von Tuma 2018 betont). Dabei nähern wir uns der Abwesenheit des Sozialen (Hirschauer 2002) im Datenmaterial soziologisch immer nur an – die Frage ist, in welcher Konsequenz wir dies auch visuell tun (nämlich bis zum filmisch-soziologischen Argument, Miko 2013; Miko-Schefzig/Learmonth/McMurray 2022). Diese Frage stößt wiederum die relevante – soziologische – Frage an: Ist die visuelle Vergesellschaftung auch im Wissenschaftssystem angekommen?

Videographie in der Tradition der Konversationsanalyse fokussiert ganz klar auf die Interaktionen vor der Kamera. Im soziologischen Film stehen zusätzlich (i) die Fokussierungsleistungen der Teilnehmer:innen (Schindler/Liegl 2013) im Zentrum. In der Datenproduktion wird ganz explizit an vielfältigen Sehtechniken im Feld gearbeitet. Außerdem geht es (ii) darum, analytische, soziologische Deutungsoptionen der sozialen Welt und ihrer Interaktionsordnungen herzustellen. Soziologischer Film kann hier eine Lücke schließen, da das Weitergeben der Kamera an das Feld und der Einsatz von filmsprachlichen Mitteln die feldspezifischen Fokussierungsleistungen »intersubjektiv nachvollziehbar« (u. a. Froschauer/Lueger 2020a) visualisieren können. Die Achse Wissen – Sehen – Filmen – dichtes Zeigen im soziologischen Film kann eine Form gegenstandsbasierter Theoriebildung und Repräsentation sein.

Der soziologische Film ist nicht nur eine Verbalisierungsressource – er ist kommunikative Visualisierung. Er funktioniert dabei wie ein Protokoll, an dem bis zum Endbericht gearbeitet wird. Er ist keine technische Konserve, sondern an ihm wird analysiert, »umgeschrieben«, d. h. umgeschnitten, und letztendlich zur Disposition gestellt. Soziologischer Film verengt nicht auf einen Blick, sondern nutzt filmsprachliche Mittel, etwa die Parallelmontage (Miko-Schefzig/Learmonth/McMurray 2022), um zu zeigen, »es kann so oder so sein«. Soziologischer Film kann gemeinsames soziologisches Sehen ermöglichen. Er ist Artefakt (Froschauer/Lueger 2020b) und Wissensbestand soziologischen Schaffens und Analysierens.

In Miko-Schefzig, Learmonth und McMurray (2022) habe ich vier zentrale Merkmale des soziologischen Films dargestellt: (i) die konzeptuelle Bedeutung

der Kamera, (ii) das Performen vor der Kamera in der soziologischen Forschungssituation, (iii) die Neuinszenierung eines Lebensereignisses (Re-staging a life event) und (iv) Schnitt als visuelle Forschungsmethode (Miko-Schefzig et al. 2022, S. 656). Diese vier Merkmale können z. B. mit folgenden methodischen Strategien umgesetzt werden:

- Die *Parallelmontage* als Möglichkeit des »Theoretical Sampling« und der Darstellung von Multiperspektivität: »What film can do (and what human vision cannot do) is – through techniques of montage – to juxtapose its two-dimensional pieces and combine them into multispatial and multitemporal viewing experiences. In this way, montage offers the possibility of breaking the boundaries of the ethnographically ›thin‹ 2-D by delivering views of a multidimensional ›thick‹ and, if you like, super-real quality« (Suhr/Willerslev 2012).
- Die *filmische Verdichtung zu Idealtypen*, um spezifische Aspekte eines sozialen Phänomens hervorzuheben (Kurt 2010).
- Das *Weitergeben der Kamera* als Methode zur Sichtbarmachung unterschiedlicher Sehtechniken des Feldes.
- Das *Zusammenschneiden des diskursiven Rahmens*, in dem die konkrete Praxis stattfindet, etwa durch die Einbindung historischen filmischen Materials. Ziel ist hier die Kommunikation einer visuellen Diskursanalyse (Traue 2013).
- *Der zeitliche und räumliche Sprung* im Schnitt, um zu zeigen, dass eine Situation keinen klaren Anfang und kein klares Ende hat, sondern geschnitten, typisiert, analysiert und aus verschiedenen Perspektiven gezeigt wird. Hier ist wichtig zu erwähnen, dass Robert J. Flahertys (1922) Film *Nanook of the North* bereits früh zur Debatte geführt hat, inwieweit Schnitt und Fiktion der Wirklichkeit der Inuit entsprechen (siehe dazu Miko-Schefzig/Learmonth/McMurray 2022)

Im folgenden Fallbeispiel möchte ich ein weiteres soziologisch-filmisches Werkzeug genauer besprechen, das ich als »Re-staging a life event« (Miko-Schefzig/Learmonth/McMurray 2022) beschrieben habe. Ich argumentiere dabei, dass diese Methode noch auf einen weiteren Aspekt soziologischer Filme hinweist, nämlich auf ihr Potenzial als Sinnes-Arrangements. Das ist in eine ähnliche Richtung gedacht wie Knoblauchs (2007) Analyse von PowerPoint-Präsentationen als soziale Ökologie und Performanz. Es geht dabei um die Tatsache, dass sowohl die Herstellung als auch die Präsentation von soziologischen Filmen in ihrem räumlichen, materialen und performativen Kontext zu bewerten sind. Diese Argumentation verdeutlicht, dass filmische Darstellungen in soziologischen Analysen weit mehr sind als lediglich audiovisuelle Produkte. Sie erzeugen ein Sinnes-Arrangement, das soziale Phänomene auf eine Weise erfassbar macht, die über das hinausgeht, was etwa ein wissenschaftlicher Artikel leisten könnte. Der Begriff des Sinnes-Arrangements wurde in Anlehnung an das Konzept

des »Artefakt-Arrangements« (Lueger/Froschauer 2020b) entwickelt. Wood und Brown (2011, S. 521) betonen das Potenzial von analytischen Einsichten durch Affekte, wenn die *kommunikative Wirkung* solcher Erkenntnisse in den Fokus gerückt wird. Das Ziel sei, »to create a filmic affect, one in which the film is an experience of its own right«.

3. Soziologischer Film als Sinnes-Arrangement: Ein analytisches Fallbeispiel und seine Kommunikation

Ein Beispiel eines soziologischen Films, der auf ethnografischer Forschung basiert, ist *Warme Gefühle – vier Liebesgeschichten aus Wien*.[6] Der Film thematisiert die Erfahrungen von homosexuellen Personen vor 1968 und nutzt filmische Mittel, um diese komplexen sozialen Konstellationen darzustellen. Durch die Kombination aus historischen Materialien und persönlichen Erzählungen zeigt der Film eine visuelle Perspektive auf die gesellschaftlichen und kulturellen Herausforderungen, denen sich diese Personen in der Vergangenheit stellen mussten.

3.1 Re-staging a life event

Ich habe an anderer Stelle den soziologischen Film als »Erlebnis- und Erfahrungspotenzial von Sinnes-Arrangements« (Miko-Schefzig 2023) konzeptualisiert. In meiner Arbeit mit Film habe ich ein solches Arrangement im Filmelement »Re-staging a life event« beschrieben (Miko-Schefzig/Learmonth/McMurray 2022), als ich die Protagonist:innen von *Warme Gefühle* zu wichtigen Ereignissen in ihrem Leben zurückgehen ließ und diesen Prozess filmisch begleitete. Dabei vereinbarten wir mit unseren Forschungspartner:innen, dass sie ein zentrales Event ihre Homosexualität betreffend auswählen sollten. Ähnlich plädieren auch Hassard et al. (2018, S. 1411) für »a more nuanced sense of affect and embodiment in video-based research«, indem sie in Anlehnung an Goffman (1959) zwischen »the dramaturgical, the phenomenological, the semiotic and the narrative body« unterscheiden. Für diesen Kontext ist das Konzept des dramaturgischen Körpers hilfreich. Der dramaturgische Körper ist kein passiver Körper, der in einem sozialwissenschaftlichen Film beobachtet werden kann, er ist »embedded in social practices« (ebd., siehe auch Miko-Schefzig 2023).

Soziologische Filme können als eine Möglichkeit verstanden werden, diese vielfältigen Einbettungen zu rekonstruieren und zu zeigen. Streecks (2017) Analogie zwischen »looking and seeing« und »pointing and showing«, mit der auch Bina Mohn (2022) in ihrer Kameraethnografie arbeitet, ist hier besonders

6 https://www.geyrhalterfilm.com/warme_gefuehle (10.12.2024).

aufschlussreich. Friedemann etwa, einer der Protagonisten in meinem Film, steht hinter einem Zaun auf dem Gefängnisgelände, in dem er unter dem Totalverbot von Homosexualität Anfang der 1970er Jahre inhaftiert war. Während der Dreharbeiten sagte Friedemann, wir sollten die Kamera zuerst auf den Zaun und später auf die Felder um den Zaun und das Gefängnis herum richten (»pointing«). Anfangs sahen wir nicht, was er meinte bzw. was er gesehen hat (»seeing«), aber dann fügte Friedemann hinzu: »Oh, ihr müsst zugeben, von hier aus gesehen ist es schön. Schön ist es da!« Die Kamera als methodisches Werkzeug wird also benutzt, um sowohl nach innen (auf das Gefängnis) als auch nach außen (auf die Felder dahinter) zu schauen (»looking«), während Friedemann uns einlädt, die Welt aus seiner Perspektive zu verschiedenen Zeiten und an verschiedenen Orten zu sehen (»seeing«) (siehe[7]: https://www.youtube.com/watch?v=b_9-yVjtYXs).

Das Filmen wird somit »a part of the unfolding of the world, not a descriptive act that records, but a generative act that produces actualities in-the-making« (Hietanen/Rokka 2018, S. 324). Die Betrachter:innen können buchstäblich sehen, dass Friedemann einen neuen (Kamera-)Blickwinkel auf sein Leben gefunden hat. Während er auf die Felder zeigt, schwenkt die Kamera in dieselbe Richtung. Friedemann lädt uns ein, zu sehen, wie es war, als schwuler Mann hinter einem Zaun festgehalten zu werden und auf die relative Schönheit dahinter zu blicken. Er zeigt (»showing«) uns, wie es war, zu einer anderen Zeit in der Geschichte schwul zu sein, indem er die Szene aus dem Hier und Jetzt betrachtet.

So gesehen kann die Kamera, die bei der Produktion eines soziologischen Films eingesetzt wird, eine wesentliche Rolle bei der soziologischen Arbeit spielen. Wie Sandercock und Attili (2012, S. 164) schlussfolgern, produziert der Akt des Filmemachens theoretisch gesehen nicht nur ein visuelles Produkt, »but also the space in which this interaction can take place«. Dieses letzte Argument halte ich für zentral: Räume für soziologische Datenproduktion (Erhebungskontext) und deren Analyse zu eröffnen. Mit anderen Worten: Dieses affektive und körperliche Spiel mit der Kamera (sowie allem, was zum Filmen dazugehört) rief bei der Arbeit mit den Protagonist:innen von *Warme Gefühle* eine Art von Erinnerungen und Körperwissen hervor, die in einem Interview kaum zum Vorschein gekommen wären. Das Ergebnis ist »organisational research in which filmmakers, social theorists, participants and viewers alike are brought together in the same analytical space« (Hassard et al. 2018, S. 1417).

Zusammenfassend lässt sich sagen, dass »Re-staging a life event« nicht einfach nur bedeutet, mit einem oder einer bestimmten Protagonist:in an einen bestimmten Ort zurückzukehren. Es handelt sich vielmehr um ein »actively

7 Der Protagonist Friedemann steht vor dem Gefängnis, in dem er Anfang der 1970er Jahre unter dem Totalverbot für Homosexualität in Österreich inhaftiert war. Die Kamera schwenkt von ihm und seiner Erzählung in Richtung Gefängniszaun. Dabei hört man seine Deutung des Ortes von damals und im Hier und Jetzt. Gleichzeitig folgt man sowohl ihm und der Erzählung als auch seinem Blick und dem von ihm wahrgenommen Umfeld.

presenting« des Körpers, der durch die Interaktion mit der Kamera zu einem dramaturgischen Körper wird. Die Kamera hebt das allgemeine Prinzip dieses dramaturgischen Körpers hervor und verstärkt es. Die Neuinszenierung der sozialen Situation bringt dabei nicht-diskursives und vorkonzeptuelles Körper- und Affektwissen an die Oberfläche – wie z. B. die Frage: Wie habe ich mich damals gefühlt? Wie fühle ich mich heute? Diese Art von Wissen kann nicht durch klassische Methoden wie ethnografische Beobachtungen oder Befragungen erfasst werden. Sie entsteht durch die Art und Weise, wie die Situation und der Raum durch den filmischen Prozess als eine Form der Ko-Konstruktion und Datenproduktion geschaffen wurden (vgl. Sandercock/Attili 2012). Film wird somit zu einer Form soziologischer Forschung, »if it is capable of […] enacting meaning through gesture, body language and other sensory information rather than being an indexical sign pointing at what goes on in front of the camera« (Wood/Salovaara/Marti 2018, S. 830).

3.2 Soziologischer Film als Herausforderung für soziologische Debattenräume (Nutzungsszenarien)

Zuletzt möchte ich mich noch den Nutzungsszenarien und Rezeptionskontexten soziologischer Filme zuwenden. Wie sollen soziologische Filme in der Soziologie genutzt, wie rezipiert, wie auf sie Bezug genommen werden? Zunächst hilft es vielleicht empirisch zu klären, welche Räume für soziologische Filme bereits eröffnet sind. Prominent ist hier auf das Harvard Ethnography Lab (SEL) zu verweisen, das bereits eine Institutionalisierung visueller Wissenschaftskommunikation ist. Das SEL ist – laut Eigendefinition – ein »experimental laboratory at Harvard University that promotes innovative combinations of aesthetics and ethnography«, und die Werke, »produced through SEL in film, video, photography, phonography, and installation has been presented in universities and academic conferences across the world« (SEL 2019). Das SEL verbindet Kunst und Forschung und strebt an, etwas Neues zu schaffen, das »is itself constitutively visual or acoustic – conducted through audiovisual media rather than purely verbal sign systems – and which may thus complement the human sciences' and humanities' almost exclusive reliance on the written word and quantification« (SEL 2019). Das SEL ist in seiner (bisherigen) Wirkung mit Sicherheit nicht mit etablierten Journals zu vergleichen, aber eine Idee ist skizziert: Ein universitärer Forschungsschwerpunkt (wobei disziplinenübergreifend und nicht ausschließlich der Soziologie zugeordnet) mit einem eigenen Journal, dem *Sensate Journal*. Die Mission der Zeitschrift besteht eindeutig darin, Beiträge zu veröffentlichen, die über das geschriebene Wort hinausgehen: »Our mission is to provide a scholarly and artistic forum for experiments in critical media practices that expand academic discourse by taking us beyond the margins of the printed page«

(Sensate Journal 2019). Jeder Beitrag durchläuft ein internes und ein externes Peer-Review-Verfahren im Doppelblindverfahren.

Am Institut für Soziologie der Universität Wien haben wir 2013 ein Sinnes-Arrangement auf der Ebene von wissenschaftlicher Debatte und Kommunikation versucht (siehe Abb. 1). Während dieser Veranstaltung präsentierten wir sowohl Filme von Studierenden als auch die Arbeit *Warme Gefühle*. Bina Elisabeth Mohn hielt einen Vortrag über Blickschneisen und zeigte einen ihrer eigenen Filme. Die Pausen zwischen den Vorführungen wurden an der Kinobar verbracht, und nach jedem Film fand eine Diskussion statt. In einem anderen Jahr wurden am Institut die Filme auf den Fluren bzw. im Stiegenhaus abgespielt. Man konnte also von Station zu Station wandern und sich die Arbeiten ansehen.

Abbildung 1: Poster zur Filmschau des soziologischen Films

In all diesen Beispielen wird etwas klar: Der soziologische Film bzw. jene, die mit ihm arbeiten, suchen nach der passenden Einbettung in die Soziologie als Fach. Diese wird in unterschiedlichen Bereichen adressiert, die – passend zu dem Rahmenthemen dieses Publikationsprojektes – sowohl in den Analysekontexten als auch den Nutzungsszenarios liegen.

4. Fazit und Ausblick

Der soziologische Film stellt eine Erweiterung traditioneller Methoden der und kommunikativer Handlungen in der Soziologie dar, indem er eine genuin visuelle Dimension in die Wissensproduktion einführt. Diese visuelle Dimension ermöglicht es, soziale Phänomene auf eine Art und Weise darzustellen und zu analysieren, die mit textbasierten Darstellungsmöglichkeiten allein nicht erreichbar ist. Der soziologische Film bietet die Möglichkeit, komplexe soziale Wirklichkeiten

nicht nur zu dokumentieren, sondern auch zu interpretieren und zu analysieren, indem er visuelle und filmische Mittel nutzt, um soziale Prozesse und Phänomene zu untersuchen und zu veranschaulichen.

Die Kombination von soziologischem Film und Videographie eröffnet neue Wege der Wissensgenerierung, die die Stärken beider Ansätze vereinen könnten. Während die Videographie detaillierte Einblicke in soziale Interaktionen bietet, ermöglicht der soziologische Film eine umfassendere und multiperspektivische Darstellung sozialer Prozesse. Diese Synergie zwischen den beiden Ansätzen kann dazu beitragen, ein tieferes Verständnis sozialer Dynamiken zu entwickeln und innovative methodische Ansätze in der soziologischen Forschung zu fördern. Ich habe dabei auch vorgeschlagen, das Verhältnis von Videographie und soziologischem Film stufenartig und komplementär zu denken.

All dies zusammennehmend, ist anzudenken, den soziologischen Film als »kommunikative Gattung« für die Soziologie zu fassen, »in deren Rahmen die analysierenden Interaktionssequenzen stattfinden (Knoblauch 1995)« (Tuma 2018, S. 442), und zwar in dreifacher Weise als (i) mittels filmsprachlicher Mittel inszenierte Analyse, (ii) als Sinnes-Arrangement, das Film dafür nutzt, soziologische Daten zu produzieren, die auf der Ebene der Schnitt- und Sehtechniken soziologische Fragestellungen bearbeitbar machen und (iii) als kommunikative Handlungsform, die eine gewichtige Rolle im soziologischen Diskurs einnimmt.

Dabei stellt der soziologische Film auf der Ebene der Analysekontexte soziale Ökologie (Knoblauch 2007, S. 191) zur Verfügung, benötigt diese aber für seine Nutzungsszenarien in der Soziologie noch im großen Maße, soll er eine institutionalisierte Form soziologischer, visueller kommunikativer Handlung werden. Anders formuliert: Die Etablierung des soziologischen Films ist auch deshalb schwierig, weil die soziale Ökologie im Wissenschaftsbereich fehlt. Ob er eine relevante kommunikative Gattung für die Soziologie wird, wird sich an dieser Frage entscheiden.

Literatur

Boxenbaum, Eva/Jones, Candace/Meyer, Renate E./Svejenova, Silviya (2018): Towards an Articulation of the Material and Visual Turn in Organization Studies. In: Organization Studies, 39, H. 5–6, S. 597–616.

Breckner, Roswitha (2012): Bildwahrnehmung – Bildinterpretation. In: Österreichische Zeitschrift für Soziologie, 37, H. 1, S. 143–164.

Breckner, Roswitha/Raab, Jürgen (2016): Materiale Visuelle Soziologie: Einführung in den Themenschwerpunkt. In: Zeitschrift für Qualitative Forschung, 17, H. 1–2, S. 5–9.

Eisewicht, Paul (2016): Die Sicht der Dinge. In: Burzan, Nicole/Hitzler, Ronald/Kirschner, Heiko (Hrsg.): Materiale Analysen. Erlebniswelten. Methodenfragen in Projekten. Wiesbaden: Springer VS, S. 111–128.

Eisewicht, Paul/Grenz, Tilo (2018): Die (Un)Möglichkeit allgemeiner Gütekriterien in der Qualitativen Forschung – Replik auf den Diskussionsanstoß zu »Gütekriterien qualitativer Forschung« von Jörg Strübing, Stefan Hirschauer, Ruth Ayaß, Uwe Krähnke und Thomas Scheffer. In: Zeitschrift für Soziologie, 47, H. 5, S. 364–373.

Elliker, Florian/Reichle, Niklaus (2024): Performative Exploration. Performative Sozialwissenschaft als Dialog und Kollaboration zwischen Kunstschaffenden und Sozialwissenschaftler:innen. In: Zeitschrift für qualitative Forschung, 25, H. 2, S. 207–225.
Froschauer, Ulrike/Lueger, Manfred (2020a): Das qualitative Interview. Zur Praxis interpretativer sozialer Systeme. Wien: WUV.
Froschauer, Ulrike/Lueger, Manfred (2020b). Materiale Organisierung der Gesellschaft: Artefaktanalyse und interpretative Organisationsforschung. Weinheim und Basel: Beltz Juventa.
Glaser, Barney/Strauss, Anselm L. (1967): The discovery of grounded theory. New York: De Gruyter.
Goffman, Erving (1959): The Presentation of Self in Everyday Life. New York: Anchor University Press.
Hassard, John/Burns, Dianne/Hyde, Paula/Burns, John-Paul (2018): A visual turn for organizational ethnography: Embodying the subject in video-based research. In: Organization Studies, 39, H. 10, S. 1403–1424.
Hietanen, Joel/Rokka, Joonas (2018): Companion for the Videography ›Monstrous Organizing – The Dubstep Electronic Music Scene‹. In: Organization, 25, H. 3, S. 320–334.
Hirschauer, Stefan (2002): Grundzüge der Ethnographie und die Grenzen verbaler Daten. In: Schaeffer, Doris/Müller-Mundt, Gabriele (Hrsg.): Qualitative Gesundheits- und Pflegeforschung. Bern: Hans Huber, S. 35–46.
Höllerer, Markus A./Jancsary, Dennis/Grafstrom, Maria (2018): »A picture is worth a thousand words«. Multimodal sensemaking of the global financial crisis. In: Organization Studies, 39, H. 5 + 6, S. 617–644.
Kaczmarek, Jerzy (2008): Soziologischer Film – theoretische und praktische Aspekte. In: Forum: Qualitative Sozialforschung – Social Research. FQS, 9, H. 3, Art. 34. http://www.qualitativeresearch.net/fqs (10.12.2024).
Knoblauch, Hubert (2007): Die Performanz des Wissens: Zeigen und Wissen in Powerpoint-Präsentationen. Konstanz: UVK, S. 117–138.
Knoblauch, Hubert (2021): Reflexive methodology and the empirical theory of science. In: Historical Social Research/Historische Sozialforschung, 46, H. 2, S. 59–79.
Knoblauch, Hubert/Baer, Alejandro/Laurier, Eric/Petschke, Sabine/Schnettler, Bernt (2008): Visual Analysis. New Developments in the Interpretative Analysis of Video and Photography. In: Forum Qualitative Sozialforschung, 9, H. 3, Art. 14. http://dx.doi.org/10.17169/fqs-9.3.1170.
Knoblauch, Hubert/Tuma, René/Schnettler, Bernt (2015): Videography: Introduction to interpretive video analysis of social situations. Frankfurt am Main: Peter Lang.
Knoblauch, Hubert/Vollmer, Theresa (2018): Soziologische Videographie. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 409–422.
Kurt, Ronald (2010): Diener zweier Damen. Videobasierte Sozialforschung zwischen Datendokumentation und Filmproduktion. In: Corsten, Michael/Krug, Melanie/Moritz, Christine (Hrsg.): Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen. Wiesbaden: VS Verlag für Sozialwissenschaften.
Linstead, Stephen (2018): Black Snow: The past lives on [Film].
MacDougall, David (1995): Beyond Observational Camera. In: Hockings, Paul (Hrsg.): Principles of Visual Anthropology. The Hague: De Gruyter, S. 115–121.
Mengis, Jeanne/Nicolini, Davide/Gorli, Mara (2018): The Video Production of Space: How Different Recording Practices Matter. In: Organizational Research Methods, 21, H. 2, S. 288–315.
Mey, Günter (2023): Wissenschaft und Kunst im Dialog? Zum Verhältnis von performativer Sozialwissenschaft und qualitativer Forschung. In: Zeitschrift für Qualitative Forschung, 24, H. 1, S. 73–89.
Miko, Katharina (2013): Visuelles Nosing Around: Zur theoretischen Fundierung visualisierter Wissenschaftskommunikation. In: Soziale Welt. Zeitschrift für sozialwissenschaftliche Forschung und Praxis, 64, H. 4, S. 153–170.
Miko-Schefzig, Katharina (2020): Der sozialwissenschaftliche Film. In: Journal für Psychologie, 28, H. 1, S. 134–155.
Miko-Schefzig, Katharina (2023): Performative Sozialforschung als Sinnes-Arrangement. In: Zeitschrift für Qualitative Forschung, 24, H. 1, S. 106–123.

Miko-Schefzig, Katharina/Learmonth, Mark/McMurray, Robert (2022): A different way of looking at things: The role of social science film in organization studies. In: Organization, 29, H. 4, S. 653–672.
Mohn, Bina E. (2013): Differenzen zeigender Ethnographie: Blickschneisen und Schnittstellen der Kamera-Ethnographie. In: Soziale Welt, 64, S. 171–189.
Mohn, Bina E. (2022): Kamera-Ethnographie: Schauen, Sehen und Wissen filmisch gestalten. In: Geimer, Alexander/Heinze, Carsten/Winter, Rainer (Hrsg.): Handbuch Filmsoziologie. Wiesbaden: Springer VS, S. 591–611.
Monaco, James (1997): Film verstehen. Reinbek bei Hamburg: Rowohlt.
Pink, Sarah (2015): Doing Sensory Ethnography. London.
Raab, Jürgen (2008): Visuelle Wissenssoziologie. Theoretische Konzeption und materiale Analysen. Konstanz: Herbert von Halem.
Sandercock, Leonie/Attili, Giovanni (2012): Unsettling a Settler Society: Film, Phronesis and Collaborative Planning in Small-Town Canada. In: Flyvbjerg, Bent/Landman, Todd/Schram, Sanford (Hrsg.): Real Social Science. Applied Phronesis. Cambridge: Cambridge University Press, S. 137–66.
Schindler, Larissa/Liegl, Michael (2013): Praxisgeschulte Sehfertigkeit: Zur Fundierung audiovisueller Verfahren in der visuellen Soziologie. In: Soziale Welt, 64, H. 1–2, S. 51–67.
Schnettler, Bernt/Plötzsch, Frederik (2007): Visuelles Wissen. In: Schützeichel, Rainer (Hrsg.): Handbuch Wissenssoziologie und Wissensforschung. Konstanz: UVK, S. 472484.
Sensate Journal (2019): About. http://sensatejournal.com/about (10.12.2024).
Sensory Ethnographic Lab (2019): Sensory Ethnographic Lab. https://sel.fas.harvard.edu (10.12.2024).
Slutskaya, Natasha/Game, Annilee M./Simpson, Ruth C. (2018): Better Together: Examining the Role of Collaborative Ethnographic Documentary in Organizational Research. In: Organizational Research Methods, 21, H. 2, S. 341–365.
Soeffner, Hans-Georg/Raab, Jürgen (2004): Sehtechniken. Die Medialisierung des Sehens: Schnitt und Montage als Ästhetisierungsmittel medialer Kommunikation. In: Soeffner, Hans-Georg (Hrsg.): Auslegung des Alltags – der Alltag der Auslegung. Konstanz: UVK, S. 254–284.
Streeck, Jürgen (2017): Self-making man: A day of action, life, and language. Cambridge: Cambridge University Press.
Suhr Christian/Willerslev, Rane (2012): Can Film Show the Invisible? The Work of Montage in Ethnographic Filmmaking. In: Current Anthropology, 53, H. 3, S. 282–301.
Toraldo, Maria L./Islam, Gazi/Mangia, Gianluigi (2018): Modes of knowing. Video research and the problem of elusive knowledges. In: Organizational Research Methods, 21. Jg., H. 2, S. 438–465.
Traue, Boris (2013): Visuelle Diskursanalyse: Ein programmatischer Vorschlag zur Untersuchung von Sicht- und Sagbarkeiten im Medienwandel. In: Zeitschrift für Diskursforschung, 2, S. 117–136.
Tuma, René (2018): Video-Interaktionsanalyse: Zur Feinauswertung von videographisch erhobenen Daten. In: Moritz, Christine/Corsten, Michael (Hrsg.): Handbuch Qualitative Videoanalyse. Wiesbaden: Springer VS, S. 423–444.
Wood, Martin/Brown, Sally (2011): Lines of Flight: Everyday Resistance along England's Backbone. In: Organization, 18, H. 4, S. 517–539.
Wood, Martin/Salovaara, Perttu/Marti, Laurent (2018): Manifesto for Filmmaking as Organisational Research. In: Organization, 25, H. 6, S. 825–835.

Audiovisuelle Forschungsdaten und ihre Kontexte teilen

Archivierung und Nachnutzung von Daten aus der ethnografischen Filmforschung

Martin Gruber und Michaela Rizzolli

Zusammenfassung: In enger Zusammenarbeit mit Forschenden entwickelt das Forschungsdatenzentrum ›Qualiservice‹ an der Universität Bremen geeignete Archivierungslösungen für audiovisuelle Daten der qualitativen Sozialforschung. Dieser Beitrag präsentiert eine Use Study zu den Videodaten des ethnografischen Films »Gbaya. Beekeeping and Honey Hunting« (Gruber 2015) aus der Adamaoua-Region Kameruns. Die Autor*innen stellen die Projektergebnisse vor, erörtern technische und konzeptionelle Herausforderungen sowie Lösungen und benennen grundlegende Anforderungen an die Archivierung und Nachnutzbarkeit ethnografischer Filmdaten. Ein Schwerpunkt liegt dabei auf der Kontextualisierung von Videodaten, wobei insbesondere Verfahren der Annotation und Segmentierung in den Blick genommen werden.

Schlüsselwörter: Ethnografischer Film, audiovisuelle Daten, Forschungsdatenzentren, Archivierung, Nachnutzung

1. Einleitung

Das steigende Interesse an der Archivierung und Nachnutzung von Forschungsdaten[1] sowie die wachsende Bedeutung audiovisueller Medien in allen wissenschaftlichen und gesellschaftlichen Bereichen führen zu einer zunehmenden Nachfrage nach geeigneten Archivierungslösungen für audiovisuelle Daten. Forschungsdatenzentren (FDZ) nehmen dabei eine Schlüsselrolle ein, stehen jedoch vor zahlreichen Herausforderungen: Neben technischen Anforderungen müssen auch datenschutzrechtliche und forschungsethische Aspekte berücksichtigt sowie die Auffindbarkeit und Nachvollziehbarkeit der Daten gewährleistet werden.

1 Die Begriffe Forschungsdaten und Forschungsmaterialien werden im Text synonym verwendet.

Bei Qualiservice[2] – dem FDZ für qualitative sozialwissenschaftliche Forschungsdaten an der Universität Bremen – werden in enger Zusammenarbeit mit Forschenden innovative Ansätze zur Kuration von Videodaten entwickelt und erprobt. Eine exemplarische »Use Study« führten die Autor*innen am Beispiel von Videoaufzeichnungen durch, die für den ethnografischen Film »Gbaya. Beekeeping and Honey Hunting« (Gruber 2015) in der Adamaoua-Region Kameruns entstanden sind. In diesem Beitrag stellen wir die Use Study sowie ausgewählte Ergebnisse der Zusammenarbeit vor. Eingangs erläutern wir den Erhebungskontext des Materials und die methodologische Spezifik des ethnografischen Films.[3] Im Anschluss stellen wir den Rahmen der Use Study vor, erläutern, was kooperative Workflowentwicklung bei Qualiservice bedeutet, und zeigen die besonderen Anforderungen auf, die ethnografische Filmdaten an die Archivierung und Nachnutzung stellen. Abschließend präsentieren wir Lösungswege und diskutieren das Potenzial audiovisueller Daten in verschiedenen Nutzungsszenarien und -kontexten. Unser besonderes Augenmerk liegt auf der Kontextualisierung von Videodaten mit Blick auf ihre Nachnutzung, wobei wir insbesondere Verfahren der Annotation und Segmentierung in den Blick nehmen. Dabei orientieren wir uns an folgenden Leitfragen: Wie kann gutes Forschungsdatenmanagement für die Bereitstellung und Nachnutzung von Videodaten aussehen? Welche technischen, rechtlichen und ethischen Anforderungen müssen erfüllt sein und unter welchen Bedingungen kann audiovisuelles Material für weitere wissenschaftliche Nutzung zugänglich gemacht werden? Auf welche Weise können Annotationen die Auffindbarkeit spezifischer Inhalte in Videodaten verbessern? Wie können ausreichende Kontextinformationen erhalten werden, um die Nachvollziehbarkeit von Videodaten zu gewährleisten?

2. Zum Erhebungskontext und zur Spezifik des ethnografischen Films

Die Verwendung des Mediums Film in der Ethnologie reicht bis zu den Anfängen der Disziplin zurück. Ende des 19. Jahrhunderts, als die Technologie Film selbst erst wenige Jahre alt war, wurden bereits Filmaufnahmen im Rahmen großer Forschungsreisen gemacht. Film wurde von Ethnolog*innen bis weit in die zweite Hälfte des 20. Jahrhunderts als objektives wissenschaftliches Dokumentationsmedium verstanden. Seit den 1960er Jahren setzte sich dann immer mehr die Verwendung von Film im Rahmen einer qualitativen und interpretativen Ethnologie

2 https://www.qualiservice.org/de (14.08.2024).

3 Während im Rahmen ethnografischer Filmforschung heute fast ausnahmslos Video-Technologie verwendet wird, bezeichnet man die zugrundeliegende Methode aufgrund des Entstehungskontexts als *Ethnografischer Film*. Wir verwenden Video und Film synonym.

durch (Henley 2000). Heute verwenden Ethnolog*innen, die mit dem Medium arbeiten, eine Vielzahl filmischer Ansätze. Zentral ist die Idee der teilnehmenden Beobachtung mit der Kamera: Die Beobachtungen und Aushandlungsprozesse während des Filmemachens werden mit der (subjektiven) Kamera festgehalten und tragen zum ethnologischen Wissen bei, das im resultierenden Film vermittelt wird (Henley 2004). Die wichtigsten filmischen Ansätze, die dabei verwendet werden, sind das beobachtende *Observational Cinema* und der teilnehmende *Partizipatorische Film* (MacDougall 2003). Üblich sind lange Einstellungen von Aktivitäten, Gesprächen und verschiedenen Formen von Interviews. Extreme Nahaufnahmen oder Kamerawinkel werden i. d. R. vermieden, um einen unprivilegierten, menschlichen Blick zu ermöglichen (MacDougall 1982).

Ethnografischer Film zielt nicht darauf ab, bereits vorhandene Forschungsergebnisse zu vermitteln, sondern ethnologisches Wissen durch den Prozess des Filmemachens zu generieren: Durch die Interaktionen und Aushandlungsprozesse im Feld genauso wie durch die mit der Kamera festgehaltenen körperlichen Erfahrungen (Pink 2009b). Die Ethnologin oder der Ethnologe führt die verschiedenen Aufgaben der Filmproduktion entweder eigenständig durch oder arbeitet mit einem kleinen Team. Oft werden lokale Forschungspartner*innen oder Vertreter*innen der untersuchten Gruppe am Prozess des Filmemachens beteiligt oder in die Lage versetzt, selbst Filme zu machen (Gruber 2016; Gruber 2022). Der Schnitt wird als wichtiges analytisches Element des Forschungsprozesses verstanden (vgl. Miko-Schefzig i. d. B.). Hierbei werden die Aufnahmen oft in einen Sinnzusammenhang gebracht, der es den Rezipient*innen ermöglicht, die gefilmte Situation bzw. die Erfahrungen der Protagonist*innen nachzuempfinden. Ethnografische Filme vermitteln ethnologisches Wissen häufig auf narrative, subjektive und implizite Weise. Eine nachträgliche Analyse der aufgenommenen Interaktionen oder Handlungen durch die filmende Person oder andere Wissenschaftler*innen, im Sinne der Videographie nach Tuma, Schnettler und Knoblauch (2013), ist beim Ethnografischen Film unüblich.

Obwohl die Forschungsteilnehmer*innen visuell in das Filmmaterial eingeschrieben sind, werden die Protagonist*innen im Ethnografischen Film eher selten anonymisiert, da dies die ästhetischen Qualitäten und den Informationsgehalt des Films erheblich beeinträchtigen könnte. Gleichzeitig müssen bei Video- und Fotoaufnahmen, insbesondere von Menschen, andere ethische Standards berücksichtigt werden als bei nicht-visuellen Daten (Pink 2007). Falls Anonymisierungen erwünscht oder notwendig sind, so erfolgen diese fast nie durch im journalistischen Dokumentarfilm gängige Methoden wie Farbbalken oder Verpixeln, sondern eher durch künstlerisch-kreative Ansätze wie Gegenlichtaufnahmen, Fiktionalisierungen oder Animationen.

Im Zentrum des Beitrags stehen die Videoaufzeichnungen für den ethnografischen Film »Gbaya. Beekeeping and Honey Hunting« (Gruber 2015). Der Film entstand im März 2015 während eines Zeitraums von vier Wochen in der

Adamaoua-Region Kameruns für die Bienenbiologin Dorothea Brückner der Universität Bremen, mit dem Ziel, die traditionelle Honigjagd und Bienenhaltung in dem Gebiet zu dokumentieren. Als Auftragsarbeit stellt der Film ein Beispiel für angewandte Visuelle Anthropologie dar (Pink 2009a). Diese ist im Gegensatz zur akademischen Forschung oft stärker auf eine spezifische Fragestellung oder auf eine Intervention hin ausgerichtet, greift aber auf dieselben Theorien und Methoden zurück und entwickelt diese weiter. Der Film war sowohl für ein bienenwissenschaftliches und ethnologisches Fachpublikum als auch für ein breites Publikum intendiert. Besonders wichtig waren der Auftraggeberin eine möglichst genaue Darstellung imkerlicher Aktivitäten und der handwerklichen Herstellung von Bienenkörben.

Der Film zeigt die Honigjagd und die Imkerei mit traditionellen Bienenstöcken, indem er den erfahrenen Imker Jean Sardi in seinem Alltag begleitet. Er ist als mehrtägige Reise konstruiert, die die Protagonist*innen in das traditionelle Jagdrevier des Imkers unternehmen, in dem er seine Bienen hält. Die Handlung beginnt am frühen Morgen im Gehöft des Imkers am Stadtrand von Ngaoundéré. Sardi und seine Begleiter*innen packen ihre Ausrüstung ein und brechen mit ihren schwer beladenen Motorrädern auf. Auf dem Weg in das entlegene Jagdgebiet hält die Gruppe an, um pflanzliche Rohstoffe für die Herstellung der Bienenstöcke zu ernten. Am Zielort angekommen, schlägt die Gruppe ihr Lager auf. Dann wird die schrittweise Konstruktion einer traditionellen Beute aus Palmwedeln, Gras, Blättern, Lianen und Lehm gezeigt. Nach der Fertigstellung wird der Korb in einem Baum befestigt. Immer wieder erklärt der Imker seine Aktivitäten dem Filmpublikum und geht dabei auf kulturelle Aspekte der Imkerei ein. Dann werden die Ernte von Honigwaben aus einem solchen Bienenstock sowie die Ernte von Honigwaben eines Bienenvolks, das unbetreut in einer Baumhöhle lebt, gezeigt (Honigjagd). Am Ende des Films fährt die Gruppe zurück in ihr Dorf, wo der Honig mit einem einfachen Sieb von dem Wachs getrennt und in Flaschen zum Verkauf abgefüllt wird.

Die hier diskutierte Filmforschung basierte auf der Methode des Ethnografischen Films, als Teildisziplin der Ethnologie, wie oben beschrieben. Die Kamerunischen Bienenbiologen Mazi Sanda und Adamou Moïse von der University of Ngaoundéré unterstützten die Feldforschung sowie die Filmaufnahmen und den Schnitt. Neben der Vermittlung von Kontextwissen und Kontakten war das Übersetzen ein wichtiger Aspekt dieser Zusammenarbeit. Die Protagonist*innen des Films sprachen in ihrer Muttersprache Gbaya und der lokal gebräuchlichen *lingua franca* Fulfulde. Adamou Moïse übersetzte als Gbaya Muttersprachler während der Dreharbeiten und während des Schnitts ins Französische. Mazi Sanda, der nur über elementare Kenntnisse in Gbaya verfügte, kommunizierte mit den Protagonist*innen in Fulfulde. Um die Inhalte der Dialoge einem breiten kamerunischen und internationalen Publikum zugänglich zu machen, wurden Untertitel erstellt, welche im Ethnografischen Film i. d. R. dem Voiceover vorgezogen

werden (siehe Henley 2021). Die Untertitel wurden in Englisch und Französisch, den beiden Amtssprachen Kameruns, verfasst.

Der Film wurde als ethnologische Veröffentlichung auf zahlreichen Filmfestivals einem Fachpublikum zugänglich gemacht. Außerdem wurde er als DVD in Kamerun und international vertrieben und ist über die ethnologische Publikationsplattform EthnOA[4] frei zugänglich (https://doi.org/10.18450/ethnoa-medien/34). Der Film und die zugrundeliegende Forschung wurden zum Ausgangspunkt für Grubers und Sandas mehrjährige kollaborative Forschung zu Honigbienen und Imkerei in Kamerun. Gruber nutzte die Filmforschung für seine ethnografische Studie zu Mensch-Honigbienen Verflechtungen in Kamerun, Deutschland und Japan. Mazi Sanda verwendete die Filmaufnahmen für seine bienenwissenschaftliche Forschung und den Film für die Lehre an der University of Ngaoundéré. Die Filmforschung und das daraus resultierende Material diente direkt oder indirekt auch als Basis für zahlreiche kulturwissenschaftliche, biologische und interdisziplinäre Veröffentlichungen (Gruber/Sanda 2019; Gruber 2018; Gruber 2023; van der Wal et al. 2022; van der Wal et al. 2023; Cram et al. 2022; Mazi/Paxton/Osterman 2022; Mohamadou/Faibawa/Mazi 2022).

Nachdem wir den Erhebungskontext der Videodaten ausführlich skizziert haben, möchten wir im Folgenden den Rahmen beschreiben, in dem die Use Study bei Qualiservice durchgeführt wurde.

3. Zum Kontext der Use Study: Videodaten archivieren bei Qualiservice

Deutschlandweit gibt es verschiedene Informationsinfrastrukturen zur Archivierung und Nachnutzung qualitativer und audiovisueller Forschungsdaten,[5] die jedoch meist auf bestimmte Disziplinen, Themen oder Datensorten spezialisiert sind (Knoblauch/Wilke 2018, S. 50; Wilke/Pröbrock/Pach 2019, S. 471; vgl. auch Wilke i. d. B.).[6] Qualiservice ist das einzige FDZ in Deutschland, das qualitative sozialwissenschaftliche Forschungsmaterialien themen- und disziplinen-

4 EthnOA ist eine Publikationsplattform fur Bild-, Audio- und Videobeiträge aus den ethnologischen Fächern auf dem Medien-Repositorium der Humboldt-Universität zu Berlin. Diese werden mittels DOI dauerhaft verfügbar gemacht und stehen im Sinne von Open Access unter einer Creative Commons Lizenz weltweit frei zur Verfügung.

5 Das Netzwerk QualidataNet, an dem Qualiservice beteiligt ist, bietet einen umfassenden Überblick über Forschungsdatenzentren, Archive und Repositorien in Deutschland, die Forschungsmaterialien der qualitativen Sozialforschung archivieren und zur Nachnutzung bereitstellen. https://www.qualidatanet.com/de (14.08.2024).

6 aviDa z. B. wurde an der TU Berlin für die Archivierung videographischer Daten der empirischen qualitativen Sozialforschung entwickelt (aktuell ohne technischen Betreiber) und Oral-History.Digital an der FU Berlin unterstützt die Erschließung wie Bereitstellung von audiovisuell aufgezeichneten narrativen Interviews.

übergreifend archiviert und für die wissenschaftliche Nachnutzung bereitstellt.[7] Der Datenbestand umfasst digitale Forschungsmaterialien, wie z. B. Beobachtungs- und Gesprächsprotokolle, Interviews, Feldnotizen, Fotos, Filme, Audiodaten sowie eine Vielzahl an Kontextmaterialien wie Interviewleitfäden, Postskripte oder Codebooks. Die bei Qualiservice archivierten Forschungsdaten sind i. d. R. nicht frei im Internet einsehbar, sondern erst nach Rücksprache sowie nach Unterzeichnung einer Nutzungsvereinbarung mit Qualiservice nutzbar.[8]

Seit 2019 ist Qualiservice am Fachinformationsdienst Sozial- und Kulturanthropologie (FID SKA) beteiligt.[9] Im Rahmen des gemeinsam mit der Universitätsbibliothek der Humboldt-Universität zu Berlin verantworteten FID werden die bestehenden Services und Verfahren so erweitert, dass auch ethnografisches Material unter kontrollierten Bedingungen archiviert und nachgenutzt werden kann (Rizzolli/Imeri/Huber 2024). Dies umfasst insbesondere Beobachtungsprotokolle, Feldnotizen und -tagebücher sowie Audio- und Videodaten.[10]

Kennzeichnend für die Archivierung bei Qualiservice ist, dass die Datenvorbereitung als kooperative Aufgabe zwischen Forschenden und FDZ gestaltet wird (Mozygemba/Kretzer 2022). Qualiservice befähigt Forschende, ihre Daten projektbegleitend so aufzubereiten und zu dokumentieren, dass die Daten kuratiert, archiviert und für die Nachnutzung verfügbar gemacht werden können. Das bedeutet auch, dass Forschende, die ihre Daten bei Qualiservice archivieren möchten, maßgeblich mitbestimmen können, welche Daten archiviert und wie sie genutzt werden dürfen. Auch die Entwicklung von Workflows zur Datenkuration erfolgt in Kooperation mit Forschenden im Rahmen sogenannter Use Studies (Heuer/Hollstein/Mozygemba 2021, S. 469), um forschungsnahe Lösungen zu finden und die spezifischen Bedarfe der Disziplinen und Datentypen in die Infrastruktur und Verfahren zu integrieren. Als Use Study umgesetzt werden Projekte, die in den jeweiligen Disziplinen besonders verbreitete Datentypen oder Forschungsdesigns abbilden und daher wertvolle Erkenntnisse für die Weiterentwicklung von Workflows bieten.

---

7 Qualiservice arbeitet eng mit verschiedenen Partnern zusammen: Mit der Staats- und Universitätsbibliothek im Bereich Metadaten-Entwicklung sowie der Veröffentlichung von Studienreports und mit dem Weltdatenarchiv PANGAEA bei der Metadatenverzeichnung und -publikation. GESIS – Leibniz-Institut für Sozialwissenschaften ist Partner bei der verteilten Archivierung von Mixed Methods Studies.

8 Je nach Beurteilung der Sensibilität können Forschungsmaterialien entweder extern heruntergeladen oder – wie dies bei Videodaten der Fall ist – vor Ort unter kontrollierten Bedingungen konsultiert, grundsätzlich aber ausschließlich für wissenschaftliche Zwecke genutzt werden.

9 https://www.evifa.de/de (14.08.2024).

10 Zusätzlich zum FID SKA kooperiert Qualiservice mit weiteren Fachinformationsdiensten (darunter Soziologie, Politikwissenschaft und Kriminologie), um Datenkurations-Workflows für spezifische Datentypen zu entwickeln und fachspezifische Beratungsangebote bereitzustellen.

Für die Use Study zu den Videodaten wählten wir das Projekt »Traditional Beekeeping and Honey Hunting in Central Cameroon«, da es idealtypisch für die Methode des Ethnografischen Films ist. Neben dem 38-minütigen Film, der über EthnOA im Sinne von Open Access frei zugänglich ist, umfasst das bei Qualiservice archivierte Material 160 Videodateien mit einer Gesamtlänge von etwa acht Stunden. Während die Dialoge im geschnittenen Film mit französischen und englischen Untertiteln versehen sind, ist das Rohmaterial[11] selbst nicht untertitelt oder übersetzt. Diese Mehrsprachigkeit und der relativ große Umfang der Videodaten machen die Navigierbarkeit des Materials schwierig. Der Film zeigt nur einen Ausschnitt aus dem reichhaltigen Datenkonvolut, das eine Fülle von Aktivitäten und Gesprächen umfasst, und bietet eine Möglichkeit, sich dem Rohmaterial zu nähern und es zu kontextualisieren. Im Rahmen der Use Study wurde das Rohmaterial aufbereitet und für die wissenschaftliche Nachnutzung bereitgestellt (Gruber 2024a).[12] Es eignet sich für vielfältige Nachnutzungsszenarien, einschließlich der Anwendung unterschiedlicher Analysetechniken und -verfahren (siehe dazu auch Abschnitt 5).

Grundsätzlich ist zu beachten, dass Video- und Filmaufnahmen von Menschen gemäß der Datenschutz-Grundverordnung (DSGVO) personenbezogene Daten darstellen, da die abgebildeten Personen identifizierbar sind. Die Verarbeitung, Archivierung und Nachnutzung solcher Daten ist nur zulässig, wenn eine informierte Einwilligung der Forschungsbeteiligten vorliegt oder eine andere Rechtsgrundlage greift.[13] Da zum Zeitpunkt der Produktion die Archivierung des Rohmaterials nicht vorgesehen war, musste vor Beginn der Use Study das Einverständnis aller Protagonist*innen nachträglich für die Archivierung und Sekundärnutzung eingeholt werden.[14] Dies war möglich, da Sanda im Rahmen seiner bienenwissenschaftlichen Forschung noch enge Kontakte zu den Teilnehmenden der Forschung pflegte.

11 Rohmaterial ist ein Begriff aus dem Kontext des Dokumentarfilms und wird hier im Sinne von Video Footage bzw. ungeschnittenem Material verwendet. Wesentlich dabei ist, dass wir Rohmaterial nicht als objektive Fakten, sondern stets als Ergebnis des sozialen Prozesses der Feldforschung verstehen (vgl. Knoblauch/Wilke i. d. B.).

12 https://doi.org/10.1594/PANGAEA.963249.

13 Zum Beispiel im Rahmen der sog. Interessensabwägung (Art. 6 Abs. 1 lit. f. DSGVO).

14 Bei bereits abgeschlossenen Projekten, für die keine Einwilligung zur Archivierung und Sekundärnutzung vorliegt, kann versucht werden, diese nachträglich einzuholen. Der Aufwand – wie sich in unserem Fall zeigte – und das Risiko, dass Beforschte gegebenenfalls nicht mehr verfügbar sind, sollte jedoch nicht unterschätzt werden. Qualiservice stellt eine Handreichung und Vorlagen für die informierte Einwilligung bereit, die an die spezifischen Anforderungen eines Projektes angepasst werden können (vgl. Kretzer et al. 2020).

Während Interviewtranskripte und andere Textdaten oftmals so anonymisierbar sind, dass sie sinnvoll nachgenutzt und als Scientific-Use-Files[15] für externe wissenschaftliche Auswertungen bereitgestellt werden können, stellt die Anonymisierung von Videodaten ein Problem dar, das bestimmte Zugangswege und Nachnutzungsszenarien ausschließt. Da es derzeit keine zufriedenstellende technische Lösung gibt, um Videodaten entsprechend zu anonymisieren oder pseudonymisieren, können solche Daten momentan nur in den Räumlichkeiten von Qualiservice an einem Gastwissenschaftler*innen-Arbeitsplatz (GWAP) ohne Internetanschluss eingesehen werden. Im Rahmen der Use Study wurde daher ein Workflow entwickelt, der es interessierten Forschenden ermöglicht, die Passung des Materials vorab zu beurteilen und auf dieser Basis Aufwand und Nutzen einer Reise nach Bremen abzuwägen.

4. Kooperative Workflowentwicklung: Annotation und Segmentierung von Videodaten

Bisher hatte Qualiservice nur vereinzelt und im Rahmen projektspezifischer Lösungen Videodaten[16] archiviert. Dabei konnten die komplexen Anforderungen im Umgang mit Videodaten nicht immer optimal erfüllt werden, insbesondere was die Navigierbarkeit des Datenmaterials sowie dessen inhaltliche Beurteilung durch an einer Nachnutzung interessierte Wissenschaftler*innen betrifft. Um individuelle Lösungen systematisch in allgemein nutzbare Workflows zu überführen und eine solide, aber flexible Grundlage für die Archivierung von Videodaten zu schaffen, wurde die hier beschriebene Use Study konzipiert. Wir sind uns dabei bewusst, dass ethnografische Forschung aufgrund ihrer Pluralität und Offenheit kaum standardisierbar ist und daher weitgehend flexible Workflows erfordert (Rizzolli/Imeri/Huber 2024; Eberhard 2020). Vor diesem Hintergrund stellen wir im Folgenden ausgewählte, wesentliche Ziele und Ergebnisse der Use Study vor.

Ein zentrales Anliegen war es, die Navigierbarkeit von Videodaten zu erhöhen und somit den Zugang und die Orientierung innerhalb größerer Videobestände

15 Scientific Use Files (SUF) sind anonymisierte Daten, die zur Sekundärnutzung für Forschung bereitgestellt werden. SUFs können nach Kontaktaufnahme mit Qualiservice und Unterzeichnung einer Nutzungsvereinbarung extern heruntergeladen werden.

16 Bei Qualiservice kann Material als Forschungs- oder Kontextmaterial archiviert und für weitere Nutzungen bereitgestellt werden. Ein Beispiel für einen Datensatz, der Videodaten enthält, ist: Weiß, Anja; Quasinowski, Benjamin; Sommer, Ilka (2022): Study Report »Globalizing medical knowledge and practice«. Transcripts, translation, audiovisual and context material for doctor-patient-interaction videoobserved at university hospitals in Ankara (Turkey), Beijing (PRChina), Groningen (Netherlands) and Würzburg (Germany). Bremen: RDC Qualiservice. http://dx.doi.org/10.26092/elib/1395.

oder längerer Videoaufnahmen zu erleichtern. Zudem sollten Videodaten durch die Bereitstellung zusätzlicher Kontextinformationen verständlicher gemacht und damit tiefere Einsichten in das Material ermöglicht werden. Dazu griffen wir auf Ansätze und Formate zurück, die Qualiservice bereits erfolgreich erprobt und implementiert hatte: Metadaten, Studienreport und Mikro-Metadaten (vgl. Tab. 1).

Tabelle 1: Formate der Kontextualisierung bei Qualiservice (in Anlehnung an Rizzolli/Imeri/Huber 2024, S. 9)

Metadaten	stellen strukturierte Informationen zu Forschungsprojekten oder Datensätzen bereit, um sie auffindbar zu machen
Studienreport	beinhaltet wichtige Informationen zum Forschungsprojekt, einschließlich Entstehungs- und Analysekontext, sowie gegebenenfalls Ideen zur Nachnutzung, um einen ersten Einblick ins Material zu ermöglichen
Mikro-Metadaten	liefern detaillierte Informationen zu einzelnen Daten, wie beispielsweise zu einer Videodatei innerhalb eines Datensatzes

Metadaten stellen strukturierte Informationen zum Forschungsprojekt und zum Material bereit. Sie sind erforderlich, damit Forschungsdaten nicht nur bei Qualiservice,[17] sondern auch in übergreifenden Portalen und Suchmaschinen (z. B. BASE, EOSC, QualidataNet) aufgefunden werden können – ähnlich wie Informationen über Publikationen in Bibliothekskatalogen. Die Metadaten zu den ethnografischen Datensätzen werden zudem in der EVIFA-Suche, dem Rechercheportal des FID SKA, nachgewiesen. Der Studienreport enthält zusätzliche und ausführliche Angaben zum Projekt, zum Entstehungs- und Analysekontext sowie zu möglichen Nachnutzungsszenarien.[18] Während Metadaten und der Studienreport einen gesamten Datensatz beschreiben, bieten die Mikro-Metadaten detaillierte Kontextinformationen auf der Ebene einzelner Dateien innerhalb eines Datensatzes. Die Mikro-Metadaten enthalten beispielsweise technische Details zu einzelnen Videodateien oder Hintergrundinformationen zu einzelnen Interviewteilnehmer*innen (z. B. Geschlecht, Alter, Beruf, Bildungsstand, sozio-ökonomischer Hintergrund). Im Unterschied zu den Metadaten und dem Studienreport sind die Mikro-Metadaten nicht öffentlich zugänglich, sondern erst nach vorheriger Online-Registrierung einsehbar.[19]

17 https://www.qualiservice.org/de/qsearch.html (08.08.2024).

18 Qualiservice stellt eine Handreichung (Heuer/Kretzer/Mozygemba/Huber/Hollstein 2020) und eine Vorlage bereit, die Forschende bei der Erstellung des Reports unterstützt. Als Beispiel vgl. den Studienreport, der im Rahmen der Use Study erstellt wurde: https://doi.org/10.26092/elib/3316.

19 Die Registrierung erfolgt derzeit über ein Online-Formular beim Informationssystem PANGAEA – dem IT-Partner von Qualiservice. Anschließend können die Mikro-Metadaten eingesehen werden. https://www.pangaea.de (08.08.2024).

Im Rahmen der Use Study erachten wir den Studienreport und die Metadaten als angemessen für die Beschreibung von Videodaten. Die Mikro-Metadaten hingegen wurden erweitert, mit weiteren Informationen angereichert und an die materialspezifischen Bedarfe von Videodaten angepasst, um deren Nutzbarkeit und Kontextualisierung zu verbessern. Dazu haben wir Verfahren der Annotation und Segmentierung eingesetzt:

- *Annotation*: Hierbei werden Videos mit ergänzenden Informationen (sog. Annotationen) versehen, um den Inhalt zu beschreiben oder bestimmte Aspekte hervorzuheben z. B. Aktivitäten, Ereignisse, Themen oder Orte.
- *Segmentierung*: Bei der Segmentierung werden Videos in kleinere, logisch zusammenhängende Abschnitte unterteilt, indem etwa ein Index mit Haupt- und Zwischenüberschriften erstellt und eine übersichtliche Struktur geschaffen wird. Zudem können Videodateien in Segmente unterteilt und mit Timecodes versehen werden. Dies unterstützt die Navigation und Analyse von langen Videodateien, da relevante Abschnitte gezielt angesteuert und ausgewertet werden können.

Für die Annotation und die Segmentierung einzelner Videodateien nutzten wir die Open Source Software ELAN, die am Max-Planck-Institut für Psycholinguistik im Sprach-Archiv (TLA – The Language Archive) entwickelt wurde.[20] Die Software ist besonders gut für Video- und Audiodateien geeignet. Mit ELAN können beliebig viele Annotationen auf verschiedenen Zeilen erstellt werden, die beispielsweise Themen, Ereignisse oder Orte im Video kennzeichnen, detaillierte Erklärungen umfassen oder den gesprochenen Inhalt wiedergeben. Durch das Markieren von Auszügen können Videos in Segmente unterteilt werden, denen wiederum Annotationen zugeordnet werden können. Außerdem erlaubt ELAN den Export von Annotationen in verschiedene Formate, was die Integration der Ergebnisse in bestehende Verfahren (Erstellung von Metadaten) erst ermöglichte. Abbildung 1 veranschaulicht, wie unterschiedliche Ebenen der Annotation in ELAN umgesetzt werden, sodass sowohl eine Segmentierung der Videoaufnahmen (Titel/Untertitel) als auch eine Anreicherung des Materials durch fachspezifische Keywords und Notizen erreicht werden.

20 ELAN (Version 6.6) [Computer software]. (2023). Nijmegen: Max-Planck-Institute for Psycholinguistics, The Language Archive. https://archive.mpi.nl/tla/elan (08.08.2024).

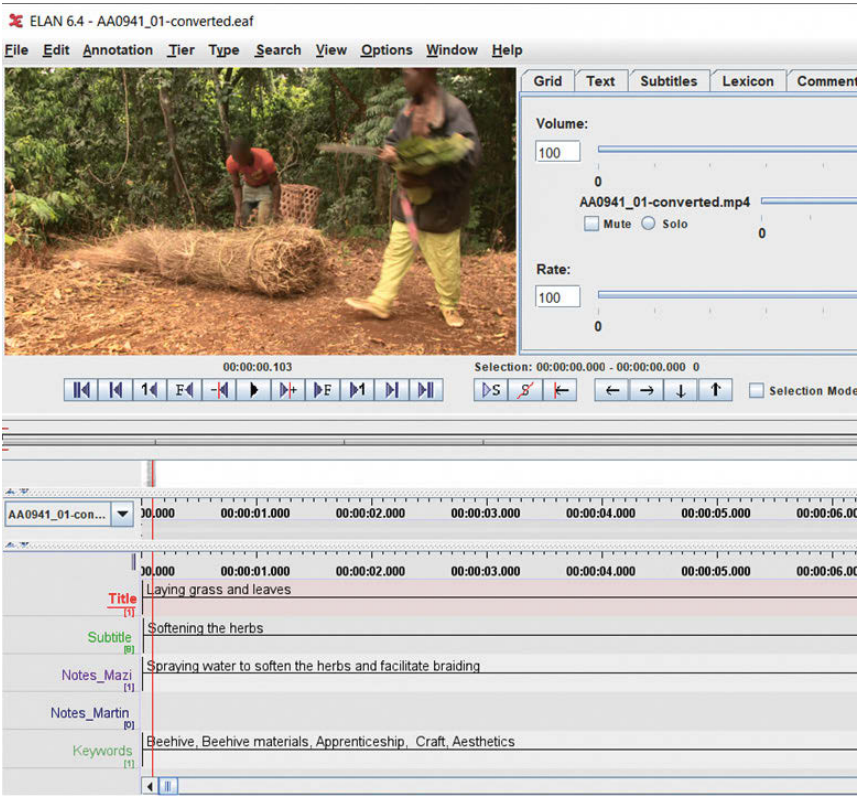


Abbildung 1: Annotation mit ELAN

Da das zu archivierende Material im Rahmen einer interdisziplinären Kooperation entstand, ist damit zu rechnen, dass Forschende aus unterschiedlichen Disziplinen an der Nachnutzung interessiert sind. Vor diesem Hintergrund entschieden wir uns, die Annotation und Segmentierung des Rohmaterials in einem entsprechend interdisziplinären und interkulturellen Team umzusetzen. Das Material wurde von dem Bienenbiologen Mazi Sanda und dem Ethnologen Martin Gruber aus ihrer jeweiligen fachspezifischen Perspektive annotiert und verschlagwortet. Die Kuration des Materials sowie die Entwicklung und Erprobung von angemessenen Verfahren wurde auf Seiten von Qualiservice von Michaela Rizzolli und Benjamin Quasinowski begleitet, die neben technischen Kenntnissen auch über fachlich-methodische Kompetenzen in der Ethnografie verfügen. Die gemeinsame Entwicklung von Kategorien und ein großer Teil der Kontextualisierung fanden während eines Gastaufenthalts von Mazi Sanda an der Universität Bremen im Jahr 2023 statt.

Bei der Annotation entschieden wir uns für drei verschiedene Arten von Kontextinformationen, die von Gruber und Sanda individuell in ELAN eingetragen wurden.

1. *Titel und Untertitel*: kurze und prägnante Beschreibungen der gefilmten Aktivitäten und Gesprächsinhalte konnten entweder für die gesamte Videodatei oder für Teilbereiche vergeben werden. Dadurch wird das Auffinden von relevanten Informationen/Segmenten im Filmmaterial erleichtert.
2. *Keywords*: fachspezifische Schlagworte (vgl. Tab. 2), die auf aktuelle Forschungsfelder oder theoretische Konzepte in der jeweiligen Disziplin hinweisen, die in der jeweiligen Datei explizit oder implizit verhandelt werden. Diese sind für Interessierte hinsichtlich der Potenziale und möglichen Szenarien der Nachnutzung in themenverwandten Fächern von zentraler Bedeutung, da sie Anknüpfungspunkte aufzeigen. Im Rahmen der Use Study entschieden wir uns vorab für die Vergabe von sieben Keywords pro Disziplin, die vor Beginn der Verschlagwortung festgelegt wurden. Mit ELAN ist eine unbegrenzte Anzahl von Keywords und die Vergabe von neuen Keywords während des Kodierens möglich.
3. *Notes*: allgemeine Kontextinformationen zu Praktiken und Ereignissen im Sinne von Memos. Unter der Kategorie wurden einerseits Informationen vermerkt, deren Bedeutung nicht als allgemein bekannt vorausgesetzt werden kann (lokales Wissen), andererseits Gedanken der Kodierenden festgehalten, die zur besseren Einordnung und zum tiefergehenden Verständnis des Materials beitragen.

Tabelle 2: Ergebnis 1: Fachspezifische Keywords

Discipline	Keywords	Description
Biology	**Beehives**	Conical containers made of grasses in which bees are kept to make and store honey
	Beeplants	Plant species that produce nectar and/or pollen sought by bees/foragers
	Landscape	A portion of the contryside shows vegetation richness and farm's lands
	Beehive materials	Sequence showing the collection of material to be used for the construction of a conical traditional beehive
	Baiting	Baiting the hive to attract bees' colony into the hive
	Honeycombs	Mass of hexagonal prismatic cells built from wax by honeybees in their hives/nests to contain their brood and stores of honey and pollen
	Honey processing	Sequence showing the chronological procedure that involves the extraction, filtration, and packaging of the harvested honey

Discipline	Keywords	Description
Anthropology	**Gbaya culture**	Oral history, stories and local knowledge that relate to or are informed by Gbaya culture
	Apprenticeship	Sequences in which a master-apprenticeship relationship is perfomed or discussed
	Craft	Materials and objects mediate between humans and honeybees. The manufacturing of the hives is based on embodies knowledge. All the sequences that show or relate to the manufacturing of hives and other objects
	Asthetics	Function and form are intertwined. Here we are interested in sequences that are related to questions of beauty and functionality
	Multispecies	Sequences in which immediate multi-species encounters are displayed. When humans and bees or other species get into contact, either physically or when human talk about the encounter
	Beekeeping	Sequences that are concerned with beekeeping
	Honey hunting	Sequences that are concerned with honey hunting

Die Annotationen wurden exportiert und als Mikro-Metadaten zur Rahmung und Beschreibung einzelner Videodateien innerhalb des Datensatzes bei Qualiservice integriert. Durch die Segmentierung war es zusätzlich möglich, nicht nur einzelne Videodateien detaillierter zu beschreiben, sondern die Videos in Abschnitte zu gliedern und diese Segmente mit fachspezifischen Keywords und weiteren Kontextinformationen anzureichern (vgl. Tab. 3). Titel und Untertitel ermöglichen einen groben Überblick über das Material und die Chronologie der Ereignisse bereits zum Zeitpunkt der Recherche. Die Annotationen können die Suche nach relevanten Datensätzen zusätzlich unterstützen und zum Verständnis des Materials bzw. der Bewertung der Passung des Materials zur eigenen Forschungsfrage, zur methodischen Herangehensweise und zum theoretischen oder konzeptionellen Rahmen beitragen.

Tabelle 3: Ergebnis 2: beschreibende Informationen innerhalb eines Videosegments

Early morning preparation					
Begin Time	**End Time**	**Subtitle**	**Notes Mazi**	**Notes Martin**	**Keywords**
00:00:00	00:29:11		Spraying water to soften the herbs and facilitate braiding	Extreme long shot of the entore process of wrapping the framework with grasses; gives an impression oft he time necessary fort he hive's construction	Beehive; Beehive materials; Apprenticeship; Craft; Aesthetics

Early morning preparation					
Begin Time	**End Time**	**Subtitle**	**Notes Mazi**	**Notes Martin**	**Keywords**
00:00:00	00:05:09	Softening the herbs		Where does the grass come from? Would it be possible to use fresh grass?	
00:05:09	00:06:19	Laying leaves			
00:06:19	00:09:48	Laying the beehive skeleton			
00:09:47	00:22:00	Braiding			
00:22:00	00:23:51		He talks how beauty and functionality go together in the construction of the hives	Significance of the tip of the hive where functionality and aesthetics unite	Aesthetics
00:23:51	00:24:17	Inside the hive			
00:24:17	00:25:01	Closing the hive with its lid			

Am Gastwissenschaftler*innen-Arbeitsplatz können die Videoaufnahmen inkl. der Annotationen in ELAN eingesehen bzw. genutzt werden (siehe Abb. 1). Dadurch wird die Navigation sowie das Durchsuchen von Material erleichtert, indem etwa die einzelnen Kapitel und Unterkapitel aufgrund der Segmentierung gezielt angesteuert werden können. Im Fall von mehrsprachigem Material ist das Vorgehen besonders gewinnbringend, da die zusätzlichen Informationen es ermöglichen, sich das Videomaterial unabhängig von Sprachkompetenzen zu erschließen. Das ist besonders interessant für Forschungsprojekte, deren Fokus nicht auf sprachlichen Äußerungen, sondern auf materieller Kultur, Praktiken oder anderen non-verbalen Aspekten liegt.

Als Teil der Workflowentwicklung wurde zudem nach Wegen einer verteilten Archivierung von Rohmaterial und Ethnografischem Film gesucht bei denen der Projektzusammenhang sichtbar und nachvollziehbar bleibt. In Zusammenarbeit mit dem Projekt EthnOA – Open Access in den ethnologischen Fächern[21] beim FID SKA konnte ein gemeinsamer Workflow für ethnografisches Filmmaterial entwickelt werden, bei dem die sensiblen Rohdaten bei Qualiservice archiviert und die frei einsehbaren Filme im EthnOA-Bereich des Medien-Repositoriums[22] der Humboldt-Universität zu Berlin veröffentlicht werden und über die Metadaten unmittelbar aufeinander verweisen.

21 https://www.evifa.de/de/services/publikationsserver-open-access (14.08.2024).

22 https://rs.cms.hu-berlin.de/ethnoa-medien/pages/home.php?login=true (14.08.2024).

5. Nachnutzungsszenarien für Daten aus der ethnografischen Filmforschung

Am Beispiel von Videoaufnahmen für einen ethnografischen Film haben wir gezeigt, wie sich die Entwicklung und Erprobung von Archivierungslösungen in einer bedeutungsstiftenden kommunikativen Triade von Primärforschende (Datengebende), Forschungsdatenzentrum und imaginierten Nachnutzenden gestalten lässt. Das im Rahmen der Use Study archivierte Material weist aufgrund der ethnografischen Reichhaltigkeit und der interdisziplinären Ausrichtung des Projekts ein hohes Nachnutzungspotenzial auf und bietet Anknüpfungspunkte für verschiedene Disziplinen, u. a. der Ethnologie, Biologie, Klima-, Umwelt- und Biodiversitätsforschung, den Area Studies oder Multispecies Research. Darüber hinaus bieten sich Re-Analysen an, die das Rohmaterial des Films mit später erhobenem Material kombinieren. Archiviertes Material kann auch für breiter angelegte Vergleichsstudien genutzt werden und bietet die Möglichkeit, kulturelle Praktiken über verschiedene Zeitperioden und Orte hinweg zu vergleichen. Rohmaterial und Film können darüber hinaus bei der Planung von ähnlichen Forschungsvorhaben verwendet werden oder im Rahmen von Qualifikationsarbeiten,[23] Lehrforschung oder Methodenlehre von Interesse sein (Witzel/Reiter 2022; Wilke/Pröbock/Pach 2019, S. 469–470; vgl. Wilke i. d. B.). Auch wenn das Teilen von Videodaten – gerade im Kontext ethnografischer Daten aus rechtlichen und/oder ethischen Gründen, nicht immer möglich sein wird, sollte eine Archivierung, die auch eine wissenschaftliche Nachnutzung ermöglicht, zumindest in Erwägung gezogen werden.

Im Rahmen der Use Study wurden wesentliche Elemente für einen angemessenen Workflow zur Kuration von Videodaten entwickelt, um diese bei Qualiservice archivieren und für weitere Forschung nutzbar machen zu können. Die Ergebnisse zeigen, wie Ansätze der Annotation und Segmentierung genutzt werden können, um audiovisuelles Material und seine Kontexte so für die Nachnutzung aufzubereiten, dass die Navigation im Filmmaterial erleichtert und ein tieferes Verständnis über disziplinäre Grenzen hinweg gefördert wird. Da Forschende – wie oben beschrieben – aufgrund datenschutzrechtlicher Bestimmungen nach Bremen reisen müssen, um audiovisuelles Material einzusehen, war es uns besonders wichtig, bereits im Zuge der Recherche möglichst viele Informationen bereitzustellen, damit Forschende besser abschätzen können, ob das Material für ihre Forschung relevant ist. Wesentlich dabei ist, dass die erprobten Ansätze und Ergebnisse nicht zwangsläufig für alle Arten der ethnografischen Filmforschung eine passende Lösung darstellen. Vielmehr zeigen sie Möglichkeiten und

23 Eine Nutzung der archivierten Daten bei Qualiservice ist nur für Wissenschaftler*innen mit Promotion möglich. Sollen Materialien im Rahmen von Qualifikationsarbeiten genutzt werden, muss die Vereinbarung von der betreuenden Person unterzeichnet werden.

Verfahren der Kontextualisierung und Aufbereitung von Videodaten auf, die im Einzelfall und in Abstimmung mit Qualiservice flexibel an die spezifischen Erfordernisse eines konkreten Projekts oder Materials angepasst werden können. Aufgrund dieser Flexibilität sind die erarbeiteten Lösungen nicht nur auf den spezifischen Anwendungsfall beschränkt, sondern auch für audiovisuelles Material aus verschiedenen Bereichen nutzbar und könnten daher auch für andere Forschungsdateninfrastrukturen von Bedeutung sein.

Größere Herausforderungen bestehen bislang noch darin, die inhaltlichen Metadaten (Annotationen) mit anderen Metadatenportalen auszutauschen und auffindbar zu machen. Im Rahmen der global operierenden Data Documentation Initiative (DDI-Alliance) koordiniert Qualiservice die Working-Group »Non-numeric data/qualitative data«,[24] die sich auf internationaler Ebene für Fragen der FAIRen[25] Beschreibung qualitativer Datensätze engagiert. Durch die Arbeit dieser Untergruppe entstehen Synergien, insbesondere im Hinblick auf die Verbesserung der Nachnutzung und der Interoperabilität audiovisueller Daten. Durch gemeinsame Anstrengungen und den Austausch von Best Practices soll die Integration und Nutzung solcher Daten weiter optimiert werden.

Darüber hinaus haben sich wichtige Fragen für die Archivierung und Nachnutzung von Videodaten ergeben, die für die zukünftige Arbeit relevant sind:

1. Wie können Verfahren der Annotation und Segmentierung in Forschungsroutinen der Videoanalyse bzw. der ethnografischen Filmforschung integriert werden, um den zeitlichen Aufwand zu minimieren und eine Verdichtung von Aufgaben zum Projektende möglichst zu vermeiden? Inwieweit können im Forschungsprozess entstandene Codierungen oder ähnliche Elemente zur Anreicherung von Metadaten genutzt werden?
2. Wäre es sinnvoll, neue Annotationen, die im Zuge zukünftiger Nachnutzungen entstehen, in die Metadaten zu integrieren (Versionierung von Datensätzen)?
3. Könnte es hilfreich sein, kleinere Ausschnitte des Materials oder »Video Snippets« frei zugänglich bereitzustellen, um potenziellen Nachnutzer*innen einen besseren Eindruck vom Material zu vermitteln?
4. Welche Möglichkeiten der verteilten Archivierung mit anderen Archiven und Forschungsdatenzentren wären für nicht-ethnografisches Videomaterial denkbar?
5. Unter welchen Bedingungen und mit welchen technischen Mitteln könnten Videodaten für die Produktion weiterer dokumentarischer oder journalistischer Filme zur Verfügung gestellt werden?

24 Weitere Informationen: https://doi.org/10.5281/zenodo.10225400.

25 http://www.go-fair.org/fair-principles (14.08.2024).

Literatur

Cram, Dominic L./van der Wal, Jessica E. M./Uomini, Natalie/Cantor, Mauricio/Afan, Anap I./Attwood, Mairenn C./Amphaeris, Jenny/Balasani, Fatima/Blair, Cameron J./Bronstein, Judith L./Buanachique, Iahaia O./Cuthill, Rion R. T./Das, Jewel/Daura-Jorge, Fábio G./Deb, Apurba/Dixit, Tanmay/Dlamini, Gcina S./Dounias, Edmond/Gedi, Isa I./Gruber, Martin/Hoffmann, Lilian S./Holzlehner, Tobias/Isack, Hussein A./Laltaika, Eliupendo A./Lloyd-Jones, David J./Lund, Jess/Machado, Alexandre M. S./Mahadevan, Machado L./Moreno, Ignacio B./Nwaogu, Chima J./Pierotti, Raymond/Rucunua, Seliano A./dos Santos, Wilson F./Serpa, Nathalia/Smith, Brian D./Sridhar, Hari/Tolkova, Irina/Tun, Tint/Valle-Pereira, João V. S./Wood, Brian M./Wrangham, Richard W./Spottiswoode, Claire N. (2022): The ecology and evolution of human-wildlife cooperation. In: People and Nature 4, H. 4, S. 841–855.

Eberhard, Igor (2020): Herausforderung ethnographische Daten: Erfahrungen und Ergebnisse aus dem Pilotprojekt ›Ethnographische Datenarchivierung‹ an der Universität Wien. In: Köstner-Pemsel, Christine/Stadler, Elisabeth/Stumpf, Markus (Hrsg.): Künstliche Intelligenz in Bibliotheken: 34. Österreichischer Bibliothekartag 2019. Graz: Unipress, S. 261–276.

Gruber, Martin (2016): Participatory ethnographic filmmaking as collaborative knowledge production. In: Visual Ethnography 5, H. 1, S. 15–44.

Gruber, Martin (2018): Hunters and guides: multispecies encounters between humans, honeyguide birds and honeybees. In: African Study Monographs 39, H. 4, S. 169–187.

Gruber, Martin (2022): Sharing the camera: a guide to collaborative ethnographic filmmaking. Sean Kingston Publishing.

Gruber, Martin (2023): Sensing bees. blogs.uni-bremen.de/beecultures (11.07.2024).

Gruber, Martin (2024a): Traditional beekeeping and honey hunting in central Cameroon. Audiovisual material, video footage [dataset]. Bremen: FDZ Qualiservice.

Gruber, Martin (2024b): Traditional beekeeping and honey hunting in central Cameroon. Ethnographic film research. Study report. Bremen: FDZ Qualiservice.

Gruber, Martin (Regie) (2015): Gbaya. Beekeeping and honey hunting [Film]. University of Bremen.

Gruber, Martin/Mazi Sanda (2019): Honey hunting and beekeeping in Adamaoua (Cameroon). Köln: Rüdiger Köppe.

Henley, Paul (2000): Ethnographic film: technology, practice and anthropological theory. In: Visual Anthropology 13, H. 2, S. 207–226.

Henley, Paul (2004): Putting film to work: observational cinema as practical ethnography. In: Pink, Sarah/Kürti, Laszlo/Afonso, Ana I. (Hrsg.): Working images: Visual research and representation in ethnography. New York/London: Routledge, S. 109–130.

Henley, Paul (2021): How to subtitle ethnographic films: Some practical suggestions. In: Visual Anthropology, 34, H. 1, S. 75–100.

Heuer, Jan-Ocko/Hollstein, Betina/Mozygemba, Kati (2021): ›Data Sharing‹ mit Qualiservice. In: Soziologie – Forum der Deutschen Gesellschaft für Soziologie 50, H. 4, S. 459–472.

Knoblauch, Hubert/Wilke, René (2018): Forschungsdateninfrastrukturen für audio-visuelle Daten der Qualitativen Sozialforschung – Bedarf und Anforderungen. In: RatSWD (Hrsg.): Archivierung und Zugang zu qualitativen Daten. RatSWD Working Paper 267/201. Berlin, Rat für Sozial und Wirtschaftsdaten (RatSWD), S. 47–58.

Kretzer, Susanne/Mozygemba, Kati/Heuer, Jan-Ocko/Huber, Elisabeth (2020): Erläuterungen zur Verwendung der von Qualiservice bereitgestellten Vorlagen für die informierte Einwilligung. Qualiservice Working Papers 2, Bremen: FDZ Qualiservice.

MacDougall, David (1982): Unprivileged camera style. In: Rain 50, S. 8–10.

MacDougall, David (2003): Beyond observational cinema. In: Hockings, Paul (Hrsg.): Principles of visual anthropology. Berlin: Mouton de Gruyter, S. 115–132.

Mazi, Sanda/Paxton, Robert J./Osterman, Julia (2022): A subset of wild bee species boosts the pollination of pigeon pea (Cajanuscajan: Fabaceae), an important crop plant of Cameroon. In: Journal of Apicultural Research 62, H. 3, S. 598–606.

Mohamadou, Moussa/Faïbawa, Esaïe/Mazi, Sanda (2022): Pollination efficiency of Ceratina cyanea (Hymenoptera: Apidae) on Helianthus annuus (Asteraceae) flowers at Dang (Ngaoundéré, Cameroon). In: Open Journal of Ecology 12, H. 1, S. 404–417.

Mozygemba, Kati/Hollstein, Betina (2023): Anonymisierung und Pseudonymisierung qualitativer textbasierter Forschungsdaten – eine Handreichung. Qualiservice Working Papers 5, Bremen: FDZ Qualiservice.

Mozygemba, Kati/Kretzer, Susanne (2022): Datenvielfalt im Data-Sharing – eine kooperative Aufgabe von Forschenden und Forschungsdatenzentrum. In: Lohmeier, Christine/Wiedemann, Thomas (Hrsg.): Datenvielfalt in kommunikationswissenschaftlichen Forschungskontexten. Wiesbaden: Springer VS, S. 157–178.

Pink, Sarah (2007): Doing visual ethnography: images, media and representation in research. 2. Auflage. London: Sage.

Pink, Sarah (2009a): Applied visual anthropology. Social interventions and visual methodologies. In: Pink, Sarah (Hrsg.): Visual interventions: applied visual anthropology. Oxford: Berghahn, S. 3–28.

Pink, Sarah (2009b): Doing sensory ethnography. Los Angeles, Calif.: Sage.

Rizzolli, Michaela/Imeri, Sabine/Huber, Elisabeth (2024): Ethnografische Forschungsmaterialien zur Archivierung und Nachnutzung vorbereiten und dokumentieren – ein Überblick für Forschende. Qualiservice Working Papers 6, Bremen: FDZ Qualiservice.

Tuma, René/Knoblauch, Hubert/Schnettler, Bernt (2013): Videographie. Einführung in die interpretative Video-Analyse sozialer Situationen. Wiesbaden: Springer VS.

Van der Wal, Jessica E. M./Spottiswoode, Claire N./Uomini, Natalie/Cantor, Mauricio/Daura-Jorge, Fábio G./Afan, Anap I./Attwood, Mairenn C./Amphaeris, Jenny/Balasani, Fatima/Begg, Colleen. M./Blair, Cameron J./Bronstein, Judith L./Buanachique, Iahaia O./Cuthill, Rion R. T./Das, Jewel/./Deb, Apurba/Dixit, Tanmay/Dlamini, Gcina S./Dounias, Edmond/Gedi, Isa I./Gruber, Martin/Hoffmann, Lilian S./Holzlehner, Tobias/Isack, Hussein A./Laltaika, Eliupendo A./Lloyd-Jones, David J./Lund, Jess/Machado, Alexandre M. S./Mahadevan, Machado L./Moreno, Ignacio B./Nwaogu, Chima J/Pereira, Valdomiro L./Pierotti, Raymond/Rucunua, Seliano A./dos Santos, Wilson F./Serpa, Nathalia/Smith, Brian D./Tolkova, I./Tun, Tint/Valle-Pereira, João V. S./Wood, Brian M./Wrangham, Richard W./Cram, Dominic L. (2022): Safeguarding human-wildlife cooperation. In: Conservation Letters 15, H. 4.

Van der Wal, Jessica E. M/Afan, Anap I./Anyawire, Mariamu/Begg, Colleen. M/, Begg, Keith. S./Dabo, Gabriel. A./Gedi, Isa. I./Harris, Jacob. A./Isack, Hussein A./Ibrahim, Joseph. I./Jamie, Gabriel. A./Kamboe, Wiro-Bless/Kilawi, Amana O./Kingston Anna/Laltaika, Eliupendo A./Lloyd-Jones, David/M'manga, George Malembo/Muhammad, Naziru Zakariah/Ngcamphalala, Celiwe A./Nhlabatsi, Salene O./Oleleteyo, Tumaini Tumbeine/Mazi, Sanda/Tsamkxao, Leon/Wood, Brian M./Spottiswoode, Claire N./Cram, Dominic L. (2023): Do honey badgers and greater honeyguide birds cooperate to access bees' nests? Ecological evidence and honey-hunter Accounts. In: Journal of Zoology 321, H. 1, S. 22–32.

Wilke, René/Pröbrock, Willi/Pach, Helen (2019): Infrastrukturen für Forschungsdaten der qualitativen Sozialforschung. In: Soziologie 48, H. 4, S. 467–486.

Witzel, Andreas/Reiter, Herwig (2022): Das problemzentrierte Interview – eine praxisorientierte Einführung. Weinheim/Basel: Beltz Juventa.

Bildungshistorische Perspektiven auf audiovisuelle Dokumente aus pädagogischen Fortbildungskontexten und der Unterrichtsforschung

May Jehle

Zusammenfassung: Die Unterrichtsvideographie verfügt in der erziehungswissenschaftlichen Forschung über eine bis in die 1960er Jahre zurückreichende Tradition. An einzelnen Hochschulstandorten und Forschungsinstitutionen erhalten gebliebene und mittlerweile als historisch zu betrachtende audiovisuelle Dokumentationen können inzwischen als Quellen einer bildungshistorischen Fachunterrichtsforschung analysiert und interpretiert werden. So können die dokumentierten pädagogischen Praktiken und Organisationsformen zum einen im Lichte (fach-)didaktischer und bildungspolitischer Diskurse diskutiert werden, zum anderen sind deren Repräsentationen hinsichtlich der Art und Weise ihrer Erzeugung zu untersuchen und zu reflektieren. Ein entsprechendes Vorgehen wird in dem Beitrag am Beispiel einer Fallstudie zu Videoaufzeichnungen aus Fortbildungsveranstaltungen für Lehrkräfte aus der DDR, die in der Transformationsphase 1990/91 für den Unterricht zur politischen Bildung qualifiziert werden sollten, konkretisiert und illustriert.

Schlüsselwörter: Unterrichtsvideographie, erziehungswissenschaftliche Forschung, bildungshistorische Fachunterrichtsforschung, ko-konstruktive Praxis, Visualität

1. Unterrichtsvideographie in der erziehungswissenschaftlichen Forschung

Audiovisuelle Daten werden in der erziehungswissenschaftlichen Forschung in einem breiten thematischen und methodischen Spektrum erhoben und ausgewertet (vgl. Mempel/Winterscheidt; Herrle/Proske/Puzicha/Zimmer; Carvalho; Hamer/Helbig/Urban i. d. B.). Innerhalb dessen kann die Unterrichtsvideographie als ein eigenes Feld betrachtet werden. Dieses hat sich von den seit Mitte der 1990er Jahren durchgeführten groß angelegten quantitativ verfahrenden Wirksamkeitsstudien der Unterrichtsqualitätsforschung bis hin zu mikroskopischen qualitativ ausgerichteten Fallstudien, die sich den vielfältigen

Interaktionszusammenhängen in Klassenzimmern widmen, erheblich ausdifferenziert (vgl. Herrle/Rauin/Engartner 2016; Dinkelaker 2020). Ein ebenso breites Verwendungsspektrum von Unterrichtsvideographien ist im Bereich der Lehrer:innenbildung zu verzeichnen (vgl. z. B. Hauenschild et al. 2020), wo bereits in den 1960er Jahren erstmals Videotechnik zum Einsatz kam. Sowohl in der Bundesrepublik als auch in der DDR stellten vor allem an pädagogischen Hochschulen eingerichtete Unterrichtsmitschauanlagen ein unterstützendes Mittel zur Ermöglichung von Hospitationen auch für größere Zahlen an Studierenden dar. Zugleich wurden die Videoaufzeichnungen an zahlreichen Standorten auch zu Forschungszwecken genutzt (vgl. Schluß/Jehle 2013). Eine beträchtliche Zahl solcher Aufzeichnungen von unterschiedlichen Standorten aus der DDR und der Bundesrepublik Deutschland konnten inzwischen in einer Reihe von Forschungsprojekten retro-digitalisiert und in Datenbanken beim Forschungsdatenzentrum Bildung (FDZ Bildung) des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) der wissenschaftlichen Forschung zugänglich gemacht werden (vgl. ebd.).[1]

Aus einer bildungshistorischen Forschungsperspektive können diese Dokumentationen Aufschluss über die Entwicklung spezifischer Unterrichtspraktiken geben, die über andere Quellen i. d. R. nur schwer zugänglich sind (vgl. z. B. Braster/Grosvenor/Pozo Andrès 2011; Grosvenor/Lawn/Rousmaniere 1999; Mietzner/Myers/Peim 2005). So wird in diesen Aufzeichnungen die historisch-spezifische Materialität und Organisiertheit schulischen Unterrichts in ihrem Vollzug dokumentiert, womit sich das bildungshistorische Interesse auch auf Fragen der Wandlungsfähigkeit oder Beständigkeit solcher Praktiken im Kontext von Transformations- und Reformprozessen richtet (vgl. Berdelmann et al. 2019; Reh/Jehle 2020). In besonderer Weise ist in der Analyse und Interpretation der Doppelcharakter der Quelle zu berücksichtigen: Auf einer inhaltlichen Ebene werden zum einen die pädagogischen Praktiken und Organisationsformen der Untersuchung und Diskussion zugänglich; zum anderen gehen in deren Form der Dokumentation spezifische institutionelle Bedingungen, Wahrnehmungs- und Darstellungsmuster sowie unterschiedliche Intentionen der verschiedenen Akteur:innen vor und hinter der Kamera ein (vgl. Warmington/Gorp/Grosvenor 2011; Reh/Jehle 2020). Daher bedürfen die audiovisuellen Repräsentationen eines mehrperspektivischen Zugangs, der unter Berücksichtigung ihrer Entstehungs- und Verwendungszusammenhänge auch die Art und Weise ihrer Erzeugung rekonstruiert und entsprechend in die Interpretation und bildungshistorische Einordnung

1 Die Datenbank »Audiovisuelle Aufzeichnungen von Schulunterricht in der DDR« enthält aktuell 271 Aufzeichnungen: https://www.fdz-bildung.de/studiendetails.php?id=55. Die Datenbank »Audiovisuelle Aufzeichnungen von Schulunterricht in der Bundesrepublik Deutschland« enthält aktuell 86 Aufzeichnungen: https://www.fdz-bildung.de/studiendetails.php?id=116 (Stand: 13.06.2024). Informationen zu Nutzungsbedingungen: https://www.fdz-bildung.de/nutzungsbedingungen-fdz-bildung.

der Quelle miteinbezieht. Diese Überlegungen sollen im Folgenden am Beispiel einer bildungshistorischen Fallstudie illustriert werden. Zur Diskussion gestellt wird damit auch ein forschungsmethodischer Zugang, der die Vielschichtigkeit dieser Quelle und damit verbundene Forschungsperspektiven auszudifferenzieren und zu erschließen sucht.

2. Entstehungs- und ursprünglicher Verwendungskontext: Fort- und Weiterbildungsveranstaltungen für Lehrkräfte aus der DDR

Wenn im Zuge der Analyse und Interpretation audiovisueller Unterrichtsaufzeichnungen die Wandlungsfähigkeit und Beständigkeit unterrichtlicher Praktiken diskutiert werden soll, erweist sich das Wissen um die ursprünglichen Entstehungs- und Verwendungskontexte in verschiedener Hinsicht als bedeutsam: Mit Blick auf das anzunehmende Professionswissen und entsprechende praktische Kompetenzen macht es einen Unterschied, ob beispielsweise routinierte Lehrkräfte zu Forschungszwecken etablierte oder neu entwickelte Konzeptionen zum Einsatz bringen oder ob die Aufzeichnungen (angehenden) Lehrkräften im Rahmen der Aus-, Fort- oder Weiterbildung in erster Linie zur Selbstreflexion dienen (Reh/Jehle 2020, S. 365).

Die drei für die Fallstudie ausgewählten Videoaufzeichnungen stammen aus Beständen des Referats für politische Bildungsarbeit der Freien Universität Berlin (FU). Dort wurde im Rahmen der Lehrkräftebildung seit Anfang der 1980er Jahre zur Verzahnung fachlicher, fachdidaktischer und unterrichtspraktischer Aspekte mit Unterrichtsvideos gearbeitet. Dafür wurde ein eigener Seminartyp mit Projektcharakter entwickelt, in dem Studierende Unterrichtseinheiten entwickelten, deren Durchführung auf Video aufgezeichnet und anschließend gemeinsam ausgewertet und diskutiert wurde. Einen Höhepunkt erlebte dieser Einsatz von Unterrichtsvideos schließlich mit dem wachsenden Bedarf an Fort- und Weiterbildung insbesondere von Lehrkräften aus der DDR in den frühen 1990er Jahren (vgl. Massing 2017, S. 244–250).

Dieser hohe (Weiter-)Qualifizierungsbedarf betraf nach der bereits im November 1989 erfolgten Aussetzung des Staatsbürgerkundeunterrichts, der als ein »Schlüsselfach« der politisch-ideologischen Erziehung interpretiert werden kann (vgl. Jehle 2018), vor allem den Fachbereich politischer Bildung. Im Februar 1990 wurde das Fach »Gesellschaftskunde« eingeführt, das im Zuge der Wiedervereinigung von bundeslandspezifischen Unterrichtsfächern abgelöst wurde. Neben den umfänglichen und breit geführten Diskussionen bezüglich der Ziele, Inhalte und methodischen Gestaltung politischen Fachunterrichts (vgl. Biskupek 2002) stellte der kurzfristige und hohe Qualifizierungsbedarf eine zentrale Herausforderung dar, in deren Bewältigung auch westdeutsche Institutionen eingebunden

waren (vgl. Jehle 2023). Am Referat für politische Bildungsarbeit der FU wurde dafür ein Weiterbildungscurriculum zu didaktischen Grundfragen und zur Unterrichtsplanung entwickelt, das die Aufzeichnung und Analyse der Unterrichtsdurchführungen als einen festen Bestandteil vorsah. Mit der Absicht, die interpretative Unterrichtsforschung voranzutreiben, wurden Transkriptionen einzelner Aufzeichnungen mit dazugehörigen fachwissenschaftlichen, -didaktischen und -methodischen Analysen publiziert (vgl. Massing 2017; Grammes/Weißeno 1993). Die Orientierung an Konzepten der westdeutschen Politikdidaktik erzeugte dabei ein spezifisches asymmetrisches Beziehungsgefüge zwischen den an den Aufzeichnungen beteiligten Protagonist:innen, die sich auch in diesen Publikationen spiegeln. Herausgestellt werden dabei im Besonderen beobachtbare Unsicherheiten und Schwierigkeiten der Lehrpersonen in der Umsetzung der für sie neuen Unterrichtsmethoden (vgl. z. B. Grammes/Weißeno 1993).

3. Weiterreichender Analysekontext: Eine bildungshistorische Analyseperspektive

Aus einer weiterreichenden bildungshistorischen Analyseperspektive bieten die Aufzeichnungen die Möglichkeit, die dokumentierten unterrichtlichen Praktiken im Kontext des beschriebenen Transformationsprozesses zu untersuchen. Bereits die ersten Rahmenpläne der Übergangskonzeption für das neue Unterrichtsfach, die von einer 60-köpfigen Kommission aus unterschiedlichen gesellschaftlichen Bereichen erarbeitet wurden (vgl. Biskupek 2002), sahen eine »grundlegende Neugestaltung der politischen Bildung und Erziehung in einem neuen Unterrichtsfach [vor, M. J.], das in seiner Zielstellung, seinem Inhalts- und Methodenkonzept sich prinzipiell vom bisherigen Staatsbürgerkundeunterricht unterscheiden wird« (Ministerium 1990a, S. 1). Diese Neuausrichtung sollte vor allem durch ein »didaktisch-methodisches Konzept« (Ministerium 1990b, S. 6) erreicht werden, »das durch einen demokratischen, kommunikations- und kooperationsfördernden Führungsstil des Lehrers und durch aktivierende Unterrichtsmethoden geprägt ist« (ebd.). Entsprechend werden in Darstellungen des Weiterbildungscurriculums der FU sowie Erfahrungsberichten die Einführung handlungs- und schülerorientierter Methoden, die für eine dialogischere und demokratischere Unterrichtskultur sorgen sollten, als ein Schwerpunkt im fachdidaktischen Bereich beschrieben. Hinsichtlich der inhaltlichen Vermittlungsebene stehen vor allem die Vermittlung von Funktionsbedingungen wie -defiziten eines politischen, demokratischen und pluralistischen Systems im Zentrum (Massing 1996; Massing 1997). Entsprechend dieser inhaltlichen Schwerpunktsetzung werden in den aufgezeichneten Unterrichtsstunden Möglichkeiten, Erfolge und Grenzen von Bürgerbewegungen in der DDR thematisiert. Die erste wurde bereits im Juni 1990 im Rahmen einer mehrwöchigen Fortbildung für

Dozent:innen aus der Lehrkräftebildung geplant und in einer achten Klasse einer Polytechnischen Oberschule (POS) in Potsdam aufgezeichnet. Gegenstand ist die lokale, tagesaktuelle Diskussion um die Umwandlung des nahegelegenen Baggersees, der trotz starker Verschmutzung zum Baden genutzt wird, in ein Naherholungsgebiet. Die Fragestellungen in den beiden anderen Aufzeichnungen beziehen sich direkt auf die Einschätzung der Erfolge von Bürgerbewegungen in der DDR. Die Konzeptionen wurden von den Teilnehmer:innen an Fort- und Weiterbildungskursen für Lehrkräfte aus der DDR entwickelt und an West-Berliner Schulen in der elften Jahrgangsstufe erprobt. In allen Konzeptionen sind Gruppenarbeitsphasen vorgesehen, die die Schüler:innen in unterschiedlicher Weise zu einer eigenständigen Beschäftigung mit der Thematik aktivieren sollen (vgl. Jehle 2022, S. 327–450).

3.1 Pädagogische Praktiken und Organisationsformen zur Konstitution von Fachlichkeit

Wenn nun also die Umsetzung der neu einzuführenden Unterrichtsgegenstände und der methodisch dialogischer und demokratischer zu gestaltenden Unterrichtspraxis in den Blick genommen werden soll, richtet sich der Fokus der Analyse auf das Verhältnis von auf der Grundlage der Planung rekonstruierbaren Intentionen und der dokumentierten Unterrichtspraxis. Auf der Grundlage einer kommunikationstheoretischen Modellierung von Unterricht (vgl. z. B. Meseth/Proske/Radtke 2011) und ethnomethodologischer Grundannahmen (vgl. z. B. Tuma/Schnettler/Knoblauch 2013, S. 89–93) wird rekonstruiert, wie die Lehrperson und die Schüler:innen im Unterrichtsgespräch ko-konstruktiv Sinn und Bedeutung in Bezug auf den Unterrichtsgegenstand hervorbringen (vgl. Jehle 2020). Der intendierte Transformationsprozess der Unterrichtspraxis bezieht sich dabei zum einen auf den Unterrichtsgegenstand selbst und das diesbezüglich zu vermittelnde Wissen. Zum anderen setzt die Zielsetzung der eigenständigen Urteilsbildung spezifische Umgangsweisen mit diesem Wissen und diese unterstützende Praktiken der Vermittlung voraus (vgl. Jehle 2024, S. 282).

In der hier vorgestellten Fallstudie sind im Zusammenhang mit den aktivierenden, die Schüler:innen einbeziehenden Unterrichtsmethoden und Frageimpulsen zur Gestaltung einer dialogischeren Unterrichtspraxis allerdings noch Entwicklungsbedarfe auf Seiten der Lehrperson zu erkennen. Am deutlichsten wird das in der ersten der drei Aufzeichnungen, in der die Lehrerin die Schüler:innen im Vorfeld zwar Anwohner:innenbefragungen zur Ausgangssituation vornehmen lässt, die Zielorientierung der Unterrichtseinheit allerdings bereits mit der Stundeneröffnung noch vor der Vorstellung der Ergebnisse vorwegnimmt:

Lw: »Und wir wolln in dieser Stunde mal überlegen, welche Möglichkeiten haben Bürgerbewegungen, hier etwas zu verändern, und welche Möglichkeiten haben wir als Schüler, auch daran teilzunehmen« (AzE 1990: # 00:00:20).

Im weiteren Verlauf holt die Lehrerin dann von den Schüler:innen zwar auch Vorschläge bezüglich des möglichen Vorgehens ein, die sie allerdings mittels ihrer Kommentierung und Suggestivfragen Bewertungen unterzieht, die den Möglichkeitsraum bereits deutlich einschränken. Ohne dass mögliche Handlungsalternativen einer ernsthaften Diskussion unterzogen werden, unterbreitet die Lehrerin schließlich selbst eine Handlungsempfehlung:

Lw: »So, dann würd ich n kleinen Vorschlag mal von euch machen, (…) Ein Vorschlag wäre vielleicht auch der, wenn ihr natürlich mutig genug dazu seid, (…) euch meinetwegen vor die Kaufhalle zu stellen oder sonstwohin und da selber Unterschriften zu sammeln (…) Es wäre also ne Möglichkeit, die ihr doch durchdenken solltet und dann vielleicht auch wirklich mal so eine Art eigene Bürgerinitiative von euch ausgehen lasst. Denn Schüler sind auch Bürger und zwar welche, die noch ne ganze Menge was-, noch ganz schön lange was wollen« (ebd.: # 00:29:13).

Die Aktivierung der Schüler:innen wird hier also über den Unterricht hinaus gedacht und findet nicht in dem Sinne statt, dass die Schüler:innen zu einer eigenständigen Auseinandersetzung mit der Sache und zu einer entsprechenden Urteilsbildung angeregt werden. Insgesamt lässt die Gesprächsführung der Lehrperson keine Anzeichen erkennen, dass die Beiträge der Schüler:innen ernst genommen und gewürdigt werden. So gilt ein konsequent aktivierender Einbezug der Schüler:innen zwar generell als anspruchsvolle Unterrichtsmethodik, da sie auf Seiten der Lehrpersonen ein ausgebildetes situatives Reaktionsvermögen in der Moderation ergebnisoffener Arbeitsprozesse erfordert. In dieser Aufzeichnung lassen die Äußerungen der Lehrperson allerdings die Ergebnisoffenheit ihrer Planung bezweifeln (vgl. Jehle 2022, S. 327–363). Somit dokumentiert die Aufzeichnung auch die Schwierigkeit, einmal etablierte Praktiken der Unterrichtsführung zu transformieren.

Verschärft zeigt sich diese Herausforderung auch in den beiden anderen in West-Berliner Schulklassen aufgezeichneten Unterrichtsstunden dadurch, dass Schüler:innen und Lehrpersonen in Bezug auf den Gegenstand über unterschiedliche biographisch begründete Deutungsperspektiven verfügen. Insbesondere in einer der Aufzeichnungen kommt es wiederholt zu Situationen, in denen die Beiträge der Schüler:innen Wahrnehmungen aus der West-Perspektive oder der westlichen medialen Berichterstattung spiegeln. Die Schwierigkeiten der Lehrperson, von diesen Äußerungen zu seinen nächsten Planungsschritten überzuleiten, lassen darauf schließen, dass in seinem Erwartungshorizont andere Beiträge vorgesehen waren. Am deutlichsten werden diese Divergenzen in beiden

Aufzeichnungen, wenn abschließend die Möglichkeiten von Bürgerbewegungen diskutiert werden. Vorangegangen war den Diskussionen in einer der Aufzeichnungen eine Gruppenarbeit, in der sich die Schüler:innen mit den Zielen von Bürgerbewegungen unter politischen, ökonomischen und ökologischen Aspekten beschäftigt haben. In der abschließenden Diskussion, ob die Bürgerbewegungen gescheitert seien, lassen die Beiträge der Schüler:innen ein Begründungsmuster erkennen, in dem das entscheidende Kriterium für Erfolg das Erreichen einer Mehrheit bei einer Wahl darstellt und an die Voraussetzung einer Parteigründung gebunden ist:

> *Sw10: »Also, die sind jetzt ja die einzigen, die jetzt zumindest auch gewählt wurden, (…) denn, wenn man halt nich, äh, sich irgendwo als Partei oder so betätigt, dann kann man eigentlich kaum ne Öffentlichkeit erreichen und auch kaum irgendwie auf sich aufmerksam machen« (AzE 1991a: # 00:52:25).*

> *Sm9: »In der Demokratie braucht man ne Mehrheit für etwas und wenn man die nicht hat, dann kann mans vergessen, kann man versuchen, die zu kriegen, aber ansonsten kann man da eigentlich wenig, wenig Chancen. Also, denk ich jedenfalls mal so« (AzE 1991a: # 00:54:19).*

Die Lehrperson dagegen verweist abschließend noch auf mögliches zukünftiges Potenzial von Bürgerbewegungen, wobei die damit verbundene Skepsis auch eine biographisch begründete Distanz zum politischen System impliziert:

> *Lm: »Man könnte allerdings eben auch sagen, angesichts der sich zuspitzenden ökonomischen und sozialen Konflikte in den neuen Bundesländern könnten die Aktivitäten und die Stärke der Bürgerbewegungen der ehemaligen DDR wieder erwachen und wieder erstarken und Bevölkerung, die Menschen in den neuen Bundesländern mobilisieren zur aktiven Tat, zu einer Bewegung der Bürger, aber dann ergibt sich auch gleich, muss man auch gleich die Frage stellen, haben denn nun, äh, diese Bürgerbewegungen angesichts der neuen Machtstrukturen in Deutschland überhaupt ne Chance oder ham se keine« (AzE 1991a: # 01:01:41).*

In der anderen Unterrichtsstunde wurde ein Gespräch mit einem Vertreter einer Bürgerbewegung geführt, für das die Schüler:innen in Kleingruppen Fragen vorbereitet haben. Dabei lassen vor allem die Fragestellungen der Schüler:innen erneut eine deutliche Orientierung an der Funktionslogik der repräsentativen Demokratie erkennen, wohingegen der Vertreter der Bürgerbewegung die Bedeutung von parteiunabhängigen Interessensvertretungen unterstreicht:

> *Sm13: »Also, unsere Hauptfrage ist eigentlich, ob es nicht ein Nachteil ist, keinen Parteiapparat zu haben …? (…) Ist es nicht so, dass man nach politischer Macht streben muss,*

um seine Vorstellungen und Ziele zu verwirklichen (…)? Und noch ne weitere Frage, ob es nicht sinnvoll wäre, eine Partei aus dem Neuen Forum zu formieren?« (AzE 1991b: # 00:38:52)

Pm: »In meinem Verständnis ist Bürgerbewegung eigentlich ein Prozess der Selbstorganisation, der die Aufgabe hat, eine neuere und breitere Öffentlichkeit zu schaffen. (…) Nun erleben wir aber, dass viele Probleme von den Parteien überhaupt nicht ernst genommen werden. (…) … weil eben im Entscheidungsprozess entweder der Bürger nicht gehört wird oder dieses schöne Wort von bürgernaher Verwaltung ein schönes Wort ist« (AzE 1991b: # 01:07:12).

So können die Schwierigkeiten in der diskursiven Weiterbearbeitung der Beiträge der Schüler:innen also auch auf unterschiedliche biographisch begründete Erfahrungen und Deutungsperspektiven in Bezug auf den Unterrichtsgegenstand zurückgeführt werden, was in den zeitgenössischen Auswertungen jedoch wenig Beachtung fand. Diese spiegeln doch eher das den Fort- und Weiterbildungsveranstaltungen inhärente asymmetrische Beziehungsverhältnis, indem sie in erster Linie die Defizite und Unsicherheiten der Lehrpersonen bei der Umsetzung der für sie neuen Unterrichtsmethoden betonen (Grammes/Weißeno 1993; Massing 1996).

3.2 Dokumentation und Visualisierung pädagogischer Praktiken

Parallel zur inhaltlichen Vermittlungsebene soll im Weiteren die Art und Weise der visuellen Dokumentation der Implementierung dieser neuen Unterrichtspraxis in den Blick genommen werden. In Anlehnung an hermeneutisch-wissenssoziologische Videoanalysen (Reichertz/Englert 2011) wird dabei grundsätzlich davon ausgegangen, dass diesen Visualisierungen komplexe Prozesse zugrunde liegen, die in technische und räumliche Möglichkeiten der Bildgestaltung eingebunden sind. Davon ausgehend sind damit aber auch Entscheidungen für spezifische Bildgestaltungen und -kompositionen verbunden, die auf zugrundeliegende Interessen, pädagogische Vorstellungen und Deutungsschemata verweisen (vgl. Jehle 2021, S. 75–76; Jehle 2022, S. 820–821).

Sofern das Interesse einer Unterrichtsvideographie auf die Vermittlungsebene gerichtet ist, sind diese Entscheidungen auch von der Absicht geleitet, diese offizielle und meist klassenöffentliche Interaktionsebene – die »Vorderbühne« des Unterrichts (Zinecker 1978) – abzubilden. Demgegenüber sind in einzelnen Aufzeichnungen der hier vorgestellten Fallstudie Fokussierungen durch die Kamera zu beobachten, die unterschiedliche Formen der Abweichung von der offiziellen Unterrichtsordnung dokumentieren. Im ersten Fall werden vor allem einzelne Handlungen von Schüler:innen fokussiert, die nicht der ausgesprochenen

Arbeitsanweisung durch die Lehrperson entsprechen. Während die Schüler:innen in Kleingruppen einen Aufruf zu einer Unterschriftensammlung verfassen sollen, zoomt die Kamera beispielsweise bis zur Nahaufnahme an die Hand einer Schülerin heran, die ihren Stift längere Zeit im Schwebezustand über dem Blatt hält, ohne etwas zu notieren (vgl. Abb. 1). Dabei signalisiert diese Handhaltung durchaus Arbeitsbereitschaft, das zum Ausdruck kommende Zögern könnte allerdings auf eine Überforderung durch die Aufgabenstellung verweisen. Als die Lehrerin die Schüler:innen schließlich zur Beendigung der Arbeit auffordert, zeigt die Kamera wiederum eine Gruppe Schülerinnen, die noch engagiert mit der Verschriftlichung beschäftigt sind. In einer zweiten Aufzeichnung wendet sich die Kamera dagegen zunehmend der »Hinterbühne« (Zinnecker 1978) der unterrichtlichen Interaktion zu. Die Aufmerksamkeit gilt vor allem einer Gruppe von Schüler:innen, in der regelmäßig belustigte Unterhaltungen zu beobachten sind, die nicht dem klassenöffentlichen Unterrichtsgespräch zugeordnet werden können. Ein Schüler und eine Schülerin, die auch auffallend physisch miteinander interagieren, werden dabei besonders detailliert in den Blick genommen (vgl. Abb. 2).

Abbildung 1 und 2: Zoom auf Abweichungen von der Unterrichtsordnung (Aufzeichnungseinheit 1990: 00:25:48; 1991a: 00:04:38)

Im Kontext der Fort- und Weiterbildungsveranstaltungen lassen sich diese Fokussierungen als spezifische Handlungen des Zeigens (vgl. Reichertz/Englert 2011, S. 28) interpretieren. Durch die explizite Dokumentation rücken sie diese Aktivitäten der Schüler:innen, die von den eigentlichen Arbeitsaufträgen abweichen, in die Aufmerksamkeit der Betrachter:innen. Den Lehrpersonen, denen diese Abweichungen möglicherweise im unterrichtlichen Vollzug entgangen sind, werden sie zur nachträglichen Reflexion zugänglich gemacht. Damit verweisen diese Visualisierungen auch auf das asymmetrische Beziehungsverhältnis in den Fort- und Weiterbildungsveranstaltungen, in denen bezüglich der neu zu implementierenden Unterrichtspraktiken ausschließlich die Lehrkräfte aus der DDR als Lernende adressiert werden.

Mit Blick auf diese neu zu implementierenden Unterrichtspraktiken sind zudem vor allem die Visualisierungen der zur Aktivierung in Gruppenarbeit

organisierten Erarbeitungsphasen von Interesse, wobei es auch das Verhältnis von Kamerapositionierung, Raum- und Sitzordnung zu berücksichtigen gilt. In allen drei Aufzeichnungen wird mit einer feststehenden, schwenkbaren und zoomfähigen Kamera aufgezeichnet, die im vorderen Bereich des Klassenzimmers auf der rechten (Fenster-)Seite positioniert ist. Die Klasse als Ganzes kann dabei nur in einer langsamen Schwenkbewegung, aber nicht in einer Einstellung in der Totalen in den Blick genommen werden. Im zeitlichen Verlauf der Aufzeichnungszeitpunkte sind dabei Transformationen in der Sitzordnung zu beobachten: Bereits in der ersten Aufzeichnung wird die bisher den Unterricht in der DDR dominierende, ausschließlich frontal ausgerichtete Anordnung der Tische aufgebrochen. So ermöglicht die Sitzordnung mit einer seitlich angeordneten Reihe im Sinne einer kommunikationsfördernden Unterrichtsgestaltung mehr Blickkontakte zwischen den Schüler:innen. Für Gruppenarbeiten ist sie allerdings noch nicht ganz prädestiniert, die Schüler:innen arbeiten in Zweier- oder Dreiergruppen nebeneinander sitzend zusammen (vgl. Abb. 3). In der zweiten Aufzeichnung sind die Tische bereits in Gruppentischen angeordnet, was insgesamt aber improvisiert wirkt (vgl. Abb. 4). Die ungenutzte Tischreihe zwischen Lehrperson und Schüler:innen lässt auf eine spontane Umstellung der Tische schließen, wobei der gesamten Sitzordnung auch im Verhältnis zur Kamerapositionierung nur wenig Aufmerksamkeit gewidmet wurde. Dass die Schüler:innen teilweise sowohl einander als auch der Kamera den Rücken zuwenden, dürfte weder der über die Gruppenarbeit hinausgehenden Interaktion in der Klasse förderlich sein noch ermöglicht es eine Dokumentation der in den Gruppen stattfindenden Interaktionen. Erst die Anordnung der Gruppentische sowie der an ihnen sitzenden Schüler:innen in der kurz darauf entstandenen dritten Aufzeichnung ermöglicht einen übersichtlichen Ein- und Überblick über die unterschiedlichen Arbeitsphasen hinweg (vgl. Abb. 5).

Abbildung 3–5: Einblicke in Raum- und Sitzordnung (Aufzeichnungseinheit 1990: 00:26:16; 1991a: 00:00:26; 1991b: 00:00:50)

In Bezug auf die räumliche Ordnung und die visuelle Dokumentation der Sozialform Gruppenarbeit sind also auch Entwicklungsprozesse auf Seiten der für die Fort- und Weiterbildungen verantwortlichen Akteur:innen zu verzeichnen. Dies verweist darauf, dass diese Sozialformen auch in der Bundesrepublik noch

keine selbstverständlich etablierten Unterrichtspraktiken darstellten, sondern dass auch dort von längerfristigen Transformationsprozessen in der Unterrichtspraxis auszugehen ist. Eine Bestätigung erfährt diese Deutungsperspektive zudem durch die im Vorangegangenen beschriebenen Handlungen des Zeigens, die mit ihrer oppositionellen Unterscheidung von Vorder- und Hinterbühne an einer tradierten Vorstellung unterrichtlicher Ordnung festhalten. Somit eröffnet diese Sekundäranalyse der Visualisierungen eine längerfristige Perspektive auf die dokumentierten Transformationsprozesse unterrichtlicher Praktiken, die diese nicht nur innerhalb eines asymmetrischen Beziehungsgefüges entlang einer politischen Zäsur wahrnimmt.

4. Bildungshistorische Forschungsperspektiven und weiterreichende (sekundär-)analytische Nutzungsszenarien

Entlang dieser Schlaglichter auf die Ergebnisse mehrperspektivisch ansetzender Analysen historischer Unterrichtsvideographien konnte am Beispiel dieser Fallstudie das Potenzial solcher (sekundär-)analytischer Nutzungsszenarien gezeigt werden. Jenseits der ursprünglich mit den Aufzeichnungen verfolgten Intentionen erlauben insbesondere die Rekonstruktionen der Visualisierungen unterrichtlicher Praxis eine Ausdifferenzierung der fachunterrichtlichen Erzählungen bezüglich der Einführung, Etablierung, Transformation und Stabilität spezifischer Unterrichtspraktiken. Die eingangs erwähnten Datenbanken des FDZ Bildung zur Archivierung solcher historischer audiovisueller Unterrichtsaufzeichnungen stellen somit einen vielversprechenden Fundus für eine bildungsgeschichtliche Fachunterrichtsforschung dar, um die dokumentierte unterrichtliche Praxis im Kontext der umfassenderen fachdidaktischen Diskussionszusammenhänge zu analysieren. Dies eröffnet die Perspektive, einer bisher vor allem konzeptionelle und organisatorische Dimensionen analysierenden Fachgeschichtsschreibung eine an Kontingenzen und Differenzen interessierte Historische Praxeologie an die Seite zu stellen (vgl. Berdelmann et al. 2019; Wähler et al. 2024). Empirische Befunde wie die relative Stabilität schulischer wie unterrichtlicher Formen als historisch gewachsene Formen und Eigendynamiken bildungspolitischer Steuerungsversuche (vgl. z. B. Tyack/Tobin 1994; Proske/Rabenstein 2018; Brückweh/Villinger/Zöller 2020) können hierbei als rahmende Bezugspunkte einer solchen Ausdifferenzierung dienen sowie auch selbst weiteren Ausdifferenzierungen unterzogen werden.

Zugleich ergibt sich hier eine produktive Schnittstelle zu einer qualitativen Unterrichtsforschung, die »*Unterricht* als *historisch gewordene Form*« (Proske/Rabenstein 2018, S. 7, kurs. i. O.) versteht und erforscht. Anhand der audiovisuellen Aufzeichnungen kann Unterricht mittels mehrperspektivischer Analysezugänge in seiner Komplexität, Materialität und Situiertheit erschlossen werden

(vgl. Dinkelaker 2020). Mit dem Fundus des FDZ Bildung ist eine erste Grundlage geschaffen, die lokale, zeitliche sowie fachliche Kontrastierungen ermöglicht. So lassen sich beispielsweise fachliche Besonderheiten in den jeweiligen Praktiken und Materialitäten im Umgang mit Wissen in verschiedenen Unterrichtsfächern herausarbeiten. Die Untersuchung der Herausbildung solcher »Traditionen schulfachlicher Kommunikations- und Vermittlungsformate« (Lorenz et al. 2024, S. 10) erweist sich nicht nur aus der Perspektive einer historischen Bildungsforschung als aufschlussreich. In ihrer Bedeutsamkeit für den fachlichen Gehalt des Unterrichts und der in ihm zu übenden Kompetenzen sind diese Formate auch »bei der Bestimmung von Bildungsstandards zu berücksichtigen und vor allem in der didaktischen Ausbildung zu reflektieren« (ebd.). Dabei kann die Analyse und Reflexion solcher unterschiedlichen Formate im Kontext der didaktischen Ausbildung nicht nur der Schulung des situativen Reaktionsvermögens angehender Lehrpersonen dienen. Sie kann auch für den Eigensinn schulischer Praxis und ihrer Akteur:innen sensibilisieren, der angenommenen Wirksamkeitslogiken bei der Implementierung von Steuerungsversuchen bisweilen auch zuwiderlaufen kann.

Die Forschung kann es sich dabei zunutze machen, dass die Audio-Visualität des Mediums die Analyse hoch komplexer Vorgänge ermöglicht, deren Sequenzialität in simultaner Koordination sozio-materiell körpergebundener Interaktionen hervorgebracht wird (vgl. Dinkelaker 2020, S. 25–27). Zugleich ist mit den Visualisierungen weitaus mehr implizites Wissen verbunden, das in den ursprünglichen Entstehungs- und Verwendungskontexten oftmals vernachlässigbar erscheint, deren Bedeutsamkeit allerdings in den Sekundäranalysen herausgearbeitet werden kann (vgl. z. B. Jehle 2021). Zu berücksichtigen sind dabei zunächst die räumlichen Gegebenheiten und technischen Möglichkeiten der Aufzeichnung. Davon ausgehend können die gewählten Kameraperspektiven und Fokussierungen – bisweilen auch die Räumlichkeiten selbst sowie ihre Einrichtung und vorhandene Gegenstände – Aufschluss über den Entstehungs- und Verwendungskontext und damit verbundene Intentionen und Erkenntnisinteressen geben, sofern hierzu nur unzureichende Informationen vorliegen. Darüber hinaus können sie aber auch auf implizite, möglicherweise gar nicht bewusst vorliegende Deutungsrahmen der Bildproduzent:innen verweisen. Ausgehend von der Frage, welchen Akteur:innen, Handlungen, Materialien und Interaktionen Aufmerksamkeit gewidmet wird, lassen sich bestimmte Vorstellungen unterrichtlicher Ordnung sowie Annahmen über Normalität und ihre Abweichungen sichtbar machen (Jehle 2022, S. 820–837). Der Fundus des FDZ Bildung könnte hier auch eine Perspektive eröffnen, in einer Längsschnittstudie die Flexibilität oder Stabilität solcher Vorstellungen auf verschiedenen Akteursebenen in unterschiedlichen Kontexten zu untersuchen und diese im Verhältnis zu bildungspolitischen oder fachdidaktischen Diskurslagen zu diskutieren.

Literatur

Aufzeichnungseinheit (1990): Baggersee (v_fu_30). Aus: Datenkollektion (2013–2015): Schluß, Henning/Jehle, May: Quellensicherung und Zugänglichmachung von Videoaufzeichnungen von Unterricht der Freien Universität Berlin. In: Audiovisuelle Aufzeichnungen von Schulunterricht in der Bundesrepublik Deutschland. Forschungsdatenzentrum Bildung am DIPF. Frankfurt am Main. https://doi.org/10.7477/19:1:18.

Aufzeichnungseinheit (1991a): Bürgerbewegungen in der DDR (v_fu_18a). Aus: Datenkollektion (2013–2015): Schluß, Henning/Jehle, May: Quellensicherung und Zugänglichmachung von Videoaufzeichnungen von Unterricht der Freien Universität Berlin. In: Audio-visuelle Aufzeichnungen von Schulunterricht in der Bundesrepublik Deutschland. Forschungsdatenzentrum Bildung am DIPF. Frankfurt am Main. https://doi.org/10.7477/19:1:8.

Aufzeichnungseinheit (1991b): Haben Bürgerbewegungen (k)eine Chance? (v_fu_oN). Aus: Datenkollektion (2013–2015): Schluß, Henning/Jehle, May: Quellensicherung und Zugänglichmachung von Videoaufzeichnungen von Unterricht der Freien Universität Berlin. In: Audiovisuelle Aufzeichnungen von Schulunterricht in der Bundesrepublik Deutschland. Forschungsdatenzentrum Bildung am DIPF. Frankfurt am Main. https://doi.org/10.7477/19:1:5.

Berdelmann, Kathrin/Fritzsche, Bettina/Rabenstein, Kerstin/Scholz, Joachim (2019): Praxeologie in der Bildungsforschung. Ein Umriss historischer und gegenwartsbezogener Forschungsperspektiven. In: Berdelmann, Kathrin/Fritzsche, Bettina/Rabenstein, Kerstin/Scholz, Joachim (Hrsg.): Transformationen von Schule, Unterricht und Profession. Erträge praxistheoretischer Forschung. Wiesbaden: Springer VS, S. 1–28.

Biskupek, Sigrid (2002): Transformationsprozesse in der politischen Bildung. Von der Staatsbürgerkunde in der DDR zum Politikunterricht in den neuen Ländern. Schwalbach/Ts.: Wochenschau.

Braster, Sjaak/Grosvenor, Ian/Pozo Andrés, Maria del Mar (Hrsg.) (2011): The Black Box of Schooling. A Cultural History of the Classroom. Brüssel: Peter Lang.

Brückweh, Kerstin/Villinger, Clemens/Zöller, Kathrin (Hrsg.) (2020): Die lange Geschichte der »Wende«. Geschichtswissenschaft im Dialog. Berlin: Christoph Links.

Dinkelaker, Jörg (2020): Potentiale der Theorieentwicklung durch erziehungswissenschaftliche Videographie. In: Corsten, Michael/Pitburg, Melanie/Wolff, Dennis/Hauenschild, Katrin/Schmidt-Thieme, Barbara/Schütte, Ulrike/Zourelidis, Sabrina (Hrsg.): Qualitative Videoanalyse in Schule und Unterricht. Weinheim und Basel: Beltz Juventa, S. 18–36.

Grammes, Tilman/Weißeno, Georg (Hrsg.) (1993): Sozialkundestunden. Politikdidaktische Auswertungen von Unterrichtsprotokollen. Opladen: Leske + Budrich.

Grosvenor, Ian/Lawn, Martin/Rousmaniere, Kate (Hrsg.) (1999): Silences and Images. The Social History of the Classroom. New York: Peter Lang.

Hauenschild, Katrin/Schmidt-Thieme, Barbara/Wolff, Dennis/Zourelidis, Sabrina (Hrsg.) (2020): Videografie in der Lehrer*innenbildung. Aktuelle Zugänge – Herausforderungen und Potenziale. Hildesheim: Universitätsverlag Hildesheim.

Herrle, Matthias/Rauin, Udo/Engartner, Tim (2016): Videos als Ressourcen zur Generierung von Wissen über Unterrichtsrealität(en). In: Rauin, Udo/Herrle, Matthias/Engartner, Tim (Hrsg.): Videoanalysen in der Unterrichtsforschung. Methodische Vorgehensweisen und Anwendungsbeispiele. Weinheim und Basel: Beltz Juventa, S. 8–28.

Jehle, May (2018): Staatsbürgerkunde – »Schlüsselfach« der politischen Erziehung in der DDR? In: Aus Politik und Zeitgeschichte 68, H. 13/14, S. 25–30.

Jehle, May (2020): Ethnomethodological Reconstruction. In: Huber, Matthias/Froehlich, Dominik E. (Hrsg.): Analyzing Group Interactions. A Guidebook for Qualitative, Quantitative and Mixed Methods. Abingdon und New York: Routledge, S. 156–168.

Jehle, May (2021): Visual ethnography in classrooms: the »action of showing« in classroom videos in contexts of social science teacher education. In: Journal of Social Science education 20, H. 4, S. 72–92.

Jehle, May (2022): Politische Bildung und Erziehung im Unterricht. Kontrastive Fallstudien anhand von Videoaufzeichnungen politischen Fachunterrichts in Ost-, West- und Gesamtberlin von 1978 bis 1993. Bad Heilbrunn: Klinkhardt.

Jehle, May (2023): »Gesellschaftskunde« als »Übergangsfach« in der Transformationsphase der DDR 1990/91. Analysen von Unterrichtsvideos im Kontext von Fort- und Weiterbildungsveranstaltungen

für Lehrkräfte aus der DDR. In: Göttlicher, Wilfried/Janík, Tomáš (Hrsg.): Politische Zäsur und Wandel des Bildungssystems. Drei Dekaden nach dem Fall des Eisernen Vorhangs – Bilanzen und Perspektiven. Wien: Lit, S. 149–175.

Lorenz, Marco/Wähler, Josefine/Reh, Sabine/Scholz Joachim (2024): Fachunterrichtsgeschichten. Studien zur Geschichte der Praxis des Fachunterrichts. Eine Einleitung. In: Wähler, Josefine/Lorenz, Marco/Reh, Sabine/Scholz, Joachim (Hrsg.): Fachunterrichtsgeschichten. Studien zur Geschichte der Praxis des Fachunterrichts. Bad Heilbrunn: Klinkhardt, S. 9–20.

Massing, Peter (1996): Das Fach Sozialkunde an der Freien Universität Berlin. In: Berliner Institut für Lehrerfort- und -weiterbildung und Schulentwicklung (Hrsg.): Weiterbildung und Wiedervereinigung. Ergebnisse und Erfahrungen in der Berliner Lehrerweiterbildung seit der Wiedervereinigung. Berlin: Selbstverlag, S. 40–47.

Massing, Peter (1997): Zwischen Verunsicherung, Politikverdrossenheit und Engagement. Politiklehrerinnen und -lehrer in den neuen Bundesländern. In: DVPB aktuell. Report zur politischen Bildung, H. 2, S. 21–26.

Massing, Peter (2017): Videobasierte Lerngelegenheiten als Grundlage eines kompetenzorientierten Lehramtsstudiums. In: Gloe, Markus/Oeftering, Tonio (Hrsg.): Perspektiven auf Politikunterricht heute. Vom sozialwissenschaftlichen Sachunterricht bis zur Politiklehrerausbildung. Festschrift für Hans-Werner Kuhn. Baden-Baden: Nomos, S. 243–255.

Meseth, Wolfgang/Proske, Matthias/Radtke, Frank-Olaf (2011): Was leistet eine kommunikationstheoretische Modellierung des Gegenstandes »Unterricht«? In: Meseth, Wolfgang/Proske, Matthias/Radtke, Frank-Olaf (Hrsg.): Unterrichtstheorien in Forschung und Lehre. Bad Heilbrunn: Klinkhardt, S. 223–440.

Mietzner, Ulrike/Myers, Kevin/Peim, Nick (Hrsg.) (2005): Visual History. Images of Education. Bern: Peter Lang.

Ministerium für Bildung (1990a): Übergangskonzeption im gesellschafts- und sozialkundlichen Unterricht (bisher Staatsbürgerkunde) für das 2. Halbjahr des Schuljahrs 1989/90. Berlin (Ost): Selbstverlag.

Ministerium für Bildung und Wissenschaft (1990b): Rahmenpläne für den Gesellschaftskundeunterricht (Erprobungslehrplan). Berlin (Ost): Selbstverlag.

Proske, Matthias/Rabenstein, Kerstin (2018b): Stand und Perspektiven qualitativ sinnverstehender Unterrichtsforschung. Eine Einführung in das Kompendium. In: Proske, Matthias/Rabenstein, Kerstin (Hrsg.): Kompendium Qualitative Unterrichtsforschung. Unterricht beobachten – beschreiben – rekonstruieren. Bad Heilbrunn: Klinkhardt, S. 7–24.

Proske, Matthias/Rabenstein, Kerstin (Hrsg.) (2018a): Kompendium Qualitative Unterrichtsforschung. Unterricht beobachten – beschreiben – rekonstruieren. Bad Heilbrunn: Klinkhardt.

Reh, Sabine/Jehle, May (2020): Visual History of Education – Audiovisuelle Unterrichtsaufzeichnungen aus der der DDR. In: Corsten, Michael/Pitburg, Melanie/Wolff, Dennis/Hauenschild, Katrin/Schmidt-Thieme, Barbara/Schütte, Ulrike/Zourelidis, Sabrina (Hrsg.): Qualitative Videoanalyse in Schule und Unterricht. Weinheim und Basel: Beltz Juventa, S. 348–370.

Reichertz, Jo/Englert, Carina J. (2011): Einführung in die qualitative Videoanalyse. Eine wissenssoziologisch-hermeneutische Fallanalyse. Wiesbaden: Springer VS.

Schluß, Henning/Jehle, May (Hrsg.) (2013): Videodokumentation von Unterricht. Zugänge zu einer neuen Quellengattung der Unterrichtsforschung. Wiesbaden: Springer VS.

Tuma, René/Schnettler, Bernd/Knoblauch, Hubert (2013): Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen. Wiesbaden: Springer VS.

Tyack, David/Tobin, William (1994): The »Grammar« of Schooling. Why Has it Been so Hard to Change? In: American Educational Research Journal 31, H. 3, S. 453–479.

Wähler, Josefine/Lorenz, Marco/Reh, Sabine/Scholz, Joachim (Hrsg.): Fachunterrichtsgeschichten. Studien zur Geschichte der Praxis des Fachunterrichts. Bad Heilbrunn: Klinkhardt.

Warmington, Paul/Gorp, Angelo van/Grosvenor, Ian (2011): Education in motion: uses of documentary film in educational research. In: Paedagogica Historica 47, H. 4, S. 457–472.

Zinnecker, Jürgen (1987): Die Schule als Hinterbühne oder Nachrichten aus dem Unterleben der Schüler. In: Reinert, Gerd-Bodo/Zinnecker, Jürgen (Hrsg.): Schüler im Schulbetrieb. Berichte und Bilder vom Lernalltag, von Lernpausen und vom Lernen in den Pausen. Reinbek bei Hamburg: Rowohlt, S. 29–121.

De- und Rekontextualisierung in der Qualitativen Sekundäranalyse von Unterrichtsvideos

Anna Hamer, Jana Helbig und Michael Urban

Zusammenfassung: Der Artikel gibt einen Einblick in die Möglichkeiten der Qualitativen Sekundäranalyse von audiovisuellen Daten anhand der Analyse eines Unterrichtsvideos. Deutlich wird, dass Videos zwar eine Vielzahl von Kontexten in sich tragen, jedoch nicht sämtliche, die für eine Interpretation relevant sind. Die mit der Trennung von Datenerhebung und -analyse einhergehende Dekontextualisierung der Daten birgt methodologische Herausforderungen, denen durch eine Rekontextualisierung im Zuge der Datenanalyse begegnet werden kann. Es wird gezeigt, wie fehlende Kontextinformationen Interpretationen sowohl limitieren, als auch unter Nutzung eines Analyseverfahrens, das die situationale Herstellung von Bedeutung in den Blick nimmt, erleichtern kann, wenn dabei entsprechend auf die im Video hergestellten Kontexte Bezug genommen wird. Zugleich wird auf die Herausforderungen hinsichtlich der Zugänglichkeit von Archiven, forschungsethischer Fragen und der Erhebung von Kontextinformationen durch die Aufzeichnung selbst eingegangen.

Schlüsselwörter: Qualitative Sekundäranalyse, Unterrichtsvideos, Rekontextualisierung

1. Einleitung

Betrachtet man die Diskussion über die Nutzung und Nutzbarkeit von qualitativen Sekundärdatenanalysen (QSA) in den letzten 20 Jahren, so fällt auf, dass der Fokus hier überwiegend auf Interviewstudien zu liegen scheint, wenn auch zunehmend gefordert wird, ethnographisches Material und Bild-/Videomaterial in die Überlegungen zu den Nutzungsmöglichkeiten einzubeziehen (Wilke/Pröbrock/Pach 2019, S. 472 f.). Insbesondere im deutschsprachigen Raum scheinen bislang keine Publikationen zu existieren, die explizit auf die QSA von audiovisuellen Daten eingehen. Somit schließt unser Beitrag einerseits an die übergreifende Diskussion über QSA an, andererseits an die wenigen Arbeiten, die im englischsprachigen Raum zur QSA von Videomaterial vorliegen (Andersson/Sørvik 2013; Faniel/Frank/Yakel 2019; Frank et al. 2018; Kim/Suzuka/Yakel 2020). Im Fokus unseres Beitrags soll der Umgang mit Kontexten in der Datenauswahl

und -analyse stehen, der am Beispiel der Analyse einer Unterrichtsszene diskutiert wird. Hierfür werden wir zunächst die QSA von audiovisuellen Daten im Diskurs der QSA (vgl. Wilke i. d. B.) verorten, auf methodologische Überlegungen zum Umgang mit Kontexten eingehen und schließlich anhand unseres eigenen Forschungsprojektes Möglichkeiten der QSA von Videodaten herausstellen.

2. Sekundäranalysen audiovisueller Daten in der qualitativen Forschung

Qualitative Forschung zeichnet sich u. a. dadurch aus, dass im Zuge des Forschungsprozesses einzelne Aspekte aus der Fülle an Daten ausgewählt werden müssen, um mit diesen forschen zu können. Nach Bauder et al. (2021, S. 905) ist es deshalb für qualitative Forschung konstitutiv, dass in Forschungsprojekten Daten für weitere Forschung hinterlassen werden. Diese Fülle an in den Daten enthaltenen Informationen, die sich entsprechend unter sehr vielen verschiedenen Perspektiven und Fragestellungen betrachten lassen, trifft besonders auf audiovisuelle Daten zu (vgl. Jehle; Gruber/Rizzoli i. d. B.). In diesen ist, neben der sprachlich fixierten Kommunikation, auch die Anordnung und Bewegung von Körpern im Raum in ihren wechselseitigen Bezugnahmen enthalten, sodass sich insbesondere Unterrichtsvideos (vgl. Mempel/Winterscheidt; Herrle/Proske/Puzicha/Zimmer; Jehle i. d. B.), in denen viele verschiedene Personen(-Gruppen) miteinander interagieren, durch eine Überkomplexität auszeichnen. Im Analyseprozess muss daher zwangsläufig eine Selektion einzelner Perspektiven und Szenen vorgenommen werden (Beeli-Zimmermann/Wannack/Staub 2020, S. 260; Dinkelaker 2018, S. 154), sodass ein Großteil der Daten bei der Analyse unberücksichtigt bleibt. Inwiefern auch identische Videoszenen durch eine Rekontextualisierung zu anderen Daten werden können, diskutieren wir weiter unten.

Wird die QSA ethnographischer Daten insgesamt skeptisch gesehen, indem darauf hingewiesen wird, dass die Feldnotizen allein für die Nachnutzung wertlos sind, da Bedeutung erst »in Interaktion mit dem körperlichen Gedächtnisspeicher« (Hirschauer 2014, S. 304) und langfristigen, kulminierten Erfahrungen im Forschungsfeld hergestellt wird, trifft dies nicht auf audiovisuelle Daten zu: Sie können, in entsprechenden digitalen Formaten, mit anderen Forschenden geteilt und wiederholt angesehen werden (Andersson/Sørvik 2013, § 5), sodass komplexe Phänomene in kleinere Einheiten ›zerlegt‹ werden können (Blikstad-Balas 2017, S. 511). Gerade dieses Zerlegen größerer Handlungsvollzüge und die Möglichkeit, sich diese wiederholt ansehen zu können, ermöglicht es dann, dasselbe Material unter verschiedenen Perspektiven zu analysieren (ebd.). Zugleich erweist sich die Erhebung von Videomaterial insbesondere im Unterricht als herausfordernd, indem verschiedene institutionelle und technische Hürden zu überwinden sind. Es gilt abzuwägen, »ob das jeweilige Erkenntnisinteresse und die

Konstitution des Untersuchungsgegenstandes einen videodatenbasierten Zugang erfordern und ob ein solcher Zugang einen substantiellen Erkenntnisgewinn erwartbar macht« (Herrle 2020, S. 166). Somit treffen viele Chancen der QSA, etwa hinsichtlich der Forschungsökonomie, der Transparenz und Qualitätssicherung sowie des Eröffnens neuer Forschungsfelder (Hollstein/Strübing 2018, S. 2) besonders auf audiovisuelle Daten zu.

Bei audiovisuellen Daten handelt es sich grundsätzlich um besonders sensible Daten, die nur schwer anonymisierbar sind (Knoblauch/Wilke 2018, S. 53). Im Zuge der für diesen Beitrag fokussierten Unterrichtsvideos besteht die Herausforderung, dass auf diesen Aufzeichnungen eine Vielzahl an Personen zu sehen ist und eine vollständige Anonymisierung dieser nicht nur schwer umsetzbar ist, sondern auch die Videos inhaltslos und somit unbrauchbar machen würde (Frank et al. 2018, S. 56; Huschka/Oellers 2013, S. 2). Das Problem der Anonymisierbarkeit wird in den Repositorien i. d. R. über ein gestuftes Zugangsverfahren gelöst, sodass die Verantwortung für einen gewissenhaften Umgang mit nicht anonymisierbaren Daten an die Sekundärdatennutzenden übertragen wird. Die QSA von Unterrichtsvideos geht somit mit ethischen Herausforderungen einher, die nicht nur innerhalb der Repositorien, sondern auch in der Analyse selbst und den damit im Zusammenhang stehenden Veröffentlichungen berücksichtigt werden müssen.

3. Methodologische Positionen zum Umgang mit Kontexten in der QSA

In der Debatte um QSA wird insbesondere das (fehlende) Kontextwissen als relevant und als mögliche Begrenzung der Nutzbarkeit von Sekundärdaten herausgestellt: Ohne Para- und Metadaten, in denen etwa die Forschungssituation beschrieben wird, wird eine weitere Nutzung der Daten als kaum möglich bzw. sinnvoll erachtet (Huschka/Oellers 2013, S. 9 f.). Dies ist insbesondere deshalb bedeutsam, da in der QSA die Datenerhebung von der Auswertung entkoppelt ist, sodass die Sekundärdatennutzenden keinen direkten Zugriff auf den Kontext der Datenerhebung bzw. ihren Entstehungszusammenhang haben (Medjedović 2022, S. 274). Die Trennung von Datenerhebung und -analyse hat methodologische und forschungspraktische Konsequenzen.

Die Frage nach dem Kontext scheint im Rahmen der Sekundärnutzung von Videoaufzeichnung zunächst trivial, sind doch viele Kontextinformationen bereits im Datenmaterial selbst enthalten. Zugleich ist zu berücksichtigen, dass es sich auch bei Videoaufzeichnungen, wie sie etwa über Unterricht angefertigt werden, sowohl um natürliche Daten (vgl. Knoblauch/Wilke i. d. B.) als auch um Artefakte handelt: Zwar werden alltägliche Interaktionen im Klassenraum aufgezeichnet, sie sind nicht speziell für die Datenerhebung konstruiert, wie dies

z. B. bei Interviews der Fall ist, und viele der für die Auswertung relevanten Kontexte (z. B. Sitzordnung, Unterrichtsgegenstand und -material etc.) sind bereits Teil der Aufzeichnung. Dennoch wird das Material für einen bestimmten Zweck erhoben; Entscheidungen zu Kameraeinstellungen und -anordnung sowie dem Aufnahmezeitraum werden auf Grundlage des Erkenntnisinteresses der Primärstudie getroffen, sodass auch audiovisuelle Daten grundsätzlich selektiv und abhängig von den technischen Eigenschaften der Kamera, ihrer Positionierung und Führung während der Aufnahme sind (Beeli-Zimmermann et al. 2020, S. 259). Auch spezifische Kontexte, wie etwa atmosphärische Eindrücke, sind nur eingeschränkt in den Videos enthalten (Andersson/Sørvik 2013, § 9).

Vorschläge zur Systematisierung von Kontexten stammen überwiegend aus Überlegungen zur Archivierung dieser Informationen. Bishop (2006) differenziert im Zuge ihrer Überlegungen für die Archivierung für den Economic and Social Data Service für qualitative Daten des UK Data Archives (ESDS Qualidata) drei Kontextebenen für die Interviewarchivierung und betrachtet dabei die unterschiedlichen Zeitbezüge der Datenerhebung und Datenauswertung in der QSA: Der *interaktionale Kontext* bezieht sich zunächst auf die kommunikativen Prozesse im Zuge der Datenerhebung. Während Primärforschende hier auf vielfältiges Erfahrungswissen, das nicht aufgezeichnet wurde, zurückgreifen können und hiervon ausgehend im Analyseprozess Bedeutung konstruieren, geht dieser Prozess in der QSA zurück auf die festgehaltenen Artefakte, sodass sowohl Primär- als auch Sekundärforschende auf der Grundlage des verfügbaren kommunikativen Kontextes Bedeutung konstruieren (ebd., S. 12). Der *situationale Kontext* betrifft Angaben über die anwesenden Personen und ihre Verbindungen zueinander, das räumliche Setting, Körpersprache sowie über die Forschenden und Zeit und Ort der Erhebung und die Auswahl der beforschten Personen und ihre Informiertheit über das Forschungsprojekt (ebd., S. 12 f.). Als dritter Kontext wird der *Projektkontext* ausgewiesen, hierunter fallen Informationen über die Finanzierung, Forschungsziele und -fragen, methodologische Grundlagen und Projektberichte, Interviewleitfäden, Einverständniserklärungen, Sampling, Anonymisierung etc. Darüber hinaus sind auch kulturelle und historische Informationen zum Entstehungszusammenhang der Datenerhebung und zur Rezeption der Primärstudie von Relevanz (ebd.). Ähnliche Systematisierungen finden sich etwa bei Heuer et al. (2020) für die Datenbank Qualiservice (Uni Bremen), die zwischen der Mikroebene der Datenauswertung, der Ebene des Forschungsprojektes und zeitgeschichtlichen, regionalen oder kulturellen Kontexten unterscheiden (ebd., S. 3), sowie bei Medjedovic (2022) und Steinhardt et al. (2020).

In der Debatte um das Für und Wider der QSA und das für die Analyse notwendige Kontextwissen scheint Konsens darüber zu bestehen, dass Daten aus Perspektive der qualitativen Sozialforschung nicht als objektive Entitäten zu verstehen sind, die im Feld vorgefunden werden, sondern als »soziale und kontextuell eingebettete Produkte« (Medjedović 2022, S. 274). Die daran anschließenden

Diskussionen kreisen überwiegend um die Frage des methodologischen Verständnisses der Konstruktion von Daten und ob es vor diesem Hintergrund möglich ist, Daten und Kontext für die Archivierung voneinander zu trennen oder ob dies zu einer Dekontextualisierung der Daten führt.

Mauthner et al. (1998) entwickeln aus ihren Reflexionen über eigene Erfahrungen mit der erneuten Interpretation selbst erhobener Daten eine skeptische Perspektive auf QSA. Deutlich wird anhand der von ihnen skizzierten Versuche der Interpretation insbesondere die Kontextgebundenheit der erhobenen Daten und die Diskrepanz der eigenen persönlichen, theoretischen und intellektuellen Perspektiven zwischen der (länger zurückliegenden) Erhebungssituation und der Sekundäranalyse. Sie unterstreichen in ihrer Argumentation, dass Kontexte eben nicht als eigenständige Entitäten zu verstehen sind, sondern Kontext und Daten untrennbar miteinander verbunden sind (ebd., S. 742). Kritisiert wird das hinter der Annahme einer möglichen Kontextualisierung von Daten stehende positivistische Verständnis, demzufolge Daten durch Hinzufügen von archivierten Kontextinformationen ›repariert‹ und somit einer weiteren Analyse zugänglich gemacht werden können. Sie kommen zu dem Schluss, dass mit einer QSA zwar vorangegangene Arbeiten reflektiert und kritisiert werden können, stehen einer eigenständigen Analyse jedoch skeptisch gegenüber und sehen die Möglichkeiten der QSA bestenfalls in historischen und methodologischen Analysen.

Auch Hirschauer (2014) äußert Skepsis hinsichtlich der Möglichkeiten der QSA, indem er sich mit dem Begriff des ›Datums‹ auseinandersetzt und deutlich macht, dass Daten erst durch ihre Analyse entstehen. Er problematisiert die der qualitativen Sozialforschung inhärente Verschränkung von Erhebung und Analyse, die inkompatibel ist mit der in der QSA notwendigerweise praktizierten Trennung dieser Prozesse; durch die Archivierung von Rohdaten werden diese dekontextualisiert und verlieren an Bedeutung. Hirschauer kommt zu dem Schluss, dass es zwar grundsätzlich möglich ist, einzelne Daten für die Archivierung zu kontextualisieren, um sie verstehbar zu machen, der Aufwand hierfür jedoch im Vergleich zu einer neuen Erhebung unverhältnismäßig groß ist (ebd., S. 307 f.). Aus dieser Perspektive bewertet er den potenziellen Schaden, die Schädigung des Arbeitsbündnisses zu den Informant*innen im Zuge der »archivarische[n] Verewigung und unkontrollierbare[n] Öffnung seiner Äußerungen für unbekannte Dritte« (ebd., S. 309 f.) höher als ihren möglichen Nutzen.

In diesen Argumentationen wird zum einen deutlich, dass gerade die Verschränkung von Datenerhebung und -analyse, die konstitutiv für die qualitative Sozialforschung ist, zu Herausforderungen in der Kontextualisierung von Daten führt. Zum anderen zeigt sich, dass in der Debatte um den Kontext von Daten unterschiedliche Verständnisse davon vorliegen, was überhaupt ein ›Kontext‹ ist und welchen epistemologischen Stellenwert dieser im Vergleich zu ›Daten‹ hat.

Während die aufgeführten skeptischen Perspektiven hervorheben, dass der Kontext untrennbarer Teil der Daten ist und Bedeutung erst im Zuge der Analyse der Daten hergestellt wird, wird in anderen Publikationen, die stärker den praktischen Umgang mit (fehlenden) Kontexten fokussieren, deutlich, dass diese Diskussion kaum unabhängig von den methodologischen Annahmen der jeweiligen QSA selbst geführt werden kann. So sind für unterschiedliche Fragestellungen und Analyseverfahren unterschiedliche Kontextinformationen relevant, insbesondere der Objektiven Hermeneutik ist inhärent, dass sie ohne Kontextinformationen auskommt, sodass für einige Forschungsziele und Analysemethoden ausreichend scheint, sich auf den innerhalb der Daten selbst hergestellten Kontext zu beziehen (Bauder et al. 2021, S. 903; van den Berg 2005, § 41). Gerade unter dieser Perspektive, dass Daten erst im Zuge der Analyse und den damit in Zusammenhang stehenden methodologischen Positionen als solche hervorgebracht werden, scheint es plausibel, dass auch die Relevanz von Kontexten sich erst innerhalb der Analyse ergibt. Während der Analyse wird ersichtlich, welche Kontexte (nicht) relevant werden, sie lassen sich nicht vorab bestimmen (Bauder et al. 2021, S. 905 f.). Gleiches gilt für die Primäranalyse, auch hier wird erst innerhalb der Analyse deutlich, ob alle für das Projekt relevanten Kontexte überhaupt erhoben wurden (Fielding 2004, S. 99). Grundsätzlich lässt sich festhalten: »Eine Sekundäranalyse lässt sich nur durchführen, wenn unterstellt wird, dass Daten auch außerhalb ihres unmittelbaren Erhebungskontextes ausgewertet und interpretiert werden können« (Medjedović 2022, S. 274).

Die Perspektive, der in unserem Beitrag gefolgt werden soll, ist diejenige, dass die Konstruktion von Daten vorrangig während ihrer Analyse erfolgt. Anders als in der von Hirschauer (2014) eingenommenen Position, hieraus die Unmöglichkeit einer (vollständigen) Kontextualisierung abzuleiten, soll unter Rückgriff auf die Argumentation von Moore (2006) deutlich werden, dass gerade in der durch die Archivierung erfolgenden (partiellen) Dekontextualisierung von Daten und der damit notwendig werdenden Rekontextualisierung Chancen für die QSA liegen. Moore kritisiert, dass die Diskussion über Kontexte zu einseitig bezogen auf die Erhebungssituation geführt und auf die Bedeutung von Reflexivität im Prozess der Datenproduktion begrenzt wird. Ausgehend von der Überlegung, dass Daten erst innerhalb des Forschungsprojektes im Zuge der Analyse hergestellt werden, schlägt Moore vor, nicht über »reuse of pre-existing data« (ebd., S. 22) zu diskutieren, sondern über die Rekontextualisierung von Daten durch den neuen Forschungsprozess. Somit wird die QSA nicht als die Analyse bereits bestehender Daten konzeptionalisiert, sondern als Prozess der Rekontextualisierung und Rekonstruktion von Daten (ebd., S. 28). Inwiefern eine Rekontextualisierung von Daten produktiv in der QSA von audiovisuellen Daten genutzt werden kann, soll am Beispiel unserer Auseinandersetzung mit Unterrichtsvideos gezeigt werden.

4. Exemplifizierungen zum Umgang mit Kontexten in der QSA von Unterrichtsvideos

Die Relevanz von Kontexten und insbesondere die Chancen einer Rekontextualisierung sollen anhand unserer Erfahrungen in der QSA von Unterrichtsvideos aufgezeigt werden: Während die Datenauswahl insbesondere durch zur Verfügung stehende Informationen über Projektkontexte und damit in Verbindung stehende interaktionale Kontexte erleichtert wird, wird im Zuge der Datenauswertung deutlich, dass fehlendes Wissen über interaktionale und situationale Kontexte zwar einerseits bestimmte Interpretationen des Datenmaterials begrenzt, andererseits jedoch auch durch eine Rekontextualisierung Interpretationen ermöglicht, indem dadurch der analytische Blick auf die sich in der Aufzeichnung vollziehende interaktionale Herstellung von Bedeutung gerichtet wird.

4.1 Kontexte bei der Datenauswahl

Für die Auswahl geeigneter Unterrichtsaufzeichnungen recherchierten wir nach Datenbanken, die Unterrichtsvideos bereitstellen. Bislang sind Videodatenbanken überwiegend in disziplinären Archiven zu finden, sodass wir Zugänge zu den Archiven HILDEonline des CeLeB Hildesheim sowie des FDZ Bildung des DIPF (siehe Bambey et al. 2018; Porzelt et al. 2021) beantragten. In beiden Archiven lagen für die Forschung nutzbare Aufzeichnungen von Unterrichtsstunden vor, während in weiteren recherchierten Archiven der Fokus entweder auf der Fallarbeit im Kontext von Lehre, Studium und Weiterbildung lag (z. B. Archiv für pädagogische Kasuistik [ApaeK]) oder Videos lediglich in transkribierten Auszügen (z. B. Hallesches Fallarchiv) vorlagen. Deutlich wird daran der bisher wenig systematisierte Umgang mit Unterrichtsvideos, der den unterschiedlichen Nutzungskontexten sowie der Problematik der Anonymisierung von Videodaten geschuldet zu sein scheint. Um die Eignung des vorliegenden Materials für die Beantwortung unserer Fragestellung zu eruieren, sichteten wir stichprobenartig das Datenmaterial, nahmen jedoch keinen Kontakt zu den Primärforschenden auf. Die Einsichtnahme in das jeweilige Videomaterial verschiedener Studien beantragten wir nach Sichtung der vorhandenen Informationen zu den Studien (Kurzbeschreibung und darin enthaltene Informationen über Ziele, Anlage und Forschungsinstrumente); geeignete Videos wählten wir schließlich insbesondere nach den aufgezeichneten Inhalten aus. Unser Interesse lag zu Beginn des Forschungsvorhabens auf Prozessen der Konstitution von Verhaltensstörungen in unterrichtlichen Interaktionen sowie

der Problematisierung von Verhalten. Dies machte eine Kamerafokussierung auf Interaktionen zwischen der Lehrkraft und den Schüler*innen sowie den Schüler*innen untereinander erforderlich. Von Vorteil war somit eine Aufzeichnung des Unterrichts aus mindestens zwei Kameraperspektiven, um die Interaktionen nachvollziehen zu können. Deutlich wird hier der Zusammenhang zwischen den in den Videoaufzeichnungen sichtbaren Inhalten und den in den Studienbeschreibungen enthaltenen Fragestellungen und Perspektiven auf Unterricht (Dinkelaker 2018, S. 156 f.).

Die Unterrichtsaufzeichnungen, die wir schließlich für unsere Studie auswählten, stammen aus dem Projekt »Kompetenzerwerb und Lernvoraussetzungen« (KuL), das zwischen 2012 und 2015 an den Universitäten Bamberg, Mannheim und Berlin durchgeführt wurde und Ungleichheiten im Grundschulbereich und die damit in Verbindung stehenden Erwartungseffekte von Lehrkräften im Blick hatte (Kristen/Kogan/Stanat 2015). Ziel der Videoaufzeichnungen war es, das Verhalten der Schüler*innen in der Interaktion mit der Klassenlehrkraft zu erfassen, sodass wir davon ausgehen konnten, dass eben diese Interaktionen auf den Videos sichtbar sind. Der Datensatz umfasst insgesamt 41 Unterrichtsaufzeichnungen aus dem Deutsch- und Mathematikunterricht erster Klassen, je Klasse wurden aus jeweils zwei Kameraperspektiven i. d. R. zwei Unterrichts(doppel)stunden gefilmt.

Eine erste Sichtung des Videomaterials ergab, dass das vorhandene Datenmaterial zwar hinsichtlich der sichtbar werdenden Interaktionen angemessen für die Bearbeitung unserer Fragestellung ist, der aufgezeichnete Umfang von zwei Doppelstunden hierfür jedoch nicht ausreicht. Somit war unser geplantes Forschungsvorhaben mit den vorhandenen Daten nicht umsetzbar bzw. schwer begründbar: Auf Grundlage des kurzen aufgezeichneten Zeitraumes scheint es zwar möglich, die situative Herstellung und Reaktionen auf Unterrichtsstörungen zu beobachten, jedoch nicht die Konstitution und Zuschreibung einer Verhaltensstörung, die über einen längeren Zeitraum erfolgt. Im Sinne der Herstellung eines ›Data Fit‹ modifizierten wir unser Erkenntnisinteresse entsprechend zu der Frage, wie Unterrichtsstörungen interaktional hervorgebracht werden und wie auf diese reagiert wird.

Die Verknüpfung verschiedener Kontexte beeinflusste somit die Auswahl der Daten: Der Projektkontext ermöglichte eine erste Vorauswahl potenziell geeigneter Studien und die sich in den Projektkurzbeschreibungen andeutenden Verständnisse von Unterricht ließen Rückschlüsse auf die Kameraanordnung und -führung zu, die schließlich anhand des Videomaterials überprüft wurden, sodass hier ebenfalls die Ebene der interaktionalen Kontexte relevant wurde. Verfügbare Inhalte, Kontexte und die Fragestellung mussten aufeinander abgestimmt werden.

4.2 Kontexte bei der Analyse

Betrachtet man audiovisuelle Daten vor dem Hintergrund der bereits in den aufgezeichneten Daten enthaltenen Kontexte, wird ein Vorteil dieser Datenform für die QSA deutlich: Durch die aufgezeichnete Komplexität von Handlungsvollzügen kann davon ausgegangen werden, dass das Videomaterial selbst bereits eine Fülle von für die Analyse relevanten Kontextinformationen und Datenerhebungsbedingungen enthält bzw. von den Forschenden leicht anzugeben ist und somit innerhalb des Materials selbst bereits eine Vielzahl von Metadaten ermittelbar ist (Knoblauch/Wilke 2018, S. 49). Mehrere Kameraperspektiven ermöglichen die Fixierung von vielfältigen potenziell relevanten Kontexten, sodass weniger auf in separaten Dokumenten festgehaltene Kontexte zurückgegriffen werden muss, wie es etwa in der Erhebung von Frank et al. (2018, S. 55 f.) aus der Perspektive der befragten Sekundärforschenden als notwendig erachtet wird. Zugleich sind die Forschenden, anders als z. B. in Gruppendiskussionen, Interviews und ethnographischen Beobachtungen, weniger in den Forschungsprozess involviert, wenn sie ›natürliche Situationen‹ aufzeichnen und somit weniger Teil des zu dokumentierenden und archivierenden Kontextes (Andersson/Sørvik 2013, § 10; Hollstein/Strübing 2018, S. 3). Dennoch wurde in unserer Analyse auch deutlich, dass gewisse Kontexte, die für die Analyse der Daten hilfreich wären, in den Videos fehlten und somit unsere Interpretationen limitierten, sodass bestimmte Lesarten nicht weiterverfolgt werden konnten. Zugleich werden Chancen für die Interpretation deutlich, die mit dieser Dekontextualisierung und der notwendigen Rekontextualisierung einhergehen. Dies soll am Beispiel einer Szene eines von uns analysierten Unterrichtsvideos expliziert werden.

Methodische Grundlage unserer Analyse war die von Herrle (2013) vorgeschlagene mikroethnographische Interaktionsforschung, die mit ihren Bezügen zur Ethnomethodologie und Konversationsanalyse die prozesshafte Konstitution sozialer Realitäten, die innerhalb von Interaktionen lokal hergestellt werden und sich innerhalb dieser (re)aktualisieren, in den Blick nimmt (Herrle 2013, S. 120 f.; Herrle/Dinkelaker 2016, S. 91). Mit dieser Prämisse kommt das Verfahren mit relativ wenigen Kontextbezügen, die außerhalb der aufgezeichneten Interaktionen verortet sind, aus, (re)produziert sich Wirklichkeit doch gerade innerhalb der aufgezeichneten Interaktionen selbst. Somit sind etwa vorliegende Informationen zu den einzelnen Schüler*innen, beispielsweise zu ihrem Leistungsstand, ihrer sozialen Herkunft oder zur Lage der Schule, wie sie etwa von Sekundärforschenden in der Befragung von Frank et al. (2016, S. 2) als relevant angegeben werden, für unsere Fragestellung und unser methodisches Vorgehen weniger bedeutsam.

Entsprechend der Vorschläge der mikroethnographischen Interaktionsanalyse wurde ein Video einer Unterrichtsstunde zunächst segmentiert, um einen Überblick über die Interaktionsorganisation in der Unterrichtsstunde zu erhalten

und erste Muster herauszuarbeiten (Herrle 2020, S. 171; Herrle/Dinkelaker 2016, S. 101 ff.). Auf dieser Grundlage wurden einzelne Szenen sequenzanalytisch ausgewertet, um durch eine Rekonstruktion der zeitlichen Abfolge der aufgezeichneten Handlungen und das Anschließen dieser aneinander im Detail nachzuvollziehen, »wie die Beteiligten durch ihr Agieren an der Herstellung des jeweiligen Phänomens beteiligt sind« (Herrle/Dinkelaker 2016, S. 92). Eine dieser Szenen, in der ein Schüler und die ihn ein- und ausschließenden Interaktionen fokussiert werden, soll knapp beschrieben werden, um daran unseren Umgang mit Kontexten zu erläutern:

> *Inhalt der in einer ersten Klasse aufgezeichneten Szene zu Beginn des Mathematikunterrichts ist die Formierung eines Sitzkreises vor der Tafel, die von der Lehrkraft angeleitet wird. Hierfür ruft sie die Schüler*innen einzeln auf und weist ihnen Plätze zu, dabei sehen die meisten von ihnen aufmerksam nach vorne und warten darauf, aufgerufen zu werden. Ein Schüler legt, kurz nachdem seine Sitznachbarin, jedoch nicht er selbst, aufgerufen wurde, den Kopf auf den Tisch. Als die Lehrkraft schließlich mit einer ausladenden Armbewegung den Rest der Klasse aufruft, nach vorne zu kommen, bleibt er an seinem Platz sitzen und wird weder von der Lehrkraft daran erinnert, ebenfalls nach vorne zu gehen noch ergreift er selbst die Initiative. Im Sitzkreis ist die Lehrkraft bemüht, dass alle dort sitzenden Schüler*innen die Tafel sehen können: »aber ihr müsst euch doch so setzen, dass ihr das Bild sehen könnt«. Während dieser Äußerung steht der Schüler von seinem Platz auf und gähnt lautstark, setzt sich dann wieder hin. Die Lehrkraft fährt mit dem Unterrichtsgespräch fort; sie hat sich auf Augenhöhe der im Kreis sitzenden Schüler*innen begeben und spricht im Flüsterton mit ihnen, während diese ihren Blick auf die Tafel gerichtet haben. Der Schüler, der von seinem Sitzplatz aus dem Unterrichtsgespräch aufgrund der vor ihm sitzenden Mitschüler*innen und dem Flüsterton der Lehrkraft vermutlich nicht folgen kann, steht langsam von seinem Platz auf und rutscht auf Knien nach vorne in Richtung Sitzkreis, weder die Lehrkraft noch die Mitschüler*innen blicken zu ihm. Er schiebt sich, noch immer kniend, zwischen einem Stuhl und Tisch in den Sitzkreis hinein. Die Lehrkraft blickt auf und weist ihn mit einem Fingerzeig und den Worten »[Vorname]. Setz dich DA auf den Stuhl. Du bist echt (.) zu SPÄT« zurück auf einen Stuhl außerhalb des Sitzkreises und fährt mit dem Unterrichtsgespräch fort.*

Betrachtet man diese Szene unter der Fragestellung, wie die wechselseitigen Bezugnahmen, die hier lediglich knapp umrissen werden können, schließlich dazu führen, dass der Schüler nicht Teil des Sitzkreises werden darf, sondern einen Platz außerhalb dieses zugewiesen bekommt, lässt sich zunächst feststellen, dass nicht auf Kontextwissen aus vorausgehenden Szenen zurückgegriffen werden kann, in denen gegebenenfalls Konflikte zwischen dem Schüler und der Lehrkraft sichtbar werden, die das Handeln beider hätten erklären können. Es ist auch nicht bekannt, ob es eine Übereinkunft darüber gibt, dass der Schüler grundsätzlich nicht Teil des Sitzkreises werden darf, wodurch sich sowohl sein Verhalten,

indem er etwa nicht nachfragt, ob er bei der ausladenden Geste der Lehrkraft mitgemeint ist, als auch das der Mitschüler*innen, die ebenfalls nicht darauf hinweisen, dass er noch auf seinem ursprünglichen Platz sitzt, sowie das der Lehrkraft, die den Schüler weder erneut auffordert nach vorne zu kommen noch ihm die Möglichkeit gibt, zu einem späteren Zeitpunkt Teil des Sitzkreises zu werden, erklären ließe.

Dieses fehlende Kontextwissen birgt andererseits die Chance, die beobachtbaren Interaktionen für sich zu fokussieren und somit aus der aufgezeichneten Situation heraus zu rekontextualisieren. Weil keine Informationen zum Schüler oder den Beweggründen der Lehrkraft und ihren Interpretationen von dem Verhalten des Schülers, etwa durch dokumentierte Feldgespräche, vorhanden sind, wird der Nachvollzug der wechselseitigen Interaktionen, die schließlich dazu führen, dass der Schüler des Stuhlkreises verwiesen wird, zur einzigen Möglichkeit der Rekonstruktion dieser sozialen Prozesse. Aus dieser Perspektive lassen sich die Interaktionen zwischen der Lehrkraft und ihm als kommunikatives Missverstehen interpretieren, sodass das Verhalten von ihm sowie der Lehrkraft vor dem Hintergrund der dem Handeln zugrundeliegenden Erwartungen Sinn ergeben: Im Nachvollzug der Interaktionen wird deutlich, wie sich das beobachtbare Verhalten des Schülers, das zunächst als Performieren von Unaufmerksamkeit oder Verweigerung (z. B. Kopf auf den Tisch legen; sicht- und hörbares Gähnen) gelesen werden kann, und das Ignorieren dieses Schülers durch die Lehrkraft gegenseitig bedingen. Zusammen mit der Abweichung von ihrer anfänglichen Praxis, die Schüler*innen einzeln aufzurufen und ihnen Plätze zuzuweisen, bei der der Schüler in seiner Körperhaltung nicht sehen kann, ob er bei der Zeigegeste mitgemeint ist, mit der die Lehrkraft die restlichen Schüler*innen in den Kreis bittet, etabliert sich in der Interaktion zwischen ihm und der Lehrkraft ein Zusammenwirken von körperlichem Kleinmachen bzw. Resignation seitens des Schülers, ein Deuten als Unaufmerksamkeit bzw. Verweigerung seitens der Lehrkraft und damit einhergehendes Ignorieren, das die Resignation des Schülers weiter verstärkt. Zugleich führen seine Versuche, durch Aufstehen auf sich und seine prekäre Sitzposition aufmerksam zu machen, nicht dazu, dass er von der Lehrkraft als Teil des Sitzkreises angesprochen wird. Sein anfängliches Verhalten, das Legen des Kopfes auf den Tisch, nachdem zwar seine Sitznachbarin, nicht jedoch er, namentlich aufgerufen wird, scheint auf Grundlage der Reaktion der Lehrkraft, die den Schüler wiederholt nicht nach vorne bittet, als Unaufmerksamkeit oder Verweigerung gedeutet zu werden. Dies wird insbesondere am Ausruf der Lehrkraft, er sei »zu spät« und somit nicht zum ›richtigen‹ Zeitpunkt Teil des Sitzkreises geworden, durch den diese dem Schüler die Verantwortung für seine Situation zuschreibt und damit den Ausschluss aus dem Sitzkreis begründet, deutlich. Während der Schüler also als Reaktion auf das als Unaufmerksamkeit oder Verweigerung gelesene Ignorieren durch die Lehrkraft

und darauffolgende Resignation praktisch unsichtbar wird, indem er in seiner Körperhaltung zusammensinkt und sich schließlich nur auf Knien dem Unterrichtsgeschehen nähert, provoziert diese als Unaufmerksamkeit oder Verweigerung gelesene Reaktion wiederum ein Ignorieren des Schülers. Damit macht die Lehrerin ihn für seine Nicht-Teilnahme verantwortlich und exkludiert ihn aus dem Sitzkreis. Das Verhalten sowohl der Lehrkraft als auch des Schülers wird sinnhaft, wenn man die dahinterliegenden impliziten Erwartungen in den Blick nimmt. Aus dieser Perspektive kann hinter dem Verhalten des Schülers die Erwartung rekonstruiert werden, dass alle Kinder namentlich aufgerufen und ihnen ein Platz zugewiesen wird, und ein Nicht-Aufrufen somit einer fehlenden Erlaubnis zur Teilnahme am Sitzplatz gleichkommt, an die sich der Schüler zunächst hält und durch sein Aufstehen und Gähnen auf sich aufmerksam macht, bis er schließlich selbst die Initiative ergreift und in den Sitzkreis vorrückt. Demgegenüber lässt sich das Handeln der Lehrkraft auf die Erwartung zurückführen, dass der Schüler selbständig Teil des Kreises werden muss und entsprechend keinen Anspruch mehr auf Teilnahme hat, wenn er sich nicht in dem ›richtigen‹ Zeitfenster in den Sitzkreis bewegt.

Anhand dieses Beispiels konnte gezeigt werden, dass gerade die durch die Archivierung erfolgte Dekontextualisierung, indem etwa keinerlei Hintergrundinformationen zu möglichen vorausgegangenen Konflikten und zum Schüler selbst bekannt sind, es erleichtert, den situativen Vollzug der Handlungen aus sich selbst heraus zu erklären. Gerade unser Unwissen über Vorgeschichten und Bewegründe des Schülers sowie der Lehrkraft, wie sie etwa durch weitere Videoaufnahmen, Feldgespräche oder Interviews hätten erhoben und archiviert werden können, lenkt den Blick auf die sich vollziehenden Interaktionen, ohne diese vorschnell mit Deutungen aus dem Feld zu erklären. So wird ohne weiteres Kontextwissen durch den sequenziellen Nachvollzug der im Video beobachtbaren Interaktionen insbesondere die Verkettung von divergierenden Erwartungen und damit verbundenen Deutungen deutlich. Diese Interpretation ließ sich innerhalb des Videos durch den Vergleich mit ähnlichen Szenen, in denen der Schüler erneut versucht, am Unterricht teilzuhaben, jedoch von der Lehrkraft zurückgewiesen wird, bestätigen. Jedoch lassen sich dadurch, dass nur Videomaterial in begrenztem Umfang vorhanden ist, darüber hinaus keine Aussagen zu typischen Mustern in der Interaktion zwischen dem Schüler und der Lehrkraft finden; ähnliche Interaktionen mit anderen Schüler*innen werden in dem Video nicht erkennbar.

In unserem Projekt limitiert somit insbesondere der geringe Umfang der erhobenen Unterrichtsstunden und die damit einhergehende Entkontextualisierung der aufgezeichneten Interaktionen selbst eine weitere Differenzierung und Theoretisierung unserer Interpretationen. Liegen solche Einschränkungen vor, können Sekundäranalysen insbesondere als Vorstudie und Sensibilisierung für anschließende eigene Erhebungen genutzt werden.

5. Fazit und Ausblick

Anhand der Skizzierung einer Sequenz aus einem Unterrichtsvideo konnte gezeigt werden, dass gerade ein fehlender Kontext, hier insbesondere Informationen zu vorausgehenden Interaktionen bzw. Konflikten sowie zu den Beweggründen des Schülers und der Lehrkraft, Interpretationsmöglichkeiten eröffnen, jedoch auch limitieren. Während weiteres Videomaterial aus derselben Klasse eine Plausibilisierung bzw. Modifikation unserer Interpretationen hätte erlauben können, erleichtert der Zugang ohne diese Informationen eine Interpretation der Szene, ohne dass vorschnell Erklärungen für das Handeln der Akteur*innen gefunden werden. Übergreifend wurde deutlich, dass verfügbares Kontextwissen und insbesondere der interaktive Kontext bereits im Prozess der Datenauswahl bedeutsam werden, und die Problematik der Trennung von Erhebungs- und Auswertungssituation auch in der QSA audiovisueller Daten relevant ist. Dies gilt auch dann, wenn eine Vielzahl bedeutsamer Kontexte aus dem Video selbst erschlossen werden kann. Im Zuge einer bewussten Rekontextualisierung, die angeleitet wird durch ein entsprechendes methodisches Vorgehen, das die situationale Herstellung von Bedeutung in den Blick nimmt und dabei entsprechend auf die im Video hergestellten Kontexte Bezug nehmen kann, ist es jedoch grundsätzlich möglich, zu tragfähigen Interpretationen zu gelangen.

Zugleich wurden an der exemplarischen Darstellung der QSA eines Unterrichtsvideos mehrere Herausforderungen deutlich, die abschließend zusammengetragen werden. Sie betreffen erstens die Zugänglichkeit der Archive, zweitens forschungsethische Fragen sowie drittens den Anspruch, mit mehreren Kameraperspektiven möglichst viel Kontext durch die Aufzeichnung selbst erheben zu können.

Bislang fehlt es in der heterogenen Landschaft der Datenarchive an einer übergreifenden Datenbank oder zumindest Transparenz über verschiedene Repositorien und dort archivierte Inhalte, durch die der Zugang zu den Daten erleichtert werden könnte. Die sich aktuell von der Nationalen Forschungsdateninfrastruktur (NFDI) und dem Konsortium für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (Konsort SWD) im Aufbau befindliche Plattform QualidataNet versucht zukünftig, diesem Problem zu begegnen. Erschwerend waren die Zugangswege zu den Archiven höchst unterschiedlich gestaltet und dabei vermutlich der mangelnden Anonymisierbarkeit audiovisueller Daten geschuldet. Dies führte dazu, dass für jede potenziell geeignete Studie ein eigener Zugang beantragt werden musste und erst mit dem Vollzugriff die tatsächliche Geeignetheit des Materials abgeschätzt werden konnte. Die Beschreibungen der einzelnen Unterrichtsvideos selbst, wie sie für das Projekt KuL vorlagen, waren nur bedingt geeignet, um auf dieser Basis die Relevanz hinsichtlich

der aufgezeichneten Interaktionen zwischen Lehrkraft und Schüler*innen einschätzen zu können. Daran wird einerseits die Unmöglichkeit deutlich, für jedes hypothetische zukünftige Sekundärforschungsprojekt die relevanten Kontexte bereitzustellen und andererseits der zeitliche Aufwand, der gerade bei der Sichtung und Auswahl relevanter Videos berücksichtigt werden muss.

Aus forschungsethischer Perspektive kann problematisiert werden, dass die QSA von audiovisuellen Daten nicht nur durch die kaum mögliche vollständige Anonymisierung herausgefordert wird. Auch die technischen Möglichkeiten, einzelne kleine Szenen verlangsamt und wiederholt zu sichten und für die Interpretation zu nutzen, begrenzt die Umsetzbarkeit eines informed consent. Durch die »magnification of events« (Blikstad-Balas 2017, S. 516) im Zuge der Videoanalyse werden kleine Ausschnitte im Detail analysiert und dabei ihrem Kontext enthoben. Es kann weder weiteres Wissen über das Handeln der Lehrkraft und des Schülers berücksichtigt werden noch kann bewertet werden, ob das analysierte Handeln repräsentativ für die Interaktionen innerhalb der Klasse ist. Für die Interpretationen ist somit entscheidend, welche Situationen aufgezeichnet wurden und, daran anschließend, welche Materialausschnitte für die Interpretation ausgewählt und im Detail bearbeitet werden. Die Problematik, dass in qualitativer Forschung aufgrund der Offenheit des Forschungsprozesses nie ein vollständiger informed consent eingeholt werden kann, stellt sich jedoch in ähnlicher Form auch für Primärdatenerhebungen (Bishop 2009, S. 263).

Grundsätzlich scheint es von Vorteil, durch mehrere Kameraperspektiven möglichst viel Kontext, in unserem Fall das weitere Unterrichtsgeschehen, einzufangen, sodass dieser einer Interpretation zugänglich wird. Jedoch gilt auch hier das von Blikstad-Balas (2017, S. 515 f.) formulierte Kontextparadox, nach dem bei einer detaillierten Fokussierung auf Details der weitere Kontext verloren geht, während bei einer Betrachtung des weiteren Kontexts die Details der Szene nicht aufgezeichnet werden. Somit ist es auch mit dem Versuch, durch verschiedene Kameraperspektiven möglichst viel Kontext aufzuzeichnen, kaum möglich, alle gegebenenfalls relevanten Facetten festzuhalten: »While video data are often described as rich, they are also always partial and limited« (ebd., S. 515).

Abschließend kann resümiert werden, dass, unter Berücksichtigung der skizzierten Herausforderungen, das Problem der innerhalb der QSA erfolgenden Trennung von Erhebungs- und Analysesituation für audiovisuelle Daten, die grundsätzlich viele relevante Kontextmaterialen bereits in sich tragen, insbesondere vor dem Hintergrund einer Rekontextualisierung grundsätzlich produktiv gelöst werden kann. Nicht für jede QSA ist ein umfassendes Kontextwissen erforderlich, wie es etwa auch Bishop und Kuula-Luumi (2017, S. 9) formulieren, vielmehr kann fehlendes Kontextwissen neue Interpretationen erleichtern.

Literatur

Andersson, E./Sørvik, G. O. (2013): Reality Lost? Re-Use of Qualitative Data in Classroom Video Studies. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 14.

Bambey, D./Corti, L./Diepenbroek, M./Dunkel, W./Hanekop, H./Hollstein, B./Imeri, S./Knoblauch, H./Kretzer, S./Meyer, C./Meyermann, A./Porzelt, M./Rittberger, M./Strübing, J./Wilke, R. (Hrsg.) (2018): Archivierung und Zugang zu qualitativen Daten (RatSWD Working Papers).

Bauder, T./Brinkmann, S./Leser, C./Silkenbeumer, M./Terstegen, S. (2021): Qualität qualitativer Forschungsdaten im Kontext von Archivierung und nachhaltiger Nutzung. In: Zeitschrift für Pädagogik 67, H. 6, S. 899–914.

Beeli-Zimmermann, S./Wannack, E./Staub, S. (2020): Videobasierte Unterrichtsforschung: Aufnahmen mit zwei Kameras – und dann? In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 21.

Bishop, L. (2006): A Proposal for Archiving Context for Secondary Analysis. In: Methodological Innovation Online 1, H. 2, S. 10–20.

Bishop, L./Kuula-Luumi, A. (2017): Revisiting Qualitative Data Reuse: A Decade On. In: Sage Open 7, S. 1–15.

Blikstad-Balas, M. (2017): Key challenges of using video when investigating social practices in education: contextualization, magnification, and representation. In: International Journal of Research & Method in Education 40, H. 5, S. 511–523.

Dinkelaker, J. (2018): Reversible Selektivität. In: Heinrich, M./Wernet, A. (Hrsg.): Rekonstruktive Bildungsforschung. Wiesbaden: Springer Fachmedien, S. 141–158.

Faniel, I. M./Frank, R. D./Yakel, E. (2019): Context from the data reuser's point of view. In: Journal of Documentation 75, H. 6, S. 1274–1297.

Fielding, N. (2004): Getting the most from archived qualitative data: epistemological, practical and professional obstacles. In: International Journal of Social Research Methodology 7, H. 1, S. 97–104.

Frank, R. D./Suzuka, K./Yakel, E. (2016): Examining the Reuse of Qualitative Research Data: Digital Video in Education. In: Archiving Conference 13, S. 146–151.

Frank, R. D./Tyler, A. R. B./Gault, A./Suzuka, K./Yakel, E. (2018): Issues of Privacy in Qualitative Video Data Reuse. In: International Journal of Digital Curation 13, H. 1, S. 47–72.

Herrle, M. (2013): Mikroethnographische Interaktionsforschung. In: Friebertshäuser, B./Seichter, S. (Hrsg.): Qualitative Forschungsmethoden in der Erziehungswissenschaft: eine praxisorientierte Einführung. Weinheim und Basel: Beltz Juventa, S. 119–152 (Grundlagentexte Methoden).

Herrle, M. (2020): Koordination als Gegenstand mikroethnographischer Interaktionsforschung. In: Corsten, M./Pierburg, M./Wolff, D./Hauenschild, K./Schmidt-Thieme, B./Schütte, U./Zourelidis, S. (Hrsg.): Qualitative Videoanalyse in Schule und Unterricht. Weinheim und Basel: Beltz Juventa, S. 166–178.

Herrle, M./Dinkelaker, J. (2016): Qualitative Analyseverfahren in der videobasierten Unterrichtsforschung. In: Rauin, U./Herrle, M./Engartner, T. (Hrsg.): Videoanalysen in der Unterrichtsforschung: methodische Vorgehensweisen und Anwendungsbeispiele. Weinheim und Basel: Beltz Juventa, S. 76–129 (Grundlagentexte Methoden).

Heuer, J.-O./Kretzer, S./Mozygemba, K./Huber, E./Hollstein, B. (2020): Kontextualisierung qualitativer Forschungsdaten für die Nachnutzung: eine Handreichung für Forschende zur Erstellung eines Studienreports. Bremen (Qualiservice Working Papers 1-2020). https://media.suub.uni-bremen.de/handle/elib/4381 (07.06.2024).

Hirschauer, S. (2014): Sinn im Archiv? Zum Verhältnis von Nutzen, Kosten und Risiken der Datenarchivierung. In: Soziologie 43, H. 3, S. 300–312.

Hollstein, B./Strübing, J. (2018): Archivierung und Zugang zu qualitativen Daten. In: Bambey, D./Corti, L./Diepenbroek, M./Dunkel, W./Hanekop, H./Hollstein, B./Imeri, S./Knoblauch, H./Kretzer, S./Meyer, C./Meyermann, A./Porzelt, M./Rittberger, M./Strübing, J./Wilke, R. (Hrsg.): Archivierung und Zugang zu qualitativen Daten, S. 1–13 (RatSWD Working Papers).

Huschka, D./Oellers, C. (2013): Einführung: Warum qualitative Daten und ihre Sekundäranalyse wichtig sind. In: Huschka, D./Knoblauch, H./Oellers, C./Solga, H. (Hrsg.): Forschungsinfrastrukturen für die qualitative Sozialforschung. Berlin: Scivero, S. 9–18.

Kim, J./Suzuka, K./Yakel, E. (2020): Reusing qualitative video data: matching reuse goals and criteria for selection. In: Aslib Journal of Information Management 72, H. 3, S. 395–419.
Knoblauch, H./Wilke, R. (2018): Forschungsdateninfrastrukturen für audio-visuelle Daten der Qualitativen Sozialforschung Bedarf und Anforderungen. In: Bambey, D./Corti, L./Diepenbroek, M./Dunkel, W./Hanekop, H./Hollstein, B./Imeri, S./Knoblauch, H./Kretzer, S./Meyer, C./Meyermann, A./Porzelt, M./Rittberger, M./Strübing, J./Wilke, R. (Hrsg.): Archivierung und Zugang zu qualitativen Daten (RatSWD Working Papers).
Mauthner, N. S./Parry, O./Backett-Milburn, K. (1998): The Data are Out there, or are They? Implications for Archiving and Revisiting Qualitative Data. In: Sociology 32, H. 4, S. 733–745.
Medjedović, I. (2022): Qualitative Daten für die Sekundäranalyse. In: Baur, N./Blasius, J. (Hrsg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer Fachmedien, S. 271–282.
Moore, N. (2006): The Contexts of Context: Broadening Perspectives in the (Re)use of Qualitative Data. In: Methodological Innovation Online 1, H. 2, S. 21–32.
Porzelt, M./Bambey, D./Meyermann, A./Rittberger, M. (2021): Qualitative Forschungsdaten der Bildungsforschung archivieren und nutzen: Das Forschungsdatenzentrum (FDZ) Bildung am DIPF. In: Richter, C./Mojescik, K. (Hrsg.): Qualitative Sekundäranalysen. Wiesbaden: Springer Fachmedien, S. 111–129.
Steinhardt, I./Fischer, C./Heimstädt, M./Hirsbrunner, S. D./İkiz-Akıncı, D./Kressin, L./Kretzer, S./Möllenkamp, A./Porzelt, M./Rahal, R.-M./Schimmler, S./Wilke, R./Wünsche, H. (2020): Das Öffnen und Teilen von Daten qualitativer Forschung: Ergebnisse eines Workshops der Forschungsgruppe »Digitalisierung der Wissenschaft« am Weizenbaum Institut in Berlin am 17. Januar 2020. In: Weizenbaum Series.
van den Berg, H. (2005): Reanalyzing Qualitative Interviews from Different Angles: The Risk of Decontextualization and Other Problems of Sharing Qualitative Data. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 6.
Wilke, R./Pröbrock, W./Pach, H. (2019): Infrastrukturen für Forschungsdaten der qualitativen Sozialforschung. In: Soziologie 48, H. 4, S. 467–486.

Qualitative Sekundäranalyse

Zu Chancen und Herausforderungen der Nachnutzbarmachung qualitativer Daten für Forschung und Methodenlehre

René Wilke

Zusammenfassung: In diesem Artikel werden Innovationspotenziale durch die Sekundäranalyse als Mittel und Gegenstand der qualitativen Methodenlehre für die qualitative Forschung benannt. Vor dem Hintergrund der Digitalisierung werden dabei insbesondere Forschungsdatenzentren (FDZ) als Ressource beleuchtet. Anhand von Studien und Publikationen wird gezeigt, welche Rolle die Sekundäranalyse bereits jetzt für die Forschung und Methodenausbildung spielt. An Beispielen wird erläutert, welche konkreten Möglichkeiten die Sekundäranalyse Dozent*innen der Methodenlehre bietet. Es wird argumentiert, dass der Methodenlehre eine besondere Rolle dabei zukommt, eine Kultur des Datenteilens in der qualitativen Forschungscommunity zu etablieren und so daran mitzuwirken, die Innovationspotenziale der Sekundäranalyse für die qualitative Forschung voll auszuschöpfen.

Schlüsselwörter: Qualitative Sekundäranalyse, Methodenlehre, Forschungsdateninfrastrukturen, Kultur des Datenteilens

1. Einleitung

Spätestens seit den vergangenen 25 Jahren ist durch die Digitalisierung ein Transformationsprozess der qualitativen Forschung in Gang gekommen (Medjedović 2020, S. 81), wodurch Datenerhebungs-, Aufbereitungs- und Analyseprozesse sich erheblich gewandelt haben. Hintergrund dieser Entwicklung sind, neben der Transformation empirischer Felder, die zur Ausbildung einer großen Zahl neuer Datensorten geführt hat (z. B. *Vernakulare Videos*; s. Knoblauch/Wilke i. d. B.), die Digitalisierung von Aufzeichnungsgeräten (z. B. Videokameras), die Entwicklung und stete Weiterentwicklung von datensortenspezifischen Analysesoftwares (wie MAXQDA und Atlas.ti; vgl. Kramer/Hitzler i. d. B.) sowie die vereinfachte Interoperabilität digitaler Daten. So lassen diese sich z. B. zu komplexen *Komposit-Daten* (vgl. Knoblauch/Wilke i. d. B.) zusammenfügen

(vgl. Herrle/Proske/Puzicha/Zimmer; Reitstätter/Pesen/vom Lehn; Thiel-Woznica; Harth i. d. B.) Die Digitalisierung hat dabei auch die Entwicklung von Forschungsdateninfrastrukturen (FDI) für die Aufbereitung und Nachnutzung von qualitativen Daten erheblich gefördert (Huschka et al. 2013; Knoblauch/Wilke 2018; Wilke/Pröbrock/Pach 2019). Für die Forschung sowie die Lehre und das Lernen (qualitativer) Methoden ergeben sich daraus große Innovationspotenziale, die es auszuschöpfen, aber auch neue Herausforderungen, denen es offensiv zu begegnen gilt (vgl. Diekmann 2023).

Neben das technologische Dispositiv, die technische Vereinfachung des Forschungsprozesses und die Vorteile für Forscher*innen, tritt seit geraumer Zeit ein ausgeprägter forschungspolitischer Wille zur Sekundäranalyse (Wilke/Knoblauch 2021, S. 431). Die durch die Gemeinsame Wissenschaftskonferenz (GWK) initiierte und durch die Deutsche Forschungsgemeinschaft (DFG) gegenwärtig umgesetzte Nationale Forschungsdateninfrastruktur (NFDI) ist Ausdruck dieser Weichenstellung (DFG 2019). Ziel dieses Programms ist u. a., auch in der qualitativen Forschungsgemeinschaft eine »Kultur des Datenteilens« zu schaffen, d. h. die Nachnutzung von Forschungsdaten, die in der quantitativen Forschung seit langem üblich ist, auch in der qualitativen Forschung zu etablieren (Imeri 2018, S. 217). Dem stehen in der und für die Forschung zwar nach wie vor erhebliche Vorbehalte gegenüber (z. B. van den Berg 2008; Hirschauer 2014; Feldman/Shaw 2019). Während allerdings einige Wissenschaftler*innen dem Auf- und Ausbau von Infrastrukturen für qualitative Forschungsdaten nach wie vor skeptisch gegenüberstehen, sie ignorieren oder auf Risikopotenziale fokussieren, scheint heute, vor dem Hintergrund des vielerorts betriebenen, ressourcenaufwändigen Auf- und Ausbaus entsprechender Infrastrukturen, gewiss, dass auch die qualitative Forschung langfristig Infrastrukturen für die Nachnutzung von Daten in ihre Forschungspraxis inkorporieren wird. Vor diesem Hintergrund erscheint es geboten, diesen Transformationsprozess in der Methodenausbildung von Nachwuchswissenschaftler*innen zu verankern und die Potenziale des Auf- und Ausbaus von FDI, stärker als dies aktuell der Fall ist, bereits in der qualitativen Methodenlehre zu reflektieren und zu nutzen.

In diesem Artikel möchte ich daher FDI als Ressource für die Lehre qualitativer Forschungsmethoden thematisieren, insbesondere für die Sekundäranalyse als Mittel und Gegenstand der qualitativen Methodenausbildung. Dabei sollen insbesondere die Potenziale der Infrastrukturierung qualitativer Forschung und Lehre (Wilke/Knoblauch 2021, S. 432 ff.) besprochen werden. Dieser Artikel möchte einen Beitrag dazu leisten, nicht zuletzt auch Nachwuchsforscher*innen in der qualitativen Forschung mit diesen Potenzialen vertraut zu machen und einen diskursiven Impuls dafür bieten, die Entwicklung der Qualitativen Sekundäranalyse und entsprechender Infrastrukturen kritisch zu begleiten.

2. Forschungsdateninfrastrukturen (FDI) für qualitative Forschungsdaten

Wie inzwischen in zahlreichen Publikationen, darunter auch eine Reihe von Beiträgen in diesem Sammelband, herausgestellt wurde (u. a. Corti et al. 2000; Heaton 2004; Bergman/Eberle 2005; Corti et al. 2005; Valles et al. 2011; Medjedović 2014; Richter/Mojescik 2021; vgl. auch Gruber/Rizzolli; Jehle; Hamer/Helbig/Urban i. d. B.), besitzen qualitative Forschungsdaten, jenseits der Projekte, in deren Rahmen sie ursprünglich erhoben wurden, Nutzungspotenziale, die es auszuschöpfen gilt. Wenn auch betont werden muss, wie die einschlägige Literatur ebenfalls zeigt, dass sich, insbesondere aus datenschutzrechtlichen und forschungsethischen Gründen, sicherlich nicht alle qualitativen Forschungsdatenkorpora gleichermaßen für die Nachnutzbarmachung eignen, so bieten doch zahlreiche dieser Daten Eigenschaften, die für ihre Eignung zur Nachnutzbarmachung sprechen. Nachnutzende können so z. B. auf die aufwändige Erhebung von zusätzlichen Daten, z. B. zur Kontrastierung mit eigenen Daten, verzichten. Dadurch können Primärforschungsprojekte erheblich entlastet und die Nachhaltigkeit von durch öffentliche Fördermittel erhobene Forschungsdaten gesteigert werden. Nachnutzung ist außerdem für die ursprünglichen Erheber*innen der Daten von Nutzen, indem sie die Aufmerksamkeit für die geleistete Forschungsarbeit zu steigern vermag. Schließlich profitiert auch die Forschungsgemeinschaft von der Bereitstellung: Allgemein durch die Effizienzsteigerung der eingesetzten Ressourcen, die für jeden Zweck prinzipiell nur einmal zur Verfügung stehen können; konkret durch die, im Falle der Nutzung von FDI zur Nachnutzbarmachung, Akkumulation hochgradig elaboriert aufbereiteter Forschungsdatenkorpora, die der Forschungsgemeinschaft in Forschung und Lehre zur Verfügung stehen. Ein weiterer Nutzen besteht sicherlich auch in der Förderung der Güte qualitativer Forschung. So wird durch die Nachnutzbarmachung der Nachvollzug von Forschungsentscheidungen und -ergebnissen allgemein gefördert. Konkret stellt die Nachnutzbarmachung, wie ich weiter unten kurz ausführen werde, erhebliche Anforderungen an das Forschungsdatenmanagement (FDM), wodurch die Auseinandersetzung mit den eigenen Daten, die Erhebung, auch in Hinblick auf Datenschutz und Forschungsethik, sowie die Transparenz des Forschungsprozesses (Belegpflicht) erheblich gesteigert wird.

Obwohl die Nachnutzung von Forschungsdaten über spezialisierte FDI in der qualitativen Forschung noch immer vergleichsweise selten ist (Gebel/Meyermann 2021, S. 31), zeigen die Ergebnisse einer Professor*innenbefragung, die im Rahmen des DFG-geförderten Infrastrukturprojekts aviDa (vgl. Knoblauch/Wilke i. d. B.) durchgeführt wurde, dass eine Kultur des Datenteilens in der audiovisuellen Forschungscommunity, wenngleich erst in geringem Ausmaß etabliert, doch durchaus auf breite Zustimmung stößt. Befragungen im Bereich der

ethnographischen Fächer bestätigen diesen Befund (Imeri 2018). Entsprechende Infrastrukturlösungen stehen bereits umfangreich bereit; weitere befinden sich derzeit im Aufbau. So gibt es inzwischen in Deutschland, neben auf quantitative Daten ausgerichtete, auch bereits etablierte qualitative Forschungsdatenzentren (FDZ), die ihre Dienste Forscher*innen und Forschungsinstituten – explizit auch für die Lehre – zur Verfügung stellen.[1]

Eine wesentliche Bedeutung dieser FDZ besteht darin, spezialisierte Formen technischer Funktionalitäten bereitzuhalten, die den Besonderheiten qualitativer Datensorten, Erhebungs- und Auswertungsverfahren sowie ihren spezifischen Kontextdaten, Medienformaten und Forschungslogiken gerecht werden. Diese technischen Elemente werden von FDZ auf ihren digitalen Plattformen bereitgestellt. Ihre konkrete Gestaltung leitet sich von den jeweiligen Besonderheiten der Datensorten der qualitativen Forschung bzw. gegebenenfalls von der spezifischen Datensorte und/oder Fachdisziplin ab, auf die ein FDZ spezialisiert ist. Zusätzlich zu diesen technischen Einrichtungen, die als Infrastrukturen unmittelbar der Archivierung und Nachnutzung von Forschungsdaten dienen, bieten FDZ allerdings auch ausgebaute Beratungsangebote hinsichtlich des FDM an. Der Fokus liegt hierbei auf den Anforderungen hinsichtlich der Nachnutzbarmachung von Forschungsdaten. Dass entsprechende Beratungsprogramme (Beratung, Handreichungen, Workshops, Vorlagen) durch die FDZ angeboten werden, ist Ausdruck der großen methodologischen Expertise, die notwendig ist, um Forschungsdaten durch Dritte, die nicht am Primärforschungsprozess beteiligt waren, nachnutzbar zu machen. Diese Angebote erweisen sich in der Praxis als notwendig, weil das Teilen und Nachnutzen von Forschungsdaten i. d. R. in der Methodenausbildung bislang kaum eine Rolle spielten (bzw. bis heute spielen) und die entsprechenden Kenntnisse daher aktuell auch nicht vorausgesetzt werden können.

3. Qualitative Sekundäranalyse in der Methodenlehre

Während im vorangegangenen Kontext die Sekundäranalyse zuletzt als *Gegenstand* der Methodenlehre angesprochen wurde, gilt es diese von der Sekundäranalyse als *Mittel* der Methodenlehre zu unterscheiden. Zum einen deshalb, weil, wenngleich beides von großer Relevanz erscheint, so doch recht unterschiedliches gemeint ist. Wenn Erstere (Sekundäranalyse als Gegenstand) darauf abzielen sollte, neben den methodologischen Anforderungen der

1 Für eine Liste der auf spezifische Datensorten oder Fachdisziplinen spezialisierte RatSWD-akkreditierten FDZ siehe (RatSWD 2022, S. 54 ff.). Der Rat für Sozial- und Wirtschaftsdaten (RatSWD) übernimmt eine durch Akkreditierung, Monitoring und Beschwerdestelle strukturierte qualitätssichernde Aufgabe für den Auf- und Ausbau der deutschen FDI-Landschaft (Bug et al. 2018, S. 583 f.).

Nachnutzbarmachung von erhobenen Forschungsdaten, die Sekundäranalyse selbst zu thematisieren und in der Methodenlehre Student*innen als eigenständiges Verfahren der empirischen Forschung zur Kenntnis zu bringen, so ist mit Letzterer (Sekundäranalyse als Mittel) die Anwendung der Sekundäranalyse zum praktischen Nachvollzug und Einüben bestimmter qualitativer Erhebungs- bzw. Auswertungsverfahren in der Methodenausbildung gemeint. Für den Gesamtzusammenhang der Entwicklung und Etablierung einer Kultur des Datenteilens in der empirischen qualitativen Forschung sind beide Bedeutungen der Sekundäranalyse für die Methodenausbildung eng miteinander verknüpft.

Die gerade für die qualitative Forschung besonders anspruchsvollen forschungspolitischen Ziele zur Nachnutz(barmach)ung von Forschungsdaten sowie die Potenziale, die sich für die Forschungsgemeinschaft daraus ergeben können, sind nur dann zu verwirklichen, wenn Nachwuchsforscher*innen diese auch tatsächlich (forschungspraktisch wie *karriere*technisch) als komplementär zur Primärerhebung begreifen (dürfen). Vor diesem Hintergrund werde ich mich in Abschnitt 3.2 auch dem Aspekt der Sekundäranalyse als Gegenstand der Methodenlehre zuwenden. Zunächst aber möchte ich die schon heute weitaus größere Bedeutung der Sekundäranalyse als *Mittel* der Methodenlehre thematisieren. Studien zur Quantifizierung der Nachnutzungspraxis im Bereich der qualitativen bzw. Mixed-Methods-Forschung zeigen, dass man der Bedeutung der Nachnutzung entsprechender Forschungsdaten und korrespondierender FDI bzw. FDZ nicht auf die Spur kommt, ohne ihre (wenigstens für einige Länder im internationalen Vergleich belegte) Bedeutung in der Methodenausbildung zu berücksichtigen.[2]

3.1 Qualitative Sekundäranalyse als Mittel des Lehrens und Lernens

Eine Studie von Libby Bishop und Arja Kuula-Luumi (2017) zeigt, dass in Großbritannien zwischen den Jahren 2002 und 2016 lediglich 15 Prozent der Nachnutzungsfälle von qualitativen und Mixed-Methods-Forschungsdatenkorpora des *UK Data Services* auf Forschungsprojekte einschließlich Promotionsprojekte

2 Vor diesem Hintergrund soll an dieser Stelle darauf hingewiesen werden, dass dieser Nachnutzungskontext, der in der Debatte darüber, ob sich der hohe Ressourcenaufwand beim Auf- und Ausbau von FDI für qualitative Forschungsdatensorten, angesichts der noch geringen Nachnutzung in der Forschung, tatsächlich lohne, nicht übersehen werden und, im Sinne des humboldtschen Bildungsideals von der Einheit von Forschung und Lehre, nicht als weniger wertvoll für die Forschungscommunity erachtet werden kann. Im Kontext des Projekts aviDa konnte ich am Fachgebiet für Empirische Methoden der Sozialforschung (Leitung: Nina Baur) an der Technischen Universität Berlin selbst ein Seminar zur Sekundäranalyse von videographischen Forschungsdaten durchführen.

entfielen (ebd., S. 4), während der Großteil der Nachnutzungsfälle im Kontext des Lehrens (13,38 Prozent) und Lernens (63,72 Prozent) zu verorten sind (ebd.). Unter Lernen (»learning«) fassen die Autorinnen dabei solche Nutzungskontexte zusammen, die von Master- und Bachelorstudent*innen angefertigte Arbeiten darstellen und in diesem Sinn, wie die Lehre selbst, der Ausbildung von (zukünftigen) Nachwuchswissenschaftler*innen und gleichsam ersten Forschungsarbeiten zuzuordnen sind.

Wenn auch für einen kleineren Zeitraum, liegen aus derselben Studie vergleichbare Zahlen, die den Befund aus Großbritannien bestätigen, auch für Finnland vor: Zwischen Mai 2014 und Februar 2016 wurden die qualitativen Datenkorpora des *Finnish Social Science Data Archive* demnach zu 41 Prozent von Grundstudium-Student*innen, zu 28 Prozent für Masterarbeiten, zu 20 Prozent für die Lehre und lediglich zu 11 Prozent für Forschungszwecke genutzt (ebd., S. 6 f.).[3]

Die Arbeit von Bishop und Kuula-Luummi weist, über die reine Quantifizierung der Nutzung hinaus, auch auf die Innovationspotenziale für die Ausbildung in und die Entwicklung von qualitativer Forschung durch die Nachnutzung in der Lehre hin. Ihre Studie zeige z. B., so die Autorinnen, dass Studierende sich merklich leichter für Methodologie und das Erlernen einer Forschungsmethode begeistern ließen, wenn sie die Möglichkeit erhielten, diese an echten Forschungsdaten erlernen zu können (ebd., S. 6; siehe auch Haaker/Morgan-Brett 2017, S. 3). Anhand einer ganzen Reihe weiterer empirischer Beispiele zeigen sie die inhaltlichen und methodischen Innovationspotenziale auf, die Nachnutzung für die Methodenausbildung bereithält. Zugleich bestünden dabei Entwicklungspotenziale für FDZ, deren Nachfrage durch ihre Nutzung in der Lehre verstärkt und etabliert werde (Bishop/Kuula-Luummi 2017, S. 6; siehe auch Haaker/Morgan-Brett 2017, S. 8).

Auch Susanne Kretzer von Qualiservice Bremen betont das Innovationspotenzial von neuen »Lehr-Lern-Arrangements« (Kretzer 2013, S. 158 ff.), die durch die Bereitstellung von zur Nachnutzung aufbereiteten Forschungsdatenkorpora für die Methodenausbildung ermöglicht würden. Diese Arrangements erweisen sich als für ebenso unterschiedliche wie gleichermaßen wertvolle Zwecke geeignet. So ermöglichten sie, so Kretzer, Student*innen den forschenden Nachvollzug des Gesamtzusammenhangs einer empirischen Studie, den Vergleich mehrerer empirischer Studien, den Nachvollzug von Erhebungsmethoden oder das Einüben von Auswertungsverfahren. Ein offensichtlicher Vorzug der Nachnutzung von Forschungsdaten im Lehr-/Lernkontext betrifft den zeitlichen Aspekt: Da Seminaren, Lehrforschungsprojekten und studentischen Qualifikationsarbeiten besonders enge zeitliche Grenzen gesetzt sind, bleibt bei der Nachnutzung, gegenüber der Datenerhebung durch die Student*innen selbst, mehr

3 Vergleichbar detaillierte Zahlen liegen für den deutschsprachigen Raum bislang nicht vor.

Zeit für die Beschäftigung mit den Forschungsdaten, deren Aufbereitung sowie für die Analyse. Mit Maureen Haaker und Bethany Morgan-Brett lässt sich ergänzen, dass die Nachnutzung von Forschungsdaten Student*innen erlaubt, praktische Forschungserfahrungen zu machen, ohne im Rahmen der eingeschränkten Zeiträume von Lehrveranstaltungen zudem dem »enormous and ethically challenging task of collecting data in the field« gerecht werden zu müssen (Haaker/Morgan-Brett 2017, S. 8). Vielmehr böte sich ihnen durch die Nachnutzung die Gelegenheit, etwas über Forschungsethik und Datenschutz zu lernen und ihre Problemlösungs- und Entscheidungskompetenz zu entwickeln: »They also tap into the creativity needed to solve problems in situations where access, resources, or knowledge may be limited. Finally, students can better appreciate how even seemingly small decisions made within a larger research project can have an enormous impact on the interpretations and results of the research as a whole« (ebd.).

Wie Kretzer betonen auch Haaker und Morgan-Brett (2017) die vielfältigen Möglichkeiten, die sich für die Lehre durch die Nachnutzung von geeigneten, mit Kontextmaterialien angereicherten Forschungsdatenkorpora ergeben, um so Student*innen die Erfahrung der »real world of research methods« (ebd., S. 2) zu ermöglichen. Student*innen können so in die Lage versetzt werden, Forschungsprozesse durch den gesamten Forschungsdatenlebenszyklus, von der Planung und Erhebung, über die Aufbereitung und Analyse, bis zur Archivierung, Publikation und Nachnutzung von Forschungsdaten hindurch zu folgen, anstatt Forschung lediglich in Form von Forschungsberichten zu rezipieren (ebd., S. 3), denen i. d. R. bestenfalls eine hochgradig aggregierte Methodenreflexion zu entnehmen ist, die nicht dazu geeignet ist, Nachwuchsforscher*innen *in situ* mit den zahlreichen praktischen Herausforderungen der Datenanalyse vertraut zu machen.

Mit besonderem Augenmerk auf das Potenzial für die kritische Einschätzung von Forschungsprozessen durch Student*innen, z. B. hinsichtlich von Fragen der Güte wie der Gegenstandsangemessenheit bestimmter Methoden oder Datensorten, beschäftigt sich Helge Kminek vom Archiv für pädagogische Kasuistik (ApaeK) (Kminek 2022) mit der Nachnutzung von qualitativen Forschungsdaten in der Methodenlehre. Er erläutert, wie Lehramtsstudent*innen aber auch »Studierende in Wissenschaft im Allgemeinen« (ebd., S. 153) von der Nachnutzung der Forschungsdatenkorpora des Archivs profitieren könnten, indem sie sich anhand empirischer Materialien mit den Forschungsmethoden der Bildungsforschung kritisch auseinandersetzten und sich diese dabei aneigneten. So könnten auch Fragestellungen bezüglich der Einschränkungen einer Methode sowie deren praktischer Einfluss auf den Forschungsgegenstand in der Lehre an echten Forschungsdaten behandelt und veranschaulicht werden. Dabei, so Kminek, könne durch die Dozent*innen gezeigt und durch die Student*innen persönlich erfahren werden, was bestimmte Erhebungsmethoden und Datensorten zu leisten

vermögen und auch, wofür sie sich gegebenenfalls nicht eignen (siehe auch Haaker/Morgan-Brett 2017, S. 2).[4]

Der Artikel von Kminek (2022) verdeutlicht überzeugend, wie sich verschiedene wesentliche methodologische Fragestellungen, die sich in Vorbereitung eines empirischen Forschungsprojekts im Kontext der Bildungsforschung (aber ähnlich auch in anderen Forschungskontexten) typischerweise stellen, durch die Nachnutzung einer der verschiedenen durch das Archiv zur Verfügung gestellten Datensorten (ebd., S. 152) im Rahmen von Lehrveranstaltungen thematisieren lassen: »Mit welcher Erhebungsmethode kann welcher Ausschnitt aus der pädagogischen Wirklichkeit festgehalten werden? Verändert eine Erhebungsmethode die pädagogische Wirklichkeit, die festgehalten werden soll, und wenn ja wie? Wie können Theorien verifiziert oder falsifiziert werden? […] In welchem Kontext eignet sich welche Variante der Datenerhebung für welche Forschungsfrage?« (ebd., S. 153 f.). Zum Spektrum dessen, was anhand von nachgenutzten Forschungsdaten in der Lehre thematisiert werden kann, zählen folglich, im Sinne einer »Einführung in die Wissenschaftstheorie und in Forschungsmethoden« (ebd., S. 155), sowohl methodologische als auch erkenntnistheoretische Aspekte. Anhand der Nachnutzung von Forschungsdaten in der Methodenlehre, so Kminek, ließen sich Methoden illustrieren, Theorien überprüfen, neue Fragestellungen entwickeln, Methoden erlernen und dabei »Erkenntnismöglichkeiten und […] Grenzen von Forschungsmethoden reflektieren« (ebd.).

Auch Hubert Knoblauch (2007) weist im Kontext der Lehre qualitativer Methoden darauf hin, dass diese »keineswegs isoliert – also bloß als Methoden – gelehrt und erlernt werden [sollten], sondern als Teil von Forschungszusammenhängen« (ebd., § 19; siehe auch Stiefel 2007, § 1). Zu diesen Forschungszusammenhängen zählt er dabei sowohl den theoretischen Rahmen, insbesondere grundlegende methodologische Annahmen der qualitativen Forschung, als auch den empirischen Rahmen des qualitativen Forschungsprozesses: »Die Durchführung qualitativer Forschung setzt die Kenntnis durchgeführter Untersuchungen in den für die entsprechenden Methoden genutzten empirischen Forschungsbereichen voraus« (Knoblauch 2007, § 19). Neben Forschungsseminaren und der Lehrforschung empfiehlt er daher die Aufarbeitung exemplarischer Studien als Mittel der Methodenlehre. Forschungsdaten, als Objektivationen des Forschungsprozesses, spielen dabei die zentrale Rolle, um die Orientierung der Lehre und des Erlernens interpretativer qualitativer Forschung an empirischen Gegenständen zu gewährleisten.

4 Haaker und Morgan-Brett erläutern, dass Student*innen anhand eines durch den UK Data Service bereitgestellten Datenkorpus nachvollziehen können, wie das für eine Pilotstudie entworfene, ursprüngliche Interviewdesign nicht zu den von den Primärforscher*innen intendierten Resultaten führte und welche Anpassungen durch die Primärforscher*innen vorgenommen wurden, um die Ziele der Befragung schließlich umsetzen zu können.

3.2 Qualitative Sekundäranalyse als Gegenstand der Methodenlehre

Während im Abschnitt 3.1 die Sekundäranalyse als *Mittel* der Methodenlehre thematisiert wurde, möchte ich im Folgenden, wie bereits angekündigt, einen kurzen Abschnitt zur Sekundäranalyse als *Gegenstand* der Methodenlehre anfügen. Die Sekundäranalyse als Mittel in der Methodenlehre stellt, wie wir an einigen Beispielen sehen konnten, sicherlich großes Potenzial für die qualitative Methodenentwicklung dar. Diese Nutzungsweise ist jedoch in erster Linie a) ein Verfahren, um Student*innen etablierte Auswertungsmethoden der Primärforschung zu vermitteln. Zudem ist sie b) nur an der Nutzung orientiert, nicht aber an der Herstellung nachnutzbarer Daten. Daher sollte die Sekundäranalyse nicht nur als Mittel, sondern auch als Gegenstand Bestandteil der Syllabi sein. Zur Sekundäranalyse als Gegenstand zählt dabei die Nachnutzbarmachung als wesentliches Element und Grundvoraussetzung.

Tatsächlich zeigen Erfahrungen an FDZ, dass die Kenntnisse, die notwendig sind, um Forschungsdaten nachnutzbar zu machen, bei Primärforscher*innen zum Teil noch nicht stark ausgeprägt sind. Dass zahlreiche Korpora für eine Öffnung und das Teilen mit anderen Forscher*innen unzureichend vorliegen, zeigte sich nicht nur im Rahmen des Projekts aviDa und im Kontext von audiovisuellen Forschungsdaten, sondern ist auch an anderen FDZ bekannt (z. B. Wilke/Pröbrock/Pach 2019, S. 476). Dieser Umstand kann dabei nicht allein mit einem Ressourcenmangel in der Primärforschung erklärt werden. Zwar ist es richtig, dass die entsprechenden Fördermodule, die z. B. von der DFG bereitgestellt werden, häufig noch nicht in Anspruch genommen werden und daher in der Primärforschung auch Ressourcen fehlen, um erhobene Forschungsdaten, neben der Auswertung, noch aufzubereiten, um sie nachnutzbar zu machen.[5] Sehr häufig fehlt aber auch das Know-how. Dies ist kaum überraschend, denn die entsprechende Arbeit ist nicht weniger anspruchsvoll wie der qualitative Forschungsprozess an sich. Und ebenso wenig eignet sich die Nachnutzbarmachung, als elementarer Bestandteil der Sekundäranalyse, zur autodidaktischen Aneignung (Knoblauch 2007, § 18).

Als Voraussetzung dafür, diese Problematik zu lösen, bedarf es einer neuen Aufmerksamkeit in der Methodenlehre. Die Berücksichtigung grundlegender Erfordernisse der Nachnutzbarmachung, als Grundvoraussetzung dafür, dass eine wachsende Zahl von Forschungsdatenkorpora für die Nachnutzung in der

5 Vor besonderen Herausforderungen steht der Prozess zudem dann, wenn die Entscheidung zum Datenteilen retrospektiv getroffen wird. Häufig ist die Aufbereitung dann nur unter sehr erschwerten Umständen möglich, die im Rahmen eines vernünftigen Kosten-Nutzen-Kalküls i. d. R. dazu führen, dass ein entsprechendes Vorhaben verworfen werden muss. Umso wichtiger erscheint es, die Nachnutzbarmachung bereits bei der Projektentwicklung mit zu denken und als essenziellen Bestandteil des Forschungsdatenmanagements zu betrachten.

qualitativen Forschung bereitgestellt werden kann, trägt dabei nicht nur zur Einhaltung rechtlicher und forschungsethischer Standards bei (Gebel/Meyermann 2021, S. 25 f.), sondern, wie wir sehen werden, indirekt auch zur Weiterentwicklung der Güte qualitativer Forschung insgesamt.

Zu den essenziellen Kompetenzen, das hat sich u. a. im Kontext der Entwicklung des FDZ aviDa deutlich gezeigt, zählen *erstens* umfangreiche Kenntnisse datenschutzrechtlicher und forschungsethischer Natur. Die Digitalisierung und die scheinbare ubiquitäre Verfügbarkeit (vernakularer) Daten in den Sozialen Medien dürfen nicht darüber hinwegtäuschen, dass die Erzeugung von Forschungsdaten, trotz gewisser Ausnahmeregelungen, rechtlichen und ethischen Auflagen unterliegt, die ohnehin zu berücksichtigen sind, aber besonders transparent werden, wenn man sich mit der Nachnutzbarmachung eines Datenkorpus beschäftigt. Ist die informierte Einwilligung z. B. schon rechtlich wie forschungsethisch obligatorisch, ist sie für die Nachnutzbarmachung über ein FDZ i. d. R. eine Grundvoraussetzung.

Ein *zweiter* Kompetenzbereich berührt organisationale und repräsentative Aspekte des FDM, die z. B. die Datenablage, -speicherung und -aufbereitung betreffen. Selbst scheinbar so triviale Aspekte wie die Ordnerstruktur bei der Speicherung von Forschungsdaten und die Dateinomenklatur können ein praktisch unüberwindliches technisches Hindernis für die Nachnutzbarmachung bzw. die Sekundäranalyse eines inhaltlich wertvollen Datenkorpus darstellen. Dies gilt gerade im Rahmen qualitativer Forschungsprojekte, mit ihren häufig diversen und in einer Vielzahl von Dateiformaten kleinteilig abgespeicherten Datensorten: von verschiedensten Textsorten (wie z. B. kurze Feldnotizen, Gedächtnisprotokolle oder seitenlange Feldbeschreibungen) über Bild- und Videodateien bis hin zu Skizzen, Mappings und erhobenen Dokumenten verschiedenster Natur (einschließlich natürlich der unterschiedlichsten Transformationsformen all dieser Daten).

Ein *dritter* (und an dieser Stelle letzter) Aspekt, den die Lehre stärker berücksichtigen sollte, ist die Notwendigkeit der umfangreichen Dokumentation der Erhebungs- und Analysekontexte von Forschung. Auch diese werden nicht erst durch die Nachnutzung relevant, sondern sind, im Sinne der Belegpflicht als Hauptelement guter wissenschaftlicher Praxis, schon immer von größter Bedeutung. Erst die systematische Aufbereitung und Bereitstellung entsprechender Materialien, wie z. B. Datenmanagementpläne, Studienberichte, umfangreich geführte Dokumentationen sowie nachvollziehbare Feldnotizen, -protokolle etc. ermöglichen den Nachvollzug von Forschungsprozessen, Forschungsentscheidungen und Forschungsergebnissen, nicht allein im Rahmen der Nachnutzbarmachung.

An dieser hier zugegebenermaßen stark verkürzten Darstellung dreier Kompetenzbereiche zum Erwerb der Fähigkeit, selbst Forschungsdaten nachnutzbar zu machen, wird ein interessanter Aspekt deutlich, der einen potenziellen

Zusammenhang zwischen einer ›Kultur des Datenteilens‹ und der Güte qualitativer Forschung verdeutlicht: Die genannten Kompetenzen sind nicht von genuiner Bedeutung für die Nachnutz(barmach)ung von Forschungsdaten, sondern stellen, wie gesagt, generelle Erfordernisse guter wissenschaftlicher Praxis dar. Die im Entstehen begriffene Kultur des Datenteilens scheint also vielmehr, wobei sie (vielleicht gerade deshalb) nachvollziehbar auch das Potenzial besitzt, als Ärgernis wahrgenommen zu werden, Licht auf Elemente des qualitativen Forschungsprozesses zu werfen, die eventuell – gerade vor dem Hintergrund der engen Verknüpfung der qualitativen Forschung mit der Leibkörperlichkeit der Forscher*innen – bislang weniger stark beleuchtet wurden. Die aktuelle Initiative, das Datenteilen auch in der qualitativen Forschung zu etablieren, kann in diesem Sinne als Beitrag zur weiteren Fortentwicklung der Güte qualitativer Forschung betrachtet werden. Daher sollte sie sich auch auf den Lehrplänen für die qualitative Methodenausbildung verstärkt wiederfinden.

4. Innovationspotenziale gemeinsam erkennen, entwickeln und nutzen

Mit diesem Beitrag soll ein Anstoß gegeben werden, Dozent*innen der qualitativen Methoden an Hochschulen und Universitäten, Forscher*innen sowie Student*innen, auf Potenziale und Herausforderungen der Qualitativen Sekundäranalyse aufmerksam zu machen. Nicht zuletzt auch, um die Innovationspotenziale des ressourcenintensiven Auf- und Ausbaus von FDI für qualitative Forschungsdaten voll auszuschöpfen, ist die Thematisierung bzw. Beschäftigung mit der Sekundäranalyse als Mittel und als Gegenstand in Forschung und Lehre von Bedeutung.

In Form einer rasch anwachsenden Zahl von Korpora qualitativer Forschungsdaten, die von Nachwuchswissenschaftler*innen, die die Sekundäranalyse als vielversprechendes Verfahren der qualitativen Forschung in ihrer Methodenausbildung kennen und zu begreifen gelernt haben, zukünftig erhoben und nachnutzbar gemacht werden könnten, kann sich diese Anstrengung für die gesamte qualitative Forschungscommunity auszahlen. Eine engere Verzahnung der Methodenausbildung mit den Serviceangeboten der FDZ sowie den übergeordneten Vernetzungsstrukturen, die derzeit im Rahmen der NFDI aufgebaut werden, stellt dabei sicherlich einen fruchtbaren Ausgangspunkt dar. Zahlreiche der in diesem Rahmen derzeit im Aufbau befindlichen Angebote für qualitative Forscher*innen sind auch für die Methodenlehre geeignet und nützliche Hilfestellungen dabei, Student*innen mit den Herausforderungen, aber auch den Innovationspotenzialen der Sekundäranalyse, vertraut zu machen.

Verantwortliche der Methodenausbildung und Dozent*innen, selbst i. d. R. Forscher*innen, sowie Forscher*innen im Allgemeinen, dürfen bei dem Prozess

der Infrastrukturierung der qualitativen Forschung nicht außenvorgelassen und sollten verstärkt in die aktuelle Entwicklung eingebunden werden. Die Ziele der forschungspolitischen Initiative zur Etablierung einer Kultur des Datenteilens können sonst nur verfehlt werden. Eine wachsende Zahl von einschlägigen Publikationen zeigt allerdings, dass die Entwicklung Fahrt aufnimmt. Besonders hervorzuheben sind dabei Arbeiten, die die Qualitative Sekundäranalyse unmittelbar aus der Praxisperspektive thematisieren und dabei direkt aus der Nutzung bestehender Infrastrukturen heraus auch konkrete Ansprüche an FDZ formulieren (Houben/Eckert 2022; Hamer/Helbig/Urban i. d. B.).

5. Fazit

In seinem Problemaufriss »Neuorientierung der Methoden-Ausbildung« diagnostiziert Andreas Diekmann (2023) (mit Schwerpunkt auf die quantitative Soziologie) einen ›kulturellen Rückstand‹ »zwischen Inhalten von Methodenkursen und dem Stand der Forschung« (ebd., S. 68). Diekmann schreibt, diese Lücke sei »auch durch den rasanten Fortschritt von Digitalisierung und Technologieentwicklung bedingt« und fordert: »Das bisherige Programm von Methodenkursen muss modernisiert und an diese Entwicklungen angepasst werden« (ebd.). Diekmanns Forderung lässt sich problemlos auch auf den vorliegenden Kontext übertragen, zumal er auch die Thematik der »open science« anspricht (ebd., S. 71). Im Kontext der Qualitativen Sekundäranalyse und der Nutzung hierfür entwickelter Infrastrukturlösungen besteht der Rückstand auf mehreren Ebenen. Einerseits sind weite Teile der qualitativen Forschungscommunity zwar bereit, eine Kultur des Datenteilens zu integrieren. Andererseits hinken die Absichtsbekundungen, wie Studien und praktische Erfahrungen an den FDZ zeigen, der Wirklichkeit noch hinterher. Und obgleich die Nachnutzung von Forschungsdaten zwar in der Methodenlehre angekommen ist, wird das volle Potenzial der digitalen Möglichkeiten noch nicht ausgeschöpft. Auch spielt die Sekundäranalyse als eigenständiges Verfahren, als Gegenstand der Methodenlehre im eigenen Recht, bislang eine völlig untergeordnete Rolle.

Mit diesem Artikel soll darauf hingewiesen werden, dass Sekundäranalyse und Nutzung von entsprechenden Infrastrukturangeboten großes Innovationspotenzial für die Methodenlehre und die Fortentwicklung der qualitativen Forschung (nicht zuletzt im Sinne ihrer Güte) besitzen. Neben die zahlreichen Möglichkeiten, die sich der qualitativen Forschung und Methodenlehre durch die Sekundäranalyse bieten, treten ihre Herausforderungen. Um sich durchzusetzen, setzt sie voraus, das gewusst wird, was nötig ist, um Forschungsdaten überhaupt erst nachnutzbar zu machen. Das entsprechende Wissen muss daher auf den Methodenlehrplänen Einzug halten.

Literatur

Bergman, Manfred M./Eberle, Thomas S. (Hrsg.) (2005): Qualitative Forschung, Archivierung, Sekundärnutzung: Eine Bestandsaufnahme. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 6, 2. http://www.qualitative-research.net/index.php/fqs/issue/view/12.

Bishop, Libby/Kuula-Luumi, Arja (2017): Revisiting Qualitative Data Reuse: A Decade On. In: Sage Open January-March, S. 1–15. https://doi.org/10.1177/2158244016685136.

Bug, Mathias/Liebig, Stefan/Oellers, Claudia/Riphahn, Regina T. (2018): Operative und strategische Elemente einer leistungsfähigen Forschungsdateninfrastruktur in den Sozial- und Wirtschaftswissenschaften«. In: Journal of Economics and Statistics 238, 6, S. 571–590. https://doi.org/10.1515/jbnst-2018-0029.

Corti, Louise/Kluge, Susanne/Mruck, Katja/Opitz, Diane (Hrsg.) (2000): Text. Archiv. Re-Analyse. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 1, 3. https://www.qualitative-research.net/index.php/fqs/issue/view/27.

Corti, Louise/Witzel, Andreas/Bishop, Libby (Hrsg.) (2005): Sekundäranalyse qualitativer Daten. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 6, 1. http://www.qualitativeresearch.net/index.php/fqs/issue/view/13.

DFG (2019): Information für die Wissenschaft Nr. 9. Ankündigung: Nationale Forschungsdateninfrastruktur, 11. Februar 2019.

Diekmann, Andreas (2023): Neuorientierung der Methoden-Ausbildung. In: Soziologie 52, 1, S. 68–71.

Feldman, Shelley/Shaw, Linda (2019): The Epistemological and Ethical Challenges of Archiving and Sharing Qualitative Data«. In: American Behavioral Scientist 63, 6, S. 699–721. https://doi.org/10.1177/0002764218796084.

Gebel, Tobias/Meyermann, Alexia (2021): Sekundäranalysen qualitativer Interviews: Eine Metaanalyse zur Praxis sekundäranalytischer Forschung zu Arbeitsorganisationen. In: Richter, Caroline/Mojescik, Katharia (Hrsg.): Qualitative Sekundäranalysen. Wiesbaden: Springer VS, S. 17–36. https://doi.org/10.1007/978-3-658-32851-1_2.

Haaker, Maureen/Morgan-Brett, Bethany (2017): Developing Research-Led Teaching: Two Cases of Practical Data Reuse in the Classroom. In: Sage Open April-June, S. 1–9. https://journals.sagepub.com/doi/10.1177/2158244017701800.

Heaton, Janet (2004): Reworking Qualitative Data. London: Sage.

Hirschauer, Stefan (2014): Sinn im Archiv? Zum Verhältnis von Nutzen, Kosten und Risiken der Datenarchivierung. In: Soziologie 43, 3, S. 300–312. https://publikationen.soziologie.de/index.php/soziologie/article/view/795.

Houben, Malin/Eckert, Judith (2022): Die Arbeit mit archivierten Interviewdaten in einem methodologischen Sekundärforschungsprojekt: Reflexionen zur Archivierung qualitativer Forschungsdaten. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 23, 1, Art. 22. https://www.qualitative-research.net/index.php/fqs/article/view/3701/4841.

Huschka, Denis/Knoblauch, Hubert/Oellers, Claudia/Solga, Heike (Hrsg.) (2013): Forschungsinfrastrukturen für die Qualitative Sozialforschung. Berlin: SCIVERO. https://www.konsortswd.de/wp-content/uploads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.

Imeri, Sabine (2018): Ordnen, archivieren, teilen. Forschungsdaten in den ethnologischen Fächern. In: Österreichische Zeitschrift für Volkskunde LXXII/121, 2, S. 213–243. https://doi.org/10.18452/21504.

Kminek, Helge (2022) Datenarchive in Lehre und Studium. Zur Nutzung vergessener Schatzkammern. In: Egloff, Birte/Richter, Sophia (Hrsg.): Erziehungswissenschaftlich denken und arbeiten. Ein Lehr- und Studienbuch. Stuttgart: Kohlhammer, S. 149–158.

Knoblauch, Hubert (2007): Thesen zur Lehr- und Lernbarkeit qualitativer Methoden. Diskussionsbeitrag zur FQS-Debatte »Lehren und Lernen der Methoden qualitativer Sozialforschung«. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 8, 1. https://www.qualitative-research.net/index.php/fqs/article/view/217/479.

Knoblauch, Hubert/Wilke, René (2018): Forschungsdateninfrastrukturen für audiovisuelle Daten der Qualitativen Sozialforschung – Bedarf und Anforderungen. In: RatSWD Working Paper 267, S. 47–58. https://www.konsortswd.de/wp-content/uploads/RatSWD_WP_267.pdf.

Kretzer, Susanne (2013): Vom Nutzen des Datasharing für die Lehre in der qualitativen Sozialforschung. In: Huschka, Denis/Knoblauch, Hubert/Oellers, Claudia/Solga, Heike (Hrsg.): Forschungsinfrastrukturen für die Qualitative Sozialforschung. Berlin: SCIVERO, S. 153–166. https://www.konsortswd.de/wp-content/uploads/forschungsinfrastrukturen_qualitative_sozialforschung.pdf.

Medjedović, Irena (2014): Qualitative Sekundäranalyse. Zum Potenzial einer neuen Forschungsstrategie in der empirischen Sozialforschung. Wiesbaden: Springer VS.

Medjedović, Irena (2020): Qualitative Sekundäranalyse. In: Mey, Günther/Mruck, Katja (Hrsg.): Handbuch Qualitative Forschung in der Psychologie. Band 2: Designs und Verfahren. 2. Auflage. Wiesbaden: Springer VS, S. 79–95.

RatSWD (2022): Tätigkeitsbericht 2020 der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ). Berlin: Rat für Sozial- und Wirtschaftsdaten (RatSWD). https://doi.org/10.17620/02671.65.

Richter, Caroline/Mojescik, Katharina (Hrsg.) (2021): Qualitative Sekundäranalysen. Daten der Sozialforschung aufbereiten und nachnutzen. Wiesbaden: Springer VS.

Stiefel, Britta (2007): »Der Einsatz archivierter Forschungsdaten in der qualitativen Methodenausbildung – Konzept und Evaluation eines Pilotmodells für forschungsnahes Lernen. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 8, 3: Art. 15. https://doi.org/10.17169/fqs-8.3.292.

Valles, Miguel S./Corti, Louise/Tamboukou, Maria/Baer, Alejandro (Hrsg.) (2011): Qualitative archives and biographical research methods. In: Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 12, 3. http://www.qualitative-research.net/index.php/fqs/issue/view/38.

van den Berg, Harry (2008): Reanalyzing qualitative interviews from different angles: The risk of decontextualization and other problems of sharing qualitative data. In: Historical Social Research/Historische Sozialforschung 33, 3, S. 179–192. https://doi.org/10.12759/hsr.33.2008.3.179-192.

Wilke, René/Knoblauch, Hubert (2021): Faire Prinzipien – für die Nachnutzung von Forschungsdaten und die Primärforschung«. In: Soziologie 50, 4, S. 431–438.

Wilke, René/Pröbrock, Willi/Pach, Helen (2019): Infrastrukturen für Forschungsdaten der qualitativen Sozialforschung. In: Soziologie 48, 4, S. 467–486.