

Hess, Miriam; Lipowsky, Frank  
**Zur (Un-)Abhängigkeit von Oberflächen- und Tiefenmerkmalen im  
Grundschulunterricht. Fragen von Lehrpersonen im öffentlichen Unterricht  
und in Schülerarbeitsphasen im Vergleich**

*Praetorius, Anna-Katharina [Hrsg.]; Grünkorn, Juliane [Hrsg.]; Klieme, Eckhard [Hrsg.]: Empirische  
Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen. 1. Auflage.  
Weinheim; Basel : Beltz Juventa 2020, S. 117-131. - (Zeitschrift für Pädagogik, Beiheft; 66)*



Quellenangabe/ Reference:

Hess, Miriam; Lipowsky, Frank: Zur (Un-)Abhängigkeit von Oberflächen- und Tiefenmerkmalen im Grundschulunterricht. Fragen von Lehrpersonen im öffentlichen Unterricht und in Schülerarbeitsphasen im Vergleich - In: Praetorius, Anna-Katharina [Hrsg.]; Grünkorn, Juliane [Hrsg.]; Klieme, Eckhard [Hrsg.]: Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen. 1. Auflage. Weinheim; Basel : Beltz Juventa 2020, S. 117-131 - URN: urn:nbn:de:0111-pedocs-258689 - DOI: 10.25656/01:25868

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-258689>

<https://doi.org/10.25656/01:25868>

in Kooperation mit / in cooperation with:

**BELTZ JUVENTA**

<http://www.juventa.de>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

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](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft

66. Beiheft

April 2020

# **ZEITSCHRIFT FÜR PÄDAGOGIK**

---

**Empirische Forschung zu Unterrichts-  
qualität. Theoretische Grundfragen und  
quantitative Modellierungen**

**BELTZ** JUVENTA

Zeitschrift für Pädagogik · 66. Beiheft



Zeitschrift für Pädagogik · 66. Beiheft

# **Empirische Forschung zu Unterrichtsqualität**

**Theoretische Grundfragen  
und quantitative Modellierungen**

Herausgegeben von  
Anna-Katharina Praetorius, Juliane Grünkorn  
und Eckhard Klieme

**BELTZ** JUVENTA

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, bleiben dem Beltz-Verlag vorbehalten.

Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genutzte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, bei der die einzelnen Zahlungsmodalitäten zu erfragen sind.



ISSN: 0514-2717

ISBN 978-3-7799-3534-6 Print

ISBN 978-3-7799-3535-3 E-Book (PDF)

Bestellnummer: 443534

1. Auflage 2020

© 2020 Beltz Juventa

in der Verlagsgruppe Beltz · Weinheim Basel

Werderstraße 10, 69469 Weinheim

Alle Rechte vorbehalten

Herstellung: Hannelore Molitor

Satz: text plus form, Dresden

Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza

Printed in Germany

Weitere Informationen zu unseren Autoren und Titeln finden Sie unter: [www.beltz.de](http://www.beltz.de)

# Inhaltsverzeichnis

*Anna-Katharina Praetorius/Juliane Grünkorn/Eckhard Klieme*  
Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen  
und quantitative Modellierungen. Einleitung in das Beiheft ..... 9

## **Themenblock I: Dimensionen der Unterrichtsqualität – Theoretische und empirische Grundlagen (englischsprachig)**

*Anna-Katharina Praetorius/Eckhard Klieme/Thilo Kleickmann/Esther Brunner/  
Anke Lindmeier/Sandy Taut/Charalambos Charalambous*  
Towards Developing a Theory of Generic Teaching Quality: Origin,  
Current Status, and Necessary Next Steps Regarding the Three Basic  
Dimensions Model ..... 15

*Thilo Kleickmann/Mirjam Steffensky/Anna-Katharina Praetorius*  
Quality of Teaching in Science Education: More Than Three  
Basic Dimensions? ..... 37

*Courtney A. Bell*  
Commentary Regarding the Section “Dimensions of Teaching Quality –  
Theoretical and Empirical Foundations” – Using Warrants and Alternative  
Explanations to Clarify Next Steps for the TBD Model ..... 56

## **Themenblock II: Angebots-Nutzungs-Modelle als Rahmung (deutschsprachig)**

*Svenja Vieluf/Anna-Katharina Praetorius/Katrin Rakoczy/Marc Kleinknecht/  
Marcus Pietsch*  
Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts:  
ein kritischer Vergleich verschiedener Modellvarianten ..... 63

*Sibylle Meissner/Samuel Merk/Benjamin Fauth/Marc Kleinknecht/  
Thorsten Bohl*  
Differenzielle Effekte der Unterrichtsqualität auf die aktive Lernzeit ..... 81

*Tina Seidel*

Kommentar zum Themenblock „Angebots-Nutzungs-Modelle als Rahmung“ – Quo vadis deutsche Unterrichtsforschung? Modellierung von Angebot und Nutzung im Unterricht .....	95
---	----

### **Themenblock III: Oberflächen- und Tiefenstruktur des Unterrichts (deutschsprachig)**

<i>Jasmin Decristan/Miriam Hess/Doris Holzberger/Anna-Katharina Praetorius</i> Oberflächen- und Tiefenmerkmale – eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung .....	102
--	-----

<i>Miriam Hess/Frank Lipowsky</i> Zur (Un-)Abhängigkeit von Oberflächen- und Tiefenmerkmalen im Grundschulunterricht – Fragen von Lehrpersonen im öffentlichen Unterricht und in Schülerarbeitsphasen im Vergleich .....	117
---	-----

<i>Christine Pauli</i> Kommentar zum Themenblock „Oberflächen- und Tiefenstruktur des Unterrichts“: Nutzen und Grenzen eines prominenten Begriffspaares für die Unterrichtsforschung – und das Unterrichten .....	132
--	-----

### **Themenblock IV: Zur Bedeutung unterschiedlicher Perspektiven bei der Erfassung von Unterrichtsqualität (englischsprachig)**

<i>Benjamin Fauth/Richard Göllner/Gerlinde Lenske/Anna-Katharina Praetorius/ Wolfgang Wagner</i> Who Sees What? Conceptual Considerations on the Measurement of Teaching Quality from Different Perspectives .....	138
--	-----

<i>Richard Göllner/Benjamin Fauth/Gerlinde Lenske/Anna-Katharina Praetorius/ Wolfgang Wagner</i> Do Student Ratings of Classroom Management Tell us More About Teachers or About Classroom Composition? .....	156
---	-----

<i>Marten Clausen</i> Commentary Regarding the Section “The Role of Different Perspectives on the Measurement of Teaching Quality” .....	173
--	-----

## **Themenblock V: Modellierung der Wirkungen von Unterrichtsqualität (englischsprachig)**

<i>Alexander Naumann/Susanne Kuger/Carmen Köhler/Jan Hochweber</i> Conceptual and Methodological Challenges in Detecting the Effectiveness of Learning and Teaching .....	179
<i>Carmen Köhler/Susanne Kuger/Alexander Naumann/Johannes Hartig</i> Multilevel Models for Evaluating the Effectiveness of Teaching: Conceptual and Methodological Considerations .....	197
<i>Oliver Lüdtke/Alexander Robitzsch</i> Commentary Regarding the Section “Modelling the Effectiveness of Teaching Quality” – Methodological Challenges in Assessing the Causal Effects of Teaching .....	210
 <b>Kommentare</b>	
<i>Ewald Terhart</i> Unterrichtsqualität zwischen Theorie und Empirie – Ein Kommentar zur Theoriediskussion in der empirisch-quantitativen Unterrichtsforschung .....	223
<i>Kurt Reusser</i> Unterrichtsqualität zwischen empirisch-analytischer Forschung und pädagogisch-didaktischer Theorie – Ein Kommentar .....	236
<i>Anke Lindmeier/Aiso Heinze</i> Die fachdidaktische Perspektive in der Unterrichtsqualitätsforschung: (bisher) ignoriert, implizit enthalten oder nicht relevant? .....	255

Miriam Hess/Frank Lipowsky

# Zur (Un-)Abhängigkeit von Oberflächen- und Tiefenmerkmalen im Grundschulunterricht

*Fragen von Lehrpersonen im öffentlichen Unterricht und in Schülerarbeitsphasen im Vergleich*

**Zusammenfassung:** In der Literatur zur Unterrichtsqualität wird davon ausgegangen, dass die Qualität von Unterricht vor allem von dessen Tiefenstruktur abhängt, während die Oberflächenstruktur hierfür einen Rahmen bildet. Um diesen Zusammenhang empirisch zu untersuchen, wird anhand von Videodaten aus dem Leseunterricht des ersten Schuljahres ( $N = 47$  Videos) analysiert, ob Oberflächenmerkmale (hier: öffentlicher Unterricht vs. Schülerarbeitsphasen) systematisch mit dem kognitiven Niveau der Fragen von Lehrpersonen als ausgewähltem Aspekt der Tiefenstruktur zusammenhängen. Chi-Quadrat-Tests zeigen, dass sich die im Unterricht gestellten Fragen teilweise deutlich zwischen öffentlichem Unterricht und Schülerarbeitsphasen unterscheiden, woraus Implikationen für die (videobasierte) Unterrichtsbeobachtung abgeleitet werden.

**Schlagnworte:** Oberflächenstruktur des Unterrichts, Tiefenstruktur des Unterrichts, Unterrichtsqualität, Fragen von Lehrpersonen, Grundschule

## 1. Einleitung

Im vorangegangenen Beitrag arbeiteten Decristan, Hess, Holzberger und Praetorius heraus, dass in der aktuellen Forschung zur Beurteilung der Unterrichtsqualität vor allem Tiefenmerkmale<sup>1</sup> des Unterrichts berücksichtigt werden. Gleichzeitig kann aber angenommen werden, dass Oberflächenmerkmale des Unterrichts einen gewissen Rahmen bereitstellen, sodass die Qualität von Unterricht bzw. deren Beurteilung im Rahmen von (videobasierten) Unterrichtsbeobachtungen auch von Aspekten auf Ebene der Oberfläche abhängig sein dürfte. So könnten beispielsweise verschiedene Sozialformen im Unterricht ein unterschiedlich hohes Potenzial zur Ermöglichung einer hohen Qualität der Interaktionen zwischen Lehrpersonen und Lernenden bieten (vgl. Abschnitt 2.2).

Anhand von Beobachtungsdaten aus dem Leseunterricht des ersten Schuljahres, welche aus dem PERLE-Projekt stammen, wird daher der Frage nachgegangen, ob die Oberflächenmerkmale des Unterrichts (hier operationalisiert anhand der Unterscheidung zwischen öffentlichem Unterricht und Schülerarbeitsphasen) systematisch mit dem kognitiven Niveau der Fragen von Lehrpersonen zusammenhängen. Fragen kön-

1 In Anlehnung an den Beitrag von Decristan, Hess, Holzberger und Praetorius (in diesem Heft) werden im Beitrag im Folgenden die Begriffe Oberflächen- und Tiefenstruktur weitgehend durch „Oberflächen-/Tiefenmerkmale“ oder „Oberflächen-/Tiefenebene“ ersetzt.

nen als ein Aspekt kognitiver Aktivierung zu den Tiefenmerkmalen des Unterrichts gezählt werden, da sie kognitive Prozesse anregen können (vgl. Abschnitt 2.3).

## 2. Theoretischer Hintergrund und Forschungsstand

### 2.1 Oberflächen- und Tiefenmerkmale

Bezugnehmend auf den Beitrag von Decristan et al. (in diesem Heft) werden Oberflächenstrukturen als methodisch-organisatorische Gestaltungsmerkmale des Unterrichts verstanden, die nicht auf direktem Weg kognitive Prozesse anregen, aber einen mehr oder weniger geeigneten Rahmen schaffen können, um die Voraussetzungen für die Anregung von Denkprozessen bereitzustellen. Tiefenmerkmale hingegen dürften direkt die mentalen Verarbeitungsprozesse der Lernenden adressieren.<sup>2</sup>

Decristan et al. (in diesem Heft) formulieren als ein Desiderat der aktuellen Forschung die Frage, inwiefern sich Oberflächenmerkmale des Unterrichts auf dessen Tiefenstruktur auswirken bzw. damit in Zusammenhang stehen. Dieser Frage soll daher im vorliegenden Beitrag nachgegangen werden. Als klassische Oberflächenmerkmale gelten die Sozialformen des Unterrichts. Daher werden hier zwei Sozialformen, der öffentliche Unterricht und die Schülerarbeitsphasen (Einzel-, Partner- und Gruppenarbeit), betrachtet. Im Folgenden wird zunächst darauf eingegangen, was diese Phasen grundlegend kennzeichnet.

### 2.2 Unterrichtsqualität in Schülerarbeitsphasen und im öffentlichen Unterricht

Unterricht kann in verschiedenen Sozialformen stattfinden, wobei grundlegend zwischen eher schülerzentrierten Formen wie Einzelarbeit, Partnerarbeit und Gruppenarbeit (Schülerarbeitsphasen) und lehrerzentrierten Formen (öffentlicher Unterricht) unterschieden werden kann. Innerhalb des öffentlichen Unterrichts können Methoden wie das Unterrichtsgespräch oder der Vortrag von Lehrpersonen zum Einsatz kommen. Durch die Sozialformen „wird der Unterricht methodisch-organisatorisch strukturiert“ (Lotz, 2013, S. 123). Da die Wahl der Sozialform allein aber noch keinen direkten Einfluss auf die Anregung kognitiver Prozesse haben dürfte, besteht in der aktuellen Literatur Konsens darüber, sie zu den Oberflächenmerkmalen des Unterrichts zu zählen (z. B. Denn, Hess & Lipowsky, 2017; Hugener, 2008).

Es kann aber angenommen werden, dass gewisse Oberflächenmerkmale eine höhere kognitive Aktivierung der Lernenden begünstigen können, da sich die Funktionen von öffentlichen Phasen und Schülerarbeitsphasen unterscheiden. Während im öffentlichen

---

2 Für eine umfassende Auseinandersetzung mit den verschiedenen Begriffsverständnissen wird an dieser Stelle auf den vorangegangenen Beitrag von Decristan et al. (in diesem Heft) verwiesen.

Unterricht beispielsweise häufig neue Inhalte eingeführt werden, dienen Schülerarbeitsphasen oftmals dem Üben, Festigen und Automatisieren (z. B. Krammer, 2009). Schülerarbeitsphasen stellen außerdem besonders hohe Anforderungen an die Lehrperson, welche die Arbeitsprozesse von vielen Schülern parallel im Blick haben müsste, um alle Lernenden mit passgenauen und auf dem jeweiligen Niveau kognitiv herausfordernden Instruktionen, Hilfen und Feedback versorgen zu können. Insbesondere wenn die Lernenden an unterschiedlichen Aufgaben arbeiten, lässt sich annehmen, dass im Sinne der Cognitive Load Theorie (Sweller, 1994) „die Organisation und die Steuerung einer so hohen Anzahl unterschiedlicher Prozesse bereits so viele kognitive Kapazitäten der Lehrperson beanspruchen, dass für eine anspruchsvollere inhaltliche Unterstützung keine freien Kapazitäten mehr zur Verfügung stehen“ (Lipowsky & Lotz, 2015, S. 179). In Schülerarbeitsphasen kommt der Lehrperson auch die Funktion eines Tutors zu. Effektive Tutoren können nach Lepper und Woolverton (2002) durch sieben Merkmale charakterisiert werden, die die Autoren im sogenannten Inspire-Modell zusammenfassen (intelligent, nurturant, socratic, progressive, indirect, reflective, encouraging). Gute Lernbegleitung ist demnach u. a. gekennzeichnet durch fundiertes, intelligentes Wissen über das Fach und dessen Didaktik, die Schaffung eines ‚nahrhaften‘, guten Lernklimas sowie ein sokratisches, also durch Fragen und Impulse aktivierendes Lehrverhalten. Damit verdeutlicht das Modell, dass die Ansprüche an das Verhalten von Lehrpersonen in Schülerarbeitsphasen sehr hoch sind.

Dass es möglich ist, Lehrpersonen zu trainieren, den hohen Ansprüchen, die Schülerarbeitsphasen an das Verhalten der Lehrpersonen stellen, besser gerecht zu werden, zeigten Galton, Hargreaves und Pell (2008). Sie beobachteten in Gruppenarbeitsphasen nachhaltigere Interaktionen auf einem höheren kognitiven Niveau als in Klassengesprächen, nachdem die teilnehmenden Lehrpersonen vorab an einem Training teilgenommen hatten, in dem es darum ging, die Fähigkeiten der Lernenden für Gruppenarbeiten zu verbessern.

Bislang gibt es nur wenige Studien, die Unterschiede in der Unterrichtsqualität zwischen öffentlichem Unterricht und Schülerarbeitsphasen systematisch untersuchen. Oft wird in Beobachtungsstudien nur eine der beiden Phasen fokussiert, sodass Vergleiche innerhalb einer Unterrichtsstunde nicht möglich sind. Auch wird die Unterrichtsqualität häufig global über den Verlauf des gesamten Unterrichts beurteilt, sodass keine direkten Vergleiche zwischen den Phasen gezogen werden können. Ergebnisse verschiedener Studien miteinander zu vergleichen, die entweder Schülerarbeitsphasen oder öffentlichen Unterricht in den Blick genommen haben, ist oftmals schwierig, da in diesem Fall meist noch sehr viele weitere Merkmale variieren (z. B. Klassenstufe, Fach, Unterrichtsinhalt etc.).

Hugener, Rakoczy, Pauli und Reusser (2006) schlagen daher vor, durch die gleichzeitige Erfassung von Oberflächen- und Tiefenmerkmalen innerhalb einer Studie zu überprüfen, „ob gewisse Lernaktivitäten mit ausgewählten Qualitätseinschätzungen zusammenhängen“, wobei zwei grundlegende Möglichkeiten bestehen: „(1) Lektionen mit hohen Qualitätseinschätzungen können daraufhin analysiert werden, ob sie sich in Bezug auf die tatsächlich inszenierten Lehr-Lernaktivitäten unterscheiden, oder (2) Lek-

tionen mit ähnlichen beobachteten Lehr-Lernaktivitäten können daraufhin untersucht werden, ob sie sich auch im Qualitätsrating gleichen“ (S. 49).

Hugener et al. (2009) gingen im Rahmen der schweizerisch-deutschen Pythagoras-Videostudie der Frage nach, ob bestimmte Inszenierungsmuster des Unterrichts (hier der Oberflächenstruktur zugeordnet) mit der Einschätzung der kognitiven Aktivierung (Tiefenstruktur) zusammenhängen. Es zeigte sich, dass problemlösend-entdeckende Unterrichtsmuster im Vergleich zu darstellenden oder fragend-entwickelnden Vorgehensweisen im hoch inferenten Rating als kognitiv aktivierender eingeschätzt wurden. Weitere Analysen ergaben aber, dass zusätzliche Aspekte kognitiver Aktivierung (Anteil anspruchsvollen Übens, durchschnittliche Länge von Schülerbeiträgen im Klassengespräch und inhaltlich-strukturelle Klarheit von Theoriephasen) von den Inszenierungsmustern unabhängig sind (Pauli, Drollinger-Vetter, Hugener & Lipowsky, 2008).

Ein weiterer Versuch, Zusammenhänge zwischen Oberflächen- und Tiefenmerkmalen zu analysieren, wurde von Lotz (2015) unternommen. In dieser Studie wurde der Leseunterricht im ersten Schuljahr anhand niedrig inferenter Kodierungen in Phasen unterteilt und zusätzlich hoch inferent in seiner Qualität beurteilt. Anschließend wurden Korrelationen zwischen der Dauer und den prozentualen Anteilen verschiedener Sozialformen und der Ausprägung der Qualitätseinschätzungen des Unterrichts berechnet. Dabei zeigen sich einzelne Zusammenhänge: Beispielsweise korrelieren die klassenbezogenen Werte für die hoch inferent erfasste Skala „Anregung von Denkprozessen“ signifikant positiv mit der absoluten Dauer sowie dem prozentualen Anteil von Schülerarbeitsphasen im Unterricht. Auch die „Förderung einer eigenständigen Auseinandersetzung mit dem Lerngegenstand“ wird erwartungsgemäß signifikant besser eingeschätzt, wenn der Anteil an öffentlichem Unterricht geringer ist. Diese Ergebnisse können als erster Hinweis darauf gewertet werden, dass sich die Qualität des Unterrichts in Abhängigkeit von seiner Oberflächenstruktur unterscheiden kann. Da das hoch inferente Rating nicht für die verschiedenen Phasen getrennt durchgeführt wurde, können die berechneten Korrelationen zwischen der Ausprägung der Unterrichtsqualität und der Dauer und der Anzahl der Unterrichtsphasen lediglich als Hinweis, aber nicht als Beleg dafür betrachtet werden, dass Tiefenmerkmale mit Oberflächenmerkmalen systematisch kovariieren. Dazu sind phasenspezifische Analysen notwendig.

Daher werden im vorliegenden Beitrag Fragen von Lehrpersonen innerhalb des öffentlichen Unterrichts mit Fragen von Lehrpersonen während Schülerarbeitsphasen verglichen. Dieses Vorgehen ermöglicht den direkten Vergleich zwischen den Phasen.

### 2.3 Fragen von Lehrpersonen im Unterricht

Fragen von Lehrpersonen kommen sowohl im Unterrichtsgespräch als auch in der individuellen Lehrer-Schüler-Interaktion vor und werden vor allem anhand ihrer Funktion, eine Antwort zu verlangen, „der im allgemeinen ein bestimmter Denkproze(ss) vorausgeht“ (Spanhel, 1980, S. 89), definiert. Da durch Fragen kognitive Prozesse angeregt

werden können (z. B. Lipowsky, 2015), lassen sie sich den Tiefenmerkmalen des Unterrichts zuordnen (vgl. Decristan et al., in diesem Heft).

Geht es um die Qualität von Fragen in ihrer Funktion zur Förderung kognitiver Prozesse, erscheint die Unterscheidung ihres kognitiven Niveaus zentral (z. B. Levin, 2005). Die Grundlage für die meisten Klassifikationssysteme zum kognitiven Niveau von Fragen ist die Lernzieltaxonomie von Bloom, Engelhart, Furst, Hill und Krathwohl (1976), nach der Wissens-, Verständnis-, Anwendungs-, Analyse-, Synthese- und Bewertungsfragen unterschieden werden. Vereinfacht werden darauf aufbauend oftmals auch lediglich die beiden Kategorien Wissens- vs. Denkfragen bzw. Fragen mit höherem vs. niedrigerem kognitiven Niveau (higher vs. lower order questions) differenziert (z. B. Gayle et al., 2006; Lotz, 2015).

Wissensfragen zielen darauf ab, dass sich die Lernenden an bereits Bekanntes oder Erarbeitetes erinnern und dieses Wissen wiedergeben. Bei Denkfragen rückt hingegen eine neuartige Situation oder ein unbekanntes Problem in den Mittelpunkt. Vorerfahrungen müssen genutzt werden, um sie auf neue Situationen oder Probleme zu übertragen. Daneben können außerdem Reflexionsfragen, organisatorische und ablaufgerichtete Fragen im Unterricht vorkommen. Reflexionsfragen zielen darauf ab, die Lernenden zum Nachdenken über (ihre eigenen) Lernprozesse anzuregen. Organisatorische und ablaufgerichtete Fragen dienen hingegen eher der Unterrichtsorganisation (Lotz, 2015).<sup>3</sup>

### 3. Fragestellung

Der Beitrag geht der übergeordneten Frage nach, inwiefern ausgewählte Oberflächen- und Tiefenmerkmale des Unterrichts systematisch miteinander zusammenhängen. Dabei soll folgende Fragestellung fokussiert werden: *Unterscheidet sich das kognitive Niveau von Fragen im öffentlichen Unterricht und in Schülerarbeitsphasen?*

Sollten sich hier Unterschiede zeigen, wäre dies ein Hinweis dafür, dass verschiedene Sozialformen ein unterschiedlich hohes Potenzial zur Anregung kognitiver Prozesse bieten. Aus methodischer Sicht ist die Fragestellung des Beitrags für die (video-basierte) Unterrichtsforschung von Interesse, da sie Aufschluss darüber geben kann, ob bei der Beurteilung bestimmter Merkmale von Unterrichtsqualität die Rahmenbedingungen des Unterrichts (noch stringenter oder umfassender) berücksichtigt werden müssten, um systematische Verzerrungen oder Fehlinterpretationen der Ergebnisse zu vermeiden.

---

3 Als organisatorische Fragen wurden in der vorliegenden Studie Fragen zur Verwendung von Hilfsmitteln, zur Vorbereitung der Aufgabenbearbeitung, zum Aufrufen von Schülern zum Vorlesen oder Helfen sowie Aufforderungen zur Bearbeitung eines Auftrags sowie Ermahnungen in Form von Fragen kodiert. Ablaufgerichtete Fragen beziehen sich auf den Arbeitsstand, das Verständnis des Arbeitsauftrags, die Wahl der Aufgabe oder Textsorte sowie die Art der Aufgabenbearbeitung.

#### 4. Datengrundlage: Die Videostudie im Fach Deutsch des PERLE-Projekts

Die den Analysen zugrundeliegenden Beobachtungsdaten stammen aus der Längsschnittstudie PERLE (Lipowsky, Faust & Kastens, 2013), in deren Rahmen die Persönlichkeits- und Lernentwicklung von Grundschulkindern sowie deren Determinanten vom ersten bis zum vierten Schuljahr untersucht wurden. Im März des ersten Schuljahres wurde im Fach Deutsch eine 90-minütige Unterrichtseinheit videografiert (Lotz & Corvacho del Toro, 2013). Zur Ermöglichung einer grundlegenden Vergleichbarkeit der videografierten Unterrichtseinheiten erhielten die Lehrpersonen inhaltliche Vorgaben zur Gestaltung der Stunde, in der u. a. eine Leseübung durchgeführt werden sollte. Über die Reihenfolge und alle weiteren didaktisch-methodischen Aspekte – wie beispielsweise die Sozialformen – konnten die Lehrpersonen selbst entscheiden.

Die Lehrpersonen waren bis auf eine Ausnahme weiblich und verfügten im Schnitt über 15 Jahre Berufserfahrung. Die Lernenden waren zum Zeitpunkt der Datenerhebung im Mittel sieben Jahre und einen Monat alt (Mädchenanteil: 52.4%).

Für die Analyse der Fragestellungen werden nur Kodierungen der Leseübungsphase genutzt. Diese wurde vorab über eine niedrig inferente Kodierung identifiziert (vgl. Lotz, 2015), wobei nur Videos mit Leseübungen von mindestens fünf Minuten einbezogen wurden ( $N = 47$ ;  $Min = 6.17$  min;  $Max = 55.50$  min;  $M = 26.76$  min;  $SD = 12.47$  min). Aufgrund der Varianz in der Unterrichtszeit wurden alle Analysen mit relativen Häufigkeiten oder prozentualen Anteilen durchgeführt.

#### 5. Methodisches Vorgehen

##### 5.1 Kodierung der Sozialformen des Unterrichts

Für die gesamte Phase der Leseübung wurden die Sozialformen niedrig inferent im Time-Sampling-Verfahren (10-Sekunden-Intervalle) kodiert (vgl. Lotz, 2013). Grundlegend werden durch das Kategoriensystem öffentliche Unterrichtsphasen von Schülerarbeitsphasen unterschieden. Zu den Schülerarbeitsphasen zählen die Kategorien „Einzelarbeit“, „Partnerarbeit“ und „Gruppenarbeit“. Der öffentliche Unterricht wurde mit den beiden Kategorien „Öffentlicher Unterricht im Sitzkreis“ und „Öffentlicher Unterricht ohne Sitzkreis“ erfasst. Für die folgenden Analysen werden nur die zusammengefassten Kategorien „Schülerarbeitsphasen“ („SAP“) und „Öffentlicher Unterricht“ („OEU“) differenziert. Zusätzlich wurde die „Restkategorie“ vergeben, wenn es während der Leseübung zu kurzen Unterbrechungen kam, in denen keine inhaltliche Arbeit stattfand.

Die Kodierungen wurden von vier trainierten Beobachterinnen durchgeführt (prozentuale Übereinstimmung  $PÜ \geq 97.52\%$ ; Cohens Kappa  $\kappa \geq .97$ ; vgl. Lotz, 2013; 2015). In den 47 hier analysierten Videos liegt der Anteil des öffentlichen Unterrichts durchschnittlich bei  $M = 39.5\%$  ( $SD = 25.5\%$ ), der Anteil an Schülerarbeitsphasen bei

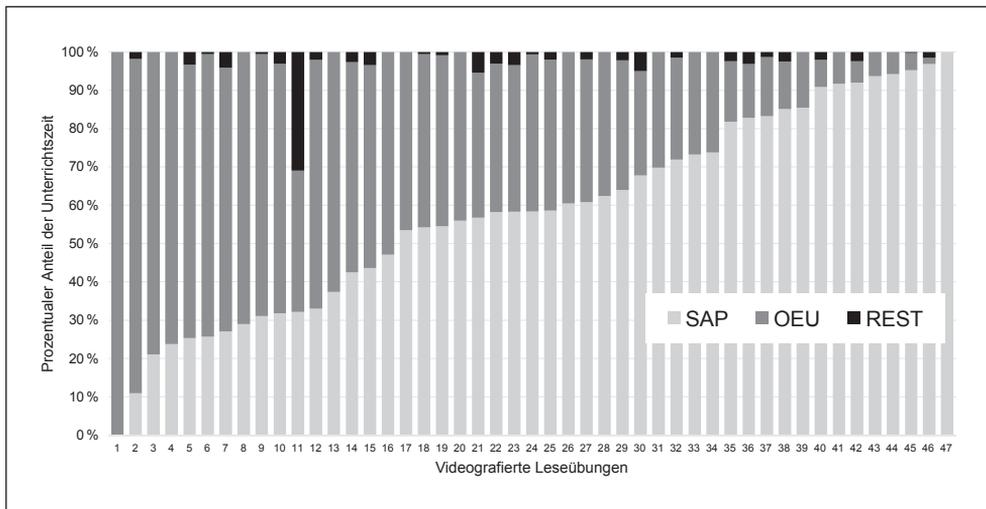


Abb. 1: Prozentuale Anteile von Schülerarbeitsphasen (SAP) und öffentlichem Unterricht (OEU) in den 47 videografierten Leseübungen

$M = 58.4\%$  ( $SD = 25.8\%$ ) und der Anteil der Restkategorie bei nur  $2.0\%$  ( $SD = 4.6$ ). Unter den 47 Videos gibt es sowohl ein Video, in dem die Leseübung zu  $100\%$  in Schülerarbeitsphasen durchgeführt wurde als auch eine Leseübung, die ausschließlich im öffentlichen Unterricht stattfand.

Da sich – wie Abbildung 1 auch grafisch verdeutlicht – auf der Oberflächenebene des Unterrichts deutliche Unterschiede zwischen den videografierten Lerngruppen zeigen, bieten sich die Daten an, um Zusammenhänge zwischen der Oberflächen- und Tiefenstruktur des Unterrichts zu untersuchen.

## 5.2 Kodierung des kognitiven Niveaus der Fragen von Lehrpersonen

Als ein exemplarischer Aspekt auf der Tiefenebene des Unterrichts wurde jede während der Leseübung vorkommende Frage der Lehrperson mit einem niedrig inferenten Event-Sampling-Verfahren identifiziert. Als Fragen wurden dabei alle Äußerungen der Lehrperson definiert, die eine Schülerantwort oder -äußerung intendieren und inhaltliche Relevanz für die Leseübung besitzen. Damit beinhaltet die Kodierung auch Impulse. Für jede Frage wurde überdies festgehalten, in welcher Sozialform und auf welchem kognitiven Niveau sie gestellt wurde (vgl. Lotz, 2015). Das vollständige Kategoriensystem mit allen Kodierregeln findet sich bei Lotz (2015).<sup>4</sup> Dieses System wurde

<sup>4</sup> Das Kategoriensystem mit den Kodierregeln ist als Anhang auch online verfügbar unter <https://www.springer.com/de/book/9783658104351> (zuletzt geprüft: 17.01.2020).

basierend auf vorhandenen Klassifikationen verschiedener Fragearten deduktiv entworfen und induktiv anhand der eigenen Videos weiterentwickelt.

Bei der „Art der Frage“ (kognitives Niveau) wurden insgesamt 26 einzelne Fragearten unterschieden, die zu den fünf übergeordneten Kategorien Wissensfragen, Denkfragen, Reflexionsfragen, organisatorische Fragen und ablaufgerichtete Fragen zusammengefasst wurden. Die Auswertung wurde auch hier von geschulten Kodiererinnen vorgenommen ( $PÜ \geq 97.40\%$ ; vgl. Lotz, 2015).

In den 47 videografierten Leseübungen wurden insgesamt 2 087 Fragen von Lehrpersonen an die Lernenden identifiziert. Im Schnitt entspricht dies 44.40 Fragen pro Video, wobei sich hier – wie bereits bei den Sozialformen – eine deutliche Varianz zeigt ( $Min = 7$ ;  $Max = 118$ ;  $SD = 30.61$ ). Relativiert an der Dauer der jeweiligen Leseübungen wurden durchschnittlich 1.70 Fragen pro Minute gestellt ( $Min = 0.38$ ;  $Max = 4.86$ ;  $SD = 1.00$ ).

### 5.3 Analysen der Zusammenhänge zwischen Oberflächen- und Tiefenmerkmalen

Zur Analyse der Zusammenhänge zwischen den Oberflächen- und Tiefenmerkmalen des Unterrichts werden zunächst deskriptive Analysen vorgenommen. Da die Fragen der Lehrpersonen im Event-Sampling-Verfahren einzeln kodiert wurden, kann für jede einzelne Frage analysiert werden, ob sie in einer Schülerarbeitsphase oder im öffentlichen Unterricht gestellt wurde. Auf dieser Basis können die Verteilungen und Häufigkeiten in Schülerarbeitsphasen und im öffentlichen Unterricht verglichen werden. Ob sich die Verteilungen systematisch unterscheiden, wird mit Chi-Quadrat-Tests überprüft.

## 6. Ergebnisse

Von den insgesamt 2 087 gestellten Fragen wurden 1 151 in Schülerarbeitsphasen gestellt, 931 im öffentlichen Unterricht und nur 5 Fragen während der Phase, die der Restkategorie zugeordnet wurde. Diese 5 Fragen werden im Folgenden nicht mehr berücksichtigt.

Im öffentlichen Unterricht wurden von den Lehrpersonen im Schnitt 2.04 Fragen pro Minute gestellt, während in Schülerarbeitsphasen die Anzahl der Fragen von Lehrpersonen mit 1.49 Fragen pro Minute etwas geringer ausfällt.<sup>5</sup> Insbesondere im öffentlichen Unterricht fällt aber eine starke Varianz zwischen den Klassen in der relativen

<sup>5</sup> Dass auch in den Schülerarbeitsphasen häufig Fragen gestellt werden, lässt sich damit erklären, dass in den videografierten Leseübungen die Erstklasslehrpersonen auch in Schülerarbeitsphasen zu einem Großteil der Zeit mit einzelnen Lernenden interagieren (vgl. auch Lotz, 2015) und dabei auch Fragen stellen (z. B. „Wie heißt dieses Wort?“; „Kommst du zu-

Durchschnittliche Anzahl an Fragen pro Minute	N	Min	Max	M	SD
Gesamte Leseübung (OEU+SAP)	47	0.38	4.86	1.70	0.98
Öffentlicher Unterricht (OEU)	46	0.00	10.50	2.04	1.75
Schülerarbeitsphasen (SAP)	46	0.13	4.69	1.49	1.10

Anmerkung: N = 46 kommt zustande, da in jeweils einer Klasse die Leseübung ausschließlich in Schülerarbeitsphasen bzw. im öffentlichen Unterricht stattfand.

Tab. 1: Deskriptive Statistik der durchschnittlichen Anzahl von Fragen im öffentlichen Unterricht und in Schülerarbeitsphasen

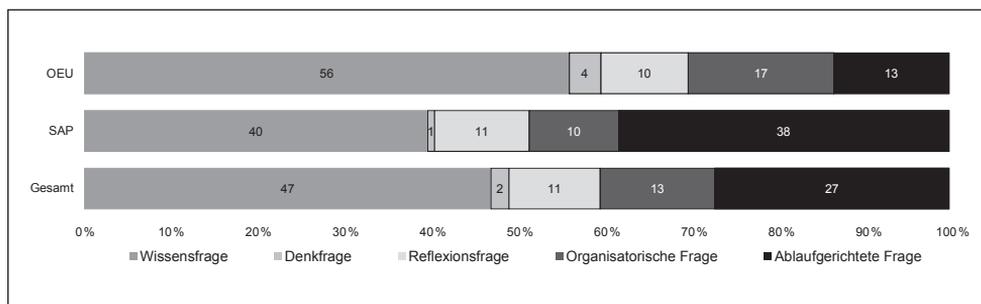


Abb. 2: Prozentuale Anteile der Arten von Fragen in Schülerarbeitsphasen (SAP) und im öffentlichen Unterricht (OEU)

Häufigkeit von Fragen von Lehrpersonen auf: Es gibt sowohl Klassen, in denen während des öffentlichen Unterrichts keine einzige Frage gestellt wird, als auch eine Klasse, in der durchschnittlich zwischen zehn und elf Fragen von Lehrpersonen pro Minute im öffentlichen Unterricht vorkommen (vgl. Tab. 1).

Abbildung 2 stellt die Verteilung der verschiedenen Arten von Fragen jeweils einzeln für den öffentlichen Unterricht, für die Schülerarbeitsphasen und für die gesamte Leseübung im Vergleich dar. Betrachtet man zunächst die Verteilung in der Leseübung insgesamt, so machen Wissensfragen mit 47.0% den größten Anteil aus, wohingegen sich nur 2.1% aller Fragen der übergeordneten Kategorie der Denkfragen zuordnen lassen. Auch ablaufgerichtete Fragen kommen mit 27.1% relativ häufig vor, wohingegen Reflexionsfragen (10.6%) und organisatorische Fragen (13.3%) seltener gestellt werden.

Der Vergleich der Verteilungen zeigt, dass die relative Häufigkeit von Reflexions- und organisatorischen Fragen in Schülerarbeitsphasen und im öffentlichen Unterricht

*recht?"; „Bist du schon fertig?“). Im Gegensatz zu den Fragen in öffentlichen Unterrichtsphasen richten sich die Fragen in Schülerarbeitsphasen meist an einzelne Schüler oder kleinere Schülergruppen.*

relativ ähnlich ausfällt. Eine Verschiebung deutet sich aber zwischen Wissensfragen und ablaufgerichteten Fragen an. Während im öffentlichen Unterricht die Wissensfragen mit 56.1% mehr als die Hälfte aller Fragen ausmachen und lediglich 13.3% ablaufgerichtete Fragen gestellt werden, ist der Anteil von Wissensfragen in den Schülerarbeitsphasen geringer (39.7%), wohingegen deutlich mehr ablaufgerichtete Fragen vorkommen (38.2%). Der ohnehin sehr geringe Anteil von Denkfragen (insgesamt 2.1%) ist in Schülerarbeitsphasen mit nur 0.8% sogar noch einmal geringer als im öffentlichen Unterricht mit zumindest 3.7%. Der geringe Anteil von Denkfragen drückt sich auch in ihrer geringen absoluten Zahl aus: Von den insgesamt über alle Videos hinweg lediglich 43 gestellten Denkfragen werden 34 (79.1%) im öffentlichen Unterricht gestellt und nur 9 (20.9%) in Schülerarbeitsphasen. Dabei muss berücksichtigt werden, dass der Anteil an Schülerarbeitsphasen mit 58.4% der Unterrichtszeit sogar höher ist als der Anteil an öffentlichem Unterricht (39.5%). Bei den ablaufgerichteten Fragen kehrt sich das Verhältnis um: 78.0% aller ablaufgerichteten Fragen werden in Schülerarbeitsphasen gestellt und nur 22.0% im öffentlichen Unterricht.

Der  $\chi^2$ -Test ergibt einen Wert von 184.60 und liegt damit deutlich über dem kritischen Wert von  $\chi^2 = 9.49$  (Kreuztabelle:  $5 \times 2$ ;  $df = 4$ ;  $p \leq .05$ ; Effektstärke  $d = 0.62$ ). Diese Abweichung der tatsächlichen Verteilung von der bei Gültigkeit der Nullhypothese erwartbaren Verteilung ist vorwiegend auf die Kategorien „Wissens-“, „Denk-“ und „ablaufgerichtete Fragen“ zurückzuführen.

Eine detailliertere Betrachtung zeigt, dass es sich bei 40.5% aller im öffentlichen Unterricht gestellten Fragen um Fragen zum „Verständnis des Inhalts/Textes“ handelt, die zur übergeordneten Kategorie der Wissensfragen zählen (z. B. „Wie fängt die Geschichte an und wie hört sie auf?“ oder „Was wollte denn die Lucy [Hauptfigur des Buchs] machen?“). In Schülerarbeitsphasen hingegen machen solche Fragen nur 18.9% der dort gestellten Fragen aus. In der individuellen Auseinandersetzung mit den Lernenden werden hingegen öfter Fragen gestellt, die den direkten Leseprozess und die Lesetechnik betreffen („Verständnis der Wörter/Laute“, z. B. „Was ist denn das für ein Buchstabe?“). Diese machen in Schülerarbeitsphasen 17.5% der gestellten Fragen aus, im öffentlichen Unterricht nur 8.9%. Die Dominanz ablaufgerichteter Fragen in Schülerarbeitsphasen lässt sich vor allem auf Fragen zum „Arbeitsstand“ zurückführen (z. B. „Wie viel habt ihr geschafft?“).

## 7. Zusammenfassung und Diskussion

Im Beitrag konnte gezeigt werden, dass sich die Arten von Fragen teilweise deutlich zwischen öffentlichem Unterricht und Schülerarbeitsphasen unterscheiden. Auffällig ist vor allem, dass Fragen im öffentlichen Unterricht stärker der Auseinandersetzung mit den Textinhalten zu dienen scheinen, wohingegen in Schülerarbeitsphasen mehr Fragen dazu genutzt werden, die Abläufe zu organisieren und aufrechtzuerhalten. Denkfragen kommen generell kaum vor, in Schülerarbeitsphasen allerdings noch seltener als im öffentlichen Unterricht. Dies deutet darauf hin, dass insbesondere in Schülerarbeitsphasen

Fragen weniger dazu genutzt werden, die Lernenden kognitiv herauszufordern, Wissen aufzubauen oder zu sichern.

Viele der gezeigten Unterschiede lassen sich gut anhand unterschiedlicher Funktionen und Erfordernisse in den beiden Phasen erklären (vgl. z. B. Krammer, 2009; Abschnitt 2.2). In Schülerarbeitsphasen dürfte es für viele Lehrpersonen wichtig sein, sich – beispielsweise anhand ablaufgerichteter Fragen – einen Überblick darüber zu verschaffen, wie die Lernenden zurechtkommen (z. B. „Hast du den Arbeitsauftrag verstanden?“). Schülerarbeitsphasen bieten außerdem die Möglichkeit, einzelne und insbesondere schwächere Lernende gezielt bei hierarchieniedrigeren Leseprozessen zu unterstützen (z. B. durch Fragen der Kategorie „Verständnis der Wörter/Laute“), wohingegen im Klassenunterricht der Fokus eher auf den Textinhalten liegt, was das häufigere Vorkommen von etwas hierarchiehöheren Fragen zum „Verständnis des Inhalts/Textes“ erklären kann.

Anhand der Ergebnisse kann aber auch kritisch diskutiert werden, ob das Potenzial der Fragen von Lehrpersonen – insbesondere in Schülerarbeitsphasen – optimal zur kognitiven Aktivierung der Lernenden genutzt wird. Evtl. erfordert gerade im ersten Schuljahr das Gestalten einer Schülerarbeitsphase noch so viel Aufwand und Aufmerksamkeit, dass den Lehrpersonen kaum Zeit oder kognitive Ressourcen bleiben, um in der Interaktion mit einzelnen Lernenden vertieft auf die Inhalte einzugehen (vgl. Abschnitt 2.2). Auch empirische Arbeiten, die explizit (auch oder ausschließlich) Schülerarbeitsphasen in den Blick nehmen, deuten darauf hin, dass Lehrpersonen deren Potenzial oftmals nicht ausschöpfen (z. B. Krammer, 2009; Lotz, 2015; van de Pol, Volman & Beishuizen, 2010). Es dominiert meist einfaches gegenüber elaborem Feedback, Hilfestellungen werden selten im Sinne des Scaffolding erteilt und die kognitive Aktivierung wird insgesamt als eher gering eingeschätzt.

Für die Lehrerinnen- und Lehrerbildung dürfte es daher wichtig sein, Lehrpersonen insbesondere für die Gestaltung von Schülerarbeitsphasen Strategien an die Hand zu geben, damit auch diese Phasen gezielter zur kognitiven Aktivierung der Lernenden genutzt werden können. Dass dies möglich ist, unterstreicht die in Abschnitt 2.2 vorgestellte Studie von Galton et al. (2008), in deren Rahmen Lehrpersonen ein erfolgreiches Training zu ihrem Lehrverhalten während Gruppenarbeiten erhielten.

Wichtig ist, dass die hier aufgezeigten Unterschiede selbst noch keinen Aufschluss über die Ursachen der Differenzen liefern können. Hier könnten weitere Untersuchungen ansetzen, in denen analysiert wird, ob beispielsweise Drittvariablen identifizierbar sind, die sowohl die Tendenz der Lehrperson erklären können, bestimmte Fragen zu stellen als auch den Unterricht vorzugsweise eher schüler- oder aber lehrerzentriert zu gestalten (z. B. Überzeugungen von Lehrpersonen; Kompositionsmerkmale der Schulklassen). Anhand von Beobachtungsdaten der PERLE-Studie konnten Denn et al. (2017) beispielsweise zeigen, dass Lehrpersonen in Klassen mit hoher Leistungsstärke häufiger schülerzentrierte Sozialformen nutzen als Lehrpersonen in Klassen mit geringerer Leistungsstärke. Kompositionsmerkmale könnten sich aber ebenso auf die Tiefenebene des Unterrichts auswirken und somit teilweise erklären, wieso es zwischen Oberflächen- und Tiefenmerkmalen zu systematischen Abhängigkeiten und Zusammenhängen

kommt. Dabei könnten sich auch Analysen dazu anschließen, inwiefern ein unterschiedliches Frageverhalten von Lehrpersonen in verschiedenen Phasen auch differenzielle Effekte auf das Lernen hat.

Einschränkend muss beachtet werden, dass mit dem kognitiven Niveau von Fragen nur ein ausgewählter Aspekt der Tiefenebene des Unterrichts untersucht wurde. Selbstverständlich gibt es noch eine Reihe weiterer Tiefenmerkmale, wozu v. a. Merkmale kognitiver Aktivierung gezählt werden können, wie beispielsweise die Art der im Unterricht gestellten Aufgaben, die Anregung der Lernenden zum Einsatz von Strategien oder die gemeinsame Reflexion des Lernprozesses (zsf. Lotz, 2015). Auch sollte bedacht werden, dass die Art der Fragen und die Häufigkeit deren Vorkommens im Unterricht allein natürlich noch keine hohe Unterrichtsqualität garantieren. Daher müssten die hier dargestellten Analysen noch vertieft und um weitere Aspekte ergänzt werden, indem sowohl anhand niedrig als auch hoch inferenter Methoden gezielt unterschiedliche Phasen beobachtet und verglichen werden. Die hier analysierten Unterrichtsphasen sind darüber hinaus relativ kurz. Der Unterrichtsinhalt war standardisiert, was die generelle Art der Fragen beeinflusst haben dürfte. Da den Lehrpersonen die methodische Umsetzung aber freigestellt war und sich hier auf der Oberflächenebene auch deutliche Unterschiede zwischen den Stunden zeigten, eignen sich die Daten dennoch gut zur Analyse der Fragestellungen.

Im vorliegenden Beitrag wurde die Verteilung von Fragen im öffentlichen Unterricht und in Schülerarbeitsphasen klassenübergreifend miteinander verglichen. Vertiefend könnte zusätzlich analysiert werden, inwiefern bei einzelnen Lehrpersonen das Frageverhalten in unterschiedlichen Phasen differiert. Eventuell lassen sich dabei sowohl Lehrpersonen identifizieren, die unabhängig von der Unterrichtsphase eher kognitiv anregende Fragen stellen, als auch Lehrpersonen, bei denen die Qualität der Fragen in Abhängigkeit von der Unterrichtsphase stark divergiert.

Zur Frage, ob sich die Tiefenebene des Unterrichts abhängig von den Oberflächenmerkmalen unterscheidet, liefern die vorliegenden Ergebnisse erste Hinweise. Dies ist insbesondere vor dem Hintergrund relevant, dass häufig die Ergebnisse unterschiedlicher Beobachtungsstudien oder auch mehrerer Unterrichtsstunden einer Lehrperson miteinander verglichen werden, ohne dass dabei dezidiert auf die Oberflächenebene des Unterrichts eingegangen wird. Die Ergebnisse der vorliegenden Studie zeigen aber – wenn auch zunächst lediglich anhand eines ausgewählten Aspekts – dass die Oberflächenmerkmale des Unterrichts nicht vernachlässigbar ist. So können Unterschiede zwischen den deskriptiven Ergebnissen verschiedener Studien auch dadurch zustande kommen, dass in jeder Studie ein anderes Unterrichtssetting beobachtet wurde. Hier wären allerdings noch mehr Studien nötig, die diese Fragestellung systematisch und für verschiedene Aspekte auf Tiefen- und der Oberflächenebene analysieren. Denn es ist sehr wahrscheinlich, dass für einige Tiefenmerkmale eine stärkere Abhängigkeit von der Oberflächenebene verzeichnet werden kann als für andere. In diese Richtung weisen auch Analysen von Praetorius (2014), die zeigen, dass insbesondere zur Beurteilung kognitiver Aktivierung mehrere Unterrichtsstunden verwendet werden sollten, um Abhängigkeiten der Qualitätsbeurteilung von der konkreten Unterrichtsgestaltung zu

minimieren. Während für die Beurteilung von Klassenführung und Schülerorientierung eine Stunde ausreicht, um Reliabilitäten  $\geq .70$  zu erreichen, werden zur reliablen Beurteilung des instabileren Merkmals der kognitiven Aktivierung eigentlich neun Stunden benötigt.

Für die Planung und Durchführung von künftigen Beobachtungsstudien kann aus den Ergebnissen der vorliegenden Studie v. a. abgeleitet werden, dass die ohnehin bereits häufig geforderte Standardisierung (z. B. Pauli, 2008) nicht nur aus inhaltlichen Gründen, sondern auch vor dem Hintergrund der Bedeutsamkeit unterschiedlicher Unterrichtsgestaltung auf Oberflächenebene wichtig sein dürfte. Natürlich bleibt dabei eine relevante Frage, wie viel Eingreifen in die übliche Planung der Lehrperson für die Forschungsfragen unabdingbar ist oder aber, inwiefern dadurch der übliche Unterricht zu stark verzerrt wird. Entscheidet man sich, wofür es durchaus gute Gründe geben kann, aber für ein geringeres Maß an Standardisierung, so sollte bei der Interpretation der Ergebnisse – selbst, wenn es dabei primär um die Tiefenmerkmale geht – der potenzielle Einfluss unterschiedlicher Oberflächenmerkmale mitbedacht oder kontrolliert werden.

## Literatur

- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1976). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz.
- Denn, A.-K., Hess, M., & Lipowsky, F. (2017). Hängen das Leistungsniveau und die Leistungsheterogenität von Grundschulklassen mit dem Anteil lehrerzentrierter Unterrichtsphasen im Deutsch- und Mathematikunterricht zusammen? Ergebnisse der PERLE-Studie. *Zeitschrift für Grundschulforschung*, 10(1), 162–176.
- Galton, M., Hargreaves, L., & Pell, T. (2009). Group work and whole-class teaching with 11- to 14-year-olds compared. *Cambridge Journal of Education*, 39(1), 119–140.
- Gayle, B. M., Preiss, R. W., & Allen, M. (2006). How effective are teacher-initiated classroom questions in enhancing student learning? In B. M. Gayle, R. W. Preiss, N. Burrell & M. Allen (Hrsg.), *Classroom communication and instructional processes: Advances through meta-analysis* (S. 279–293). Mahwah: Erlbaum.
- Hugener, I. (2008). *Inszenierungsmuster im Unterricht und Lernqualität. Sichtstrukturen schweizerischen und deutschen Mathematikunterrichts in ihrer Beziehung zu Schülerwahrnehmung und Lernleistung – eine Videoanalyse*. Münster: Waxmann.
- Hugener, I., Pauli, C., Reusser, K., Lipowsky, F., Rakoczy, K., & Klieme, E. (2009). Teaching patterns and learning quality in Swiss and German mathematics lessons. *Learning and Instruction*, 19(1), 66–78.
- Hugener, I., Rakoczy, K., Pauli, C., & Reusser, K. (2006). Videobasierte Unterrichtsforschung: Integration verschiedener Methoden der Videoanalyse für eine differenzierte Sicht auf Lehr- und Lernprozesse. In S. Rahm, I. Mammes & M. Schratz (Hrsg.), *Schulpädagogische Forschung. Unterrichtsforschung, Perspektiven innovativer Ansätze* (S. 41–53). Innsbruck: Studien-Verlag.
- Krammer, K. (2009). *Individuelle Lernunterstützung in Schülerarbeitsphasen. Eine videobasierte Analyse des Unterstützungsverhaltens von Lehrpersonen im Mathematikunterricht*. Münster: Waxmann.
- Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. Aronson (Hrsg.), *Improving academic achievement: Impact of psychological factors on education* (S. 135–158). San Diego: Academic Press.

- Levin, A. (2005). *Lernen durch Fragen. Wirkung von strukturierenden Hilfen auf das Generieren von Studierendenfragen als begleitende Lernstrategie*. Münster: Waxmann.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–105). Heidelberg: Springer.
- Lipowsky, F., Faust, G., & Kastens, C. (2013). *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren*. Münster: Waxmann.
- Lipowsky, F., & Lotz, M. (2015). Ist Individualisierung der Königsweg zum Lernen? Eine Auseinandersetzung mit Theorien, Konzepten und empirischen Befunden. In G. Mehlhorn, K. Schöppe & F. Schulz (Hrsg.), *Begabungen entwickeln & Kreativität fördern* (S. 155–219). München: kopaed.
- Lotz, M. (2013). Die Kodierung der Sozialformen. In M. Lotz, F. Lipowsky & G. Faust (Hrsg.), *Technischer Bericht zu den PERLE-Videostudien* (S. 123–142). Frankfurt a. M.: Gesellschaft zur Förderung Pädagogischer Forschung (GFPF).
- Lotz, M. (2015). *Kognitive Aktivierung im Leseunterricht der Grundschule. Eine Videostudie zur Gestaltung und Qualität von Leseübungen im ersten Schuljahr*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://www.springer.com/de/book/9783658104351> [25.09.2019].
- Lotz, M., & Corvacho del Toro, I. (2013). Die Videostudie im Fach Deutsch: „Lucy rettet Mama Krokodil“. In M. Lotz, F. Lipowsky & G. Faust (Hrsg.), *Technischer Bericht zu den PERLE-Videostudien* (S. 29–36). Frankfurt a. M.: Gesellschaft zur Förderung Pädagogischer Forschung (GFPF).
- Pauli, C. (2008). Unterrichtsbeobachtung. In F. Hellmich (Hrsg.), *Lehr-Lernforschung und Grundschulpädagogik* (S. 143–155). Bad Heilbrunn: Klinkhardt.
- Pauli, C., Drollinger-Vetter, B., Hugener, I., & Lipowsky, F. (2008). Kognitive Aktivierung im Mathematikunterricht. *Zeitschrift für Pädagogische Psychologie*, 22(2), 127–133.
- Praetorius, A.-K. (2014). *Messung von Unterrichtsqualität durch Ratings*. Münster: Waxmann.
- Spanhel, D. (1980). Analyse der verbalen Kommunikation im Unterricht. In K. Boeckmann (Hrsg.), *Analyse von Unterricht in Beispielen* (S. 83–97). Stuttgart: Klett.
- Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction* 4(4), 295–312.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296.

**Abstract:** According to published literature about instructional quality it can be assumed that deep structure features of teaching are the crucial factors for the assessment of instructional quality. Surface structure features provide a framework for the deep structure features. To analyse this assumed relationship empirically, this paper, using video data of  $N = 47$  reading lessons in first grade, examines whether surface features (here: whole class instruction vs. student work phases) have a connection to selected aspects of deep structure (here: cognitive level of teacher questions). Chi-square tests show that teacher questions differentiate considerably between whole class instruction and student work phases. Implications for video-based observation of instruction can be derived from this outcome (e.g. the significance of standardization).

**Keywords:** Surface Structure Features of Teaching, Deep Structure Features of Teaching, Instructional Quality, Teacher Questions, Elementary School

**Anschrift der Autor\_innen**

Dr. Miriam Hess, Friedrich-Alexander-Universität Erlangen-Nürnberg,  
Institut für Grundschulforschung,  
Regensburger Str. 160, 90478 Nürnberg, Deutschland  
E-Mail: miriam.hess@fau.de

Prof. Dr. Frank Lipowsky, Universität Kassel,  
Fachgebiet für Empirische Schul- und Unterrichtsforschung,  
Nora-Platiel-Straße 1, 34127 Kassel, Deutschland  
E-Mail: lipowsky@uni-kassel.de