

Wöhlke, Carina

Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte

Berlin : Logos Verlag 2020, XV, 273 S. - (Studien zum Physik- und Chemielernen; 298) - (Dissertation, Universität Hamburg, 2020)



Quellenangabe/ Reference:

Wöhlke, Carina: Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte. Berlin : Logos Verlag 2020, XV, 273 S. - (Studien zum Physik- und Chemielernen; 298) - (Dissertation, Universität Hamburg, 2020) - URN: urn:nbn:de:0111-pedocs-334030 - DOI: 10.25656/01:33403; 10.30819/5149

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-334030>

<https://doi.org/10.25656/01:33403>

in Kooperation mit / in cooperation with:



<https://www.logos-verlag.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
Internet: www.pedocs.de

Studien zum Physik- und Chemielernen

H. Niedderer, H. Fischler, E. Sumfleth [Hrsg.]

298

Carina Wöhlke

Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte



λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Hans Niedderer, Helmut Fischler und Elke Sumfleth

Diese Reihe im Logos-Verlag bietet ein Forum zur Veröffentlichung von wissenschaftlichen Studien zum Physik- und Chemielernen. In ihr werden Ergebnisse empirischer Untersuchungen zum Physik- und Chemielernen dargestellt, z. B. über Schülervorstellungen, Lehr-/Lernprozesse in Schule und Hochschule oder Evaluationsstudien. Von Bedeutung sind auch Arbeiten über Motivation und Einstellungen sowie Interessensgebiete im Physik- und Chemieunterricht. Die Reihe fühlt sich damit der Tradition der empirisch orientierten Forschung in den Fachdidaktiken verpflichtet. Die Herausgeber hoffen, durch die Herausgabe von Studien hoher Qualität einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Förderung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Hans Niedderer

Helmut Fischler

Elke Sumfleth

Studien zum Physik- und Chemielernen

Band 298

Carina Wöhlke

**Entwicklung und Validierung eines Instruments
zur Erfassung der professionellen
Unterrichtswahrnehmung angehender
Physiklehrkräfte**

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Hans Niedderer, Helmut Fischler, Elke Sumfleth [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Copyright Logos Verlag Berlin GmbH 2020

Alle Rechte vorbehalten.

ISBN 978-3-8325-5149-0

ISSN 1614-8967



Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90

Fax: +49 (0)30 / 42 85 10 92

<https://www.logos-verlag.de>

Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte

Dissertation zur Erlangung des akademischen Doktors der Philosophie am Fachbereich
Didaktik der gesellschaftlichen und mathematisch-naturwissenschaftlichen Fächer,
Fakultät für Erziehungswissenschaft, Universität Hamburg

vorgelegt von

Anna Carina Wöhlke

Geboren am 03. Februar 1990 in Hamburg

Gutachter_innen der Arbeit:

Prof. Dr. Dietmar Höttecke

Universität Hamburg, Fakultät für Erziehungswissenschaft

Prof. Dr. Sandra Sprenger

Universität Hamburg, Fakultät für Erziehungswissenschaft

Apl. Prof. Dr. Friederike Korneck

Goethe Universität Frankfurt, Institut für Didaktik der Physik

Hamburg, Dezember 2019

Danksagung

In den letzten vier Jahren hatte ich die Möglichkeit, viel in Bezug auf die Didaktik der Physik und über mich selber zu lernen. Wären einige Personen nicht gewesen, wäre mir diese Entwicklung nicht möglich gewesen. Aus diesem Grund möchte ich mich an dieser Stelle bei allen bedanken, die es in den letzten Jahren gut mit mir gemeint haben und mich in dem Unterfangen „Dissertation“ unterstützt haben.

Es folgt ein expliziter Dank an die drei wichtigsten Personen:

Ich danke meinem Betreuer **Dietmar Höttecke** für eine Betreuung, die vom ersten Augenblick an und bis zum letzten Wort wohlwollend, wertschätzend, unterstützend und fachlich sehr konstruktiv war.

Ich danke meinem Partner **Paul**, weil er mich nicht nur emotional durch diese Zeit getragen hat, sondern mit wenigen gezielten Fragen (häufig unwissentlich) diverse Knoten gelöst und mich fortwährend motiviert hat.

Ich danke meiner lieben **Mama** für alles, was sie mir mit auf den Weg des Lebens gegeben hat. Der größte Antrieb, nicht aufzugeben, auch wenn alles aussichtslos erschien, ist ihr Zuspruch gewesen.

Selbstverständlich gilt mein Dank auch dem **Bundesministerium für Bildung und Forschung**, das durch gezielte Förderung von Projekten (wie beispielsweise ProfaLe) Menschen wie mir die Chance auf akademische Weiterbildung gibt, aber auch der Bildung im Allgemeinen die Türen zur Weiterentwicklung öffnet.

Während ich versuchte, diese Danksagungen zu verfassen, ist mir aufgefallen, wie groß das Netz an unterstützenden Menschen während meiner Promotion war. In der unten stehenden Abbildung ist ein solches Netz abgebildet. Selbstverständlich kann dieses Netz nicht vollständig sein, zeigt aber wie viele kleine Bausteine etwas Großes ergeben. Mein Dank gilt daher neben meinen Gutachterinnen Sandra Sprenger und Friederike Korneck

auch den Teilnehmenden meiner Studie, Viktoria Alpers, Hannes Sander, Janne Langhof, Markus Feser, Olaf Uhden, Nadja Strunk, Annemarie Klemp, Nele Kroll, Johanna Ratzek, Timo Hackemann, Jenna Koenen, Britta Lübke, Carola Großmann, allen (ehemaligen) ProfaLe-Mitarbeitenden, Thomas Plotz, Valentin Wider, Anna Nowak, Maren Kempin, Marco Seiter, Christoph Kulgemeyer, Horst Schecker, Martin Schwichow, Knut Schwippert, Peter Reinhold, Josef Riese, Yvonne Gramzow, Heiko Krabbe, der Fakultät für Erziehungswissenschaft der Universität Hamburg und meiner Gemeinde und Jugendgruppe in Klein Flottbek.

Es sei auch allen Personen und Institutionen, die ich nicht namentlich genannt habe, abschließend gedankt.

				F															
				R															
C				I															
H				E		M									S				
R				D		A									A				
I				E	K	R									N				
S		T	O	R	N	T									D				
T	M	I	L	I	U	I	H		J		N				R		T	H	
O	A	M	A	K	T	N	E	J	A	N	E			V	A	N	H	A	
P	R	O	F	E	S	S	I	O	N	A	L	V	I	S	I	O	N		
H	K			K	C	C	K	H	N	D	E	A	C	P	N	M	N		
K	U			O	H	H	O	A	E	J		L	K	R	A	A	E		
U	S			R	W	W	K	N		A		E	I	E		S	S		
L				N	I	I	R	N				N		N					
G				E	P	C	A	A				T		G					
E				C	P	H	B					I		E					
M				K	E	O	B					N		R					
E					R	W	E												
Y					T														
E																			
R																			

Inhaltsverzeichnis

Danksagung	III
Abstract	IX
Abstract (english)	XI
Vorwort – Einordnung des Dissertationsprojekts in das Projekt „ProfLe“	XIII
1 Einleitung	1
I Theoretischer Hintergrund und Stand der Forschung	5
2 Theoretischer Hintergrund zur professionellen Unterrichtswahrnehmung	7
2.1 Professional Vision	8
2.1.1 Die Anfänge	8
2.1.2 Selective Attention und Knowledge Based Reasoning	10
2.2 PID	16
2.3 Kritik und Modifikation bisheriger Konzeptualisierungen	21
2.3.1 Eigene Konzeptualisierung	23
2.3.2 Professionelle Kompetenzen unter Kontextbeachtung	27
2.4 Angrenzende Konzepte: Reflexions- und Diagnosekompetenz	29
2.5 Zusammenfassung	33
3 Stand der empirischen Forschung	35
3.1 Studien zur professionellen Unterrichtswahrnehmung	36
3.2 Studien zur Instrumententwicklung zur Erfassung der professionellen Unterrichtswahrnehmung	40
3.2.1 Grenzen der präsentierten Instrumente	46
3.2.2 Zusammenfassung und Folgerung	46

II	Empirie	49
4	Empirische Vorüberlegungen	51
4.1	Forschungsfrage	51
4.2	Validität (Gültigkeit)	53
4.3	Forschungsdesign	57
4.4	Operationalisierung	60
4.4.1	Kritische Klassenraumsituation	60
4.4.2	Gegenstand der kritischen Situation	61
4.4.3	Der Wahrnehmungs-Prozess	66
4.4.4	Operationalisierung des Konzepts	67
5	Entwicklungsstudie I	69
5.1	Methode	69
5.1.1	Schritt 1 – Pilotierung mit einem ersten Video	70
5.1.2	Schritt 2 und 3 – Die Drehbücher	72
5.1.3	Schritt 4 – Das Kodierschema und die inhaltliche Validierung	84
5.1.4	Ergebnisse und Diskussion der inhaltlichen Validierung der Drehbücher	91
5.1.5	Schritt 5 – Fragebogen für die curriculare und ökologische Validität	96
5.1.6	Ergebnisse der curricularen und ökologischen Validierung der Drehbücher	97
5.2	Abschließende Diskussion und Zusammenfassung	100
5.2.1	Curriculare Validität	100
5.2.2	Ökologische Validität	100
6	Entwicklungsstudie II	103
6.1	Entwicklungsstudie II.1	103
6.1.1	Methode	104
6.1.2	Teilanalyse A: Prüfung der Passung und Schwierigkeit des Tests	106
6.1.2.1	Ergebnisse – Passung und Schwierigkeit des Tests	107
6.1.2.2	Interpretation und Diskussion – Passung und Schwierigkeit des Tests	109
6.1.3	Teilanalyse B: Entwicklung von Distraktoren	113
6.1.3.1	Ergebnisse – Entwicklung von Distraktoren	113
6.1.3.2	Interpretation und Diskussion – Entwicklung von Distraktoren	121
6.2	Entwicklungsstudie II.2	122
6.2.1	Methode	122
6.2.2	Teilanalyse A – Prüfung der Passung und Schwierigkeit des Tests	123
6.2.2.1	Ergebnisse – Passung und Schwierigkeit des Tests	123
6.2.2.2	Interpretation und Diskussion – Passung und Schwierigkeit des Tests	131

6.2.3	Teilanalyse B – Entwicklung von Distraktoren	134
6.2.3.1	Ergebnisse – Entwicklung von Distraktoren	134
6.2.3.2	Interpretation und Diskussion – Entwicklung von Distrak- toren	140
7	Hauptstudie	143
7.1	Methode	143
7.1.1	Das PUPiL-Instrument	143
7.1.2	Die Stichprobe	144
7.1.3	Die Erhebung zur Validierung	145
7.1.4	Die Kodierung für die Analyse der Daten	149
7.2	Der PUPiL-Test	158
7.2.1	Ergebnisse der IRT-Skalierung	158
7.2.2	Abhängigkeit der physikspezifischen professionellen Unterrichtswahr- nehmung von weiteren Variablen	171
7.2.3	Interpretation und Diskussion	174
7.2.4	Zusammenfassung	180
7.3	Der fachdidaktische Wissenstest	182
7.3.1	Ergebnisse der IRT-Skalierung	183
7.3.2	Interpretation und Diskussion	189
7.3.3	Zusammenfassung	191
7.4	Fassen eines Handlungsplans	192
7.4.1	Ergebnisse	192
7.4.2	Interpretation und Diskussion	194
7.4.3	Zusammenfassung	197
8	Abschließende Zusammenfassung und Diskussion	199
8.1	Zusammenfassung	199
8.2	Limitationen der Studie	204
8.2.1	Folgerungen für die Validität der Interpretationen	208
8.3	Diskussion	209
8.3.1	Implikationen für (physik-)didaktische Forschung	209
8.3.2	Implikationen für die Aus- und Fortbildung von Physiklehrkräften .	215
9	Schlusswort	219
	Literaturverzeichnis	223
A	Entwicklungsstudie I – Arbeitsblätter der Drehbücher	234
A.1	Drehbuch 2	235
A.2	Drehbuch 3	237
A.3	Drehbuch 4	242

B	Entwicklungsstudie I – Fragebögen für die curriculare und ökologische Validierung	244
C	Entwicklungsstudie II – Interviewleitfaden	249
D	Entwicklungsstudie II – Distraktoren Video 1, 3 und 4	258
E	Hauptstudie	268
	E.1 Fragen zu demographischen Daten 1	269
	E.2 Fragen zu demographischen Daten 2	271
F	Lister der digitalen Anlagen der vorliegenden Arbeit	272
G	Lister der aus der Dissertation hervorgegangenen Publikationen	XIX
	Abbildungsverzeichnis	XXI
	Tabellenverzeichnis	XXIII

Abstract

Die professionelle Unterrichtswahrnehmung ist eine situationsspezifische Fähigkeit von Lehrkräften, die für eine adaptive Lernunterstützung während des Unterrichts wichtig ist. Dieser Arbeit liegt die Forschungsfrage zugrunde, ob diese Fähigkeit mit einem physikspezifischen Fokus mithilfe von inszenierten Unterrichtsvideos und einem teilweise geschlossenen Fragebogen valide erfasst werden kann. Hypothesen, die helfen, diese Forschungsfrage zu untersuchen, lauten:

H1: Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften und von (angehenden) Geographielehrkräften sowie (angehenden) Sachunterrichtslehrkräften unterscheidet sich signifikant voneinander.

H2: Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften verändert sich im Studienverlauf signifikant.

Um die Forschungsfrage beantworten zu können, wird die professionelle Unterrichtswahrnehmung als ein dreidimensionaler Prozess operationalisiert, welcher mit den Teilprozessen Erkennen (Perception), Interpretation (Interpretation) und Handlungsplan Fassen (Decision Making) bezüglich physikspezifischer kritischer Unterrichtssituationen zwischen den Dispositionen und der Performanz von Lehrkräften mediiert (Blömeke, Gustafsson & Shavelson, 2015).

In der Entwicklungsstudie I wurden fünf Drehbücher für die Unterrichtsvideos entwickelt und inhaltlich, curricular sowie ökologisch durch Experten validiert. Die drei ausgewählten, validierten Drehbücher enthalten 38 kritische Situationen und wurden verfilmt.

In der Entwicklungsstudie II wurden die Videos als Stimuli für Phasen des lauten Denkens mit Studierenden (Fächer: Physik oder Geographie) und Physiklehrkräften (insgesamt $N = 22$) genutzt, um aus den Verbaldaten geschlossene Fragen für den Fragebogen in der Hauptstudie zu entwickeln und erste explorative Erkenntnisse über die physikspezifische professionelle Unterrichtswahrnehmung zu erhalten. Es entstanden 38 geschlossene Fragen in einem Ordered Multiple Choice Design mit vier Leveln.

In der Hauptstudie wurde der teilweise geschlossene Test $N = 160$ Testpersonen vorgelegt. Um neben der Inhaltsvalidität auch Aussagen über die Konstruktvalidität treffen zu können, wurden zusätzlich Daten zu dem Studium, der Unterrichtserfahrung und dem physikdidaktischen Wissen erhoben. Die Analyse mithilfe einer IRT-Skalierung ergab 19 Items (EAP-Reliabilität liegt bei 0.698). Weitere Analysen (zur Korrelation mit dem physikdidaktischen Wissen und der Abhängigkeit von studiumsbezogenen Daten) stützen die Konstruktvalidität des Tests. In Bezug auf H1 hat sich eine signifikante Abhängigkeit

des Personenschätzers für die professionelle Unterrichtswahrnehmung mit dem studierten Fach und dem Studiengang ergeben. Anders als erwartet (H2) hängt die professionelle Unterrichtswahrnehmung aber nicht signifikant vom Studienfortschritt ab. Außerdem hat sich gezeigt, dass ein schwacher Zusammenhang ($\tau = 0.093$) zwischen der professionellen Unterrichtswahrnehmung und dem physikdidaktischen Wissen zur Signifikanz tendiert ($p = 0.110$).

Abstract (english)

Professional vision is a situation-specific skill of teachers that is important for providing adaptive learning support during class. This study addresses the research question of whether this skill, focussing specifically on physics, can be validly assessed with the aid of scripted instructional videos and a questionnaire involving in part closed-ended questions. Hypotheses that help investigate this research question are:

- H1: The professional vision of (preservice) physics teachers and (preservice) geography teachers and of (preservice) teachers of general studies differs significantly.
- H2: The professional vision of (preservice) physics teachers changes significantly over the course of studies.

In order to be able to answer the research question, professional vision is operationalised as a three-dimensional process, which, in physics-specific critical situations, mediates between the dispositions and performance of teachers by means of the sub-processes of perception, interpretation and decision making (Blömeke, Gustafsson & Shavelson, 2015). In the first development study, five scripts for instructional videos were developed and validated by experts in terms of content, curriculum and ecology. The three selected, validated scripts contain 38 critical situations, and were filmed.

In the second development study, the videos were used as stimuli for periods of thinking aloud with students (subjects: physics or geography) and physics teachers ($N = 22$ in total), in order to develop closed-ended questions for the main study questionnaire from the verbal data and to obtain initial exploratory findings on professional vision in a physics-specific context. This resulted in 38 closed-ended questions in an ordered multiple choice design with four levels.

In the main study, the test of partially closed-ended questions was presented to $N = 160$ test subjects. In order to be able to make statements about the construct validity in addition to the content validity, additional data was collected on the course of study, and on the subjects' teaching experience and physics pedagogical content knowledge. Analysis using IRT scaling revealed 19 items (EAP reliability was 0.698). Further analyses (concerning the correlation with physics pedagogical content knowledge and the dependence on study-related data) support the construct validity of the test. With regard to H1, the estimator for professional vision was seen to depend to a significant degree on the studied subject and the course of studies. However, contrary to expectations (H2), professional vision is not significantly dependent on the degree of progress of the studies. In addition,

it was shown that a weak correlation ($\tau = 0.093$) seen between professional vision and physics pedagogical content knowledge tends towards being significant ($p = 0.110$).

Vorwort – Einordnung des Dissertationsprojekts in das Projekt „Profale“

Dieses Dissertationsprojekt ist im Rahmen der ersten Förderphase (2015–2018) des Projekts „Profale“ (**P**rofessionelles Lehrerhandeln zur Förderung **f**achlichen **L**ernens unter sich verändernden gesellschaftlichen Bedingungen) in der sogenannten „Qualitäts Offensive Lehrerbildung“, die vom Bundesministerium für Bildung und Forschung finanziert wurde, entstanden.

Das Projekt Profale hatte „eine Verbesserung der Lehrerbildung durch die curricular-inhaltliche Koordination der fachlichen, fachdidaktischen, pädagogischen und schulpraktischen Ausbildungsanteile der Ausbildung“ und die Überprüfung der Wirksamkeit dieser Maßnahmen zum Ziel¹. Das Projekt bestand aus vier sogenannten Handlungsfeldern, die verschiedene Schwerpunkte hatten und jeweils auf verschiedenen Fächern fußten.

Im *Handlungsfeld 1 – Kooperationen zwischen Fächern und Fachdidaktiken* war das anvisierte Ziel die Vernetzung des fachlichen und fachdidaktischen Wissens durch die Kooperation der jeweiligen Fachbereiche². Das Ziel ließ sich auf struktureller Ebene folgendermaßen differenzieren:

- Die Vernetzung der fachwissenschaftlichen und fachdidaktischen Ausbildungsinhalten und -strukturen soll verbessert werden.
- Die inhaltliche und methodische Zusammenarbeit von in der Fachdidaktik und Fachwissenschaft Tätigen soll verstetigt werden.
- Innovative Lehrveranstaltungskonzepten, in denen z.B. fachwissenschaftliche Inhalte unter den Aspekten von Aneignung und Vermittlung beleuchtet werden, sollen entwickelt und verstetigt werden.

Dazu wurden bestehende Kommunikationsstrukturen verbessert oder neue aufgebaut. Solche Kommunikation fand zum einen in individuellen Gesprächen zwischen den Lehrenden

¹Die hier zusammengefassten Informationen über das Projekt sind der Projekt-Homepage <https://www.profale.uni-hamburg.de> entnommen (zuletzt aktualisiert am 24.09.2019). Es wurden einige Satzteile direkt übernommen, weil sie die Ziele des Projekts am besten beschreiben. Auf ein explizites Kennzeichnen wurde dennoch im Sinne der besseren Lesbarkeit verzichtet.

²Im sogenannten „Hamburger Modell“ sind die Fachdidaktiken der erziehungswissenschaftlichen Fakultät und nicht der Fakultät des jeweiligen Faches zugeordnet. Dadurch sind die fach- und allgemein-didaktischen Ausbildungsanteile gut verknüpft und teilweise aufeinander abgestimmt. Dafür fehlen entsprechende Verknüpfungen zu den Fachdisziplinen.

im Lehramtsstudium und zum anderen in Gremien statt.

Eine weitere Maßnahme war die Entwicklung und Erprobung konkreter Lehrveranstaltungen innerhalb der Module des Lehramtsstudiums der beteiligten Fächer. Für die Verknüpfung wurden die Veranstaltungen entweder im Teamteaching von Dozierenden in Fächern und Fachdidaktiken abgehalten oder es wurden ergänzende Lehrveranstaltungen mit besonderer Berücksichtigung der Lehramtsstudierenden an bestehende Fachveranstaltungen angegliedert. Ob im Teamteaching oder ergänzend, wurde in den Lehrveranstaltungen untersucht, wie kognitive und affektiv-motivationale Dispositionen mit Fokus auf fachlichem Lernen im Laufe des Bachelorstudiums schrittweise aufgebaut werden können.

Die im Handlungsfeld Beteiligten haben einen zusammenfassenden Beitrag zu den entstandenen Lehrveranstaltungen veröffentlicht (Höttecke et al., 2018).

Ein zweiter Schwerpunkt im Handlungsfeld 1 war die Struktur und Entwicklung der professionellen Unterrichtswahrnehmung der angehenden Lehrkräfte. Diese und ihre Messbarkeit wurde in zwei Dissertationsprojekten untersucht, in denen die Forschenden eng miteinander kooperierten. Sowohl für die Geographiedidaktik (Nina Scholten, Sandra Sprenger) als auch für die Physikdidaktik (Carina Wöhlke, Dietmar Höttecke) wurden Instrumente entwickelt, mit denen die professionellen Unterrichtswahrnehmung in den entsprechenden Fächern erhoben werden kann. Es wurde insbesondere untersucht, ob die professionelle Unterrichtswahrnehmung fachspezifische Anteile hat und diese durch die entwickelten Instrumente erfassbar sind.

Im *Handlungsfeld 2 – Sprachlich-kulturelle Heterogenität* wurde untersucht, über welche Kompetenzen zur Gestaltung von Lehr-Lern-Gelegenheiten Studierende verfügen, die gemeinsames sprachliches und fachliches Lernen ermöglichen. Des Weiteren wurde erforscht, wie Lehramtsstudierende explizit und wirksam in die Lage versetzt werden können, sprachlich-kulturelle Heterogenität als Herausforderung und Chance zur Förderung von Bildungsprozessen wahrzunehmen (weitere Informationen sind auf der Homepage zu finden).

Ziel von *Handlungsfeld 3 – Inklusion* ist die Verschränkung der Wissensbestände und spezifischen Herangehensweisen der verschiedenen Bereiche (Schulpädagogik, Sonderpädagogik, Fachdidaktiken der Fächer: Deutsch, Mathematik, Geschichte, Sachunterricht), um die notwendigen Kompetenzen angehender Lehrkräfte der neuen und vielfältigen Herausforderung „Inklusion“ anzupassen (weitere Informationen sind auf der Homepage zu finden).

Im *Handlungsfeld 4 – Phasenübergreifende Kooperation* sollen im Rahmen des Kernpraktikums Kooperationsbeziehungen zwischen den universitären Lehrenden der Fachdidaktiken und den betreuenden und beratenden Personen der dritten Phase ausgebaut und neu aufgebaut werden (weitere Informationen sind auf der Homepage zu finden).

Die vorliegende Studie trägt in Form einer Dissertation im Handlungsfeld 1 dazu bei, über ein Instrument valide Aussagen und somit Erkenntnisse über die Messbarkeit, Struktur und Entwicklung der professionellen Unterrichtswahrnehmung von Physiklehramtsstudierenden ableiten zu können. Diese Erkenntnisse können auf lange Sicht in der Verbesserung des Lehrangebots an Universitäten in Bezug auf die Professionalisierung der Lehramts-

studierenden resultieren und hilft dem Gesamtprojekt damit teilweise beim Erreichen des Ziels, professionelles Lehrerhandeln zur Förderung fachlichen Lernens zu fördern.

1 Einleitung

*„The range of what we think and do
is limited by what we fail to notice.
And because we fail to notice
that we fail to notice,
there is little we can do
to change
until we notice
how failing to notice
shapes our thoughts and deeds.“
(Goleman, 1985, p. 24)*

Das oben stehende Zitat adressiert ein Problem, das für viele Professionen grundlegend ist. Sofern eine Profession Problemlöseprozesse umfasst, können nur diejenigen Probleme gelöst werden, die auch wahrgenommen werden. Damit ist das, was erreicht werden kann, dadurch begrenzt, was wahrgenommen wird.

Für Lehrkräfte ist es eine der alltäglichen Herausforderungen, Unterrichtssituationen wahrzunehmen, die für die Lernprozesse der Lernenden relevant sind. Diese Verarbeitung von Unterrichtssituationen durch Lehrkräfte beschreibt unter anderem das Konzept der professionellen Unterrichtswahrnehmung. Sherin, Jacobs & Philipp (2011) beschreiben die Unterrichtswahrnehmung beispielsweise als „process through which teachers manage the ‘blooming, buzzing confusion of sensory data’ with which they are faced, that is, the ongoing information with which they are presented during instruction“ (ebd.). Da die professionelle Unterrichtswahrnehmung eine grundlegende Fähigkeit von Lehrkräften ist, wurde in den letzten 20 Jahren genauer gefasst, was diese Fähigkeit ausmacht und inwiefern Lehramtsstudierende und Lehrkräfte über diese situative Fähigkeit verfügen (vgl. u.a. Floro & Bostic, 2017; Scheiner, 2016; Scholten, Höttecke & Sprenger, 2018; Seidel & Stürmer, 2014; Sherin, 2001; van Es & Sherin, 2002).

Die professionelle Unterrichtswahrnehmung wird als ein mehrdimensionaler, situativer Prozess verstanden, der innerhalb von Sekundenbruchteilen abläuft und kaum planbar ist. Häufig wurde allerdings nur die professionelle Unterrichtswahrnehmung in Bezug auf wenig fachspezifische Aspekte wie Lernbegleitung, Lernatmosphäre und Klassenführung untersucht. Es fehlt bisher an Erkenntnissen darüber, wie die professionelle Unterrichtswahrnehmung fachspezifischer Aspekte (beispielsweise Erkenntnisgewinnung und Schülervorstellungen) im Unterricht bei Lehramtsstudierenden ausgeprägt ist. Verschiedene Unterrichtsziele fordern dabei unterschiedliche Fokusse auf einen oder mehrere dieser As-

pekte. Es stellen sich – vor allem aus fachdidaktischer Sicht – die Fragen, wie eine Lehrkraft eigentlich wahrnimmt und was eine Lehrkraft in der Fülle an Ereignissen im Strom eines Unterrichts wahrnimmt. Wenn fachliche Lernprozesse der Lernenden nicht wahrgenommen werden, können Lehrkräfte diese nicht adäquat unterstützen.

Die Wahrnehmung beider oben genannten Arten von Aspekten (eher generische und physikspezifische Aspekte³) sind notwendig für einen lernförderlichen Unterricht. Eine Lehrkraft muss beispielsweise sowohl wahrnehmen, ob Maßnahmen der Aufmerksamkeitssteuerung angezeigt sind (Lernbegleitung) als auch, ob Lernende ein fachlich mehr oder weniger angemessenes Konzept des Magnetismus äußern (Schülervorstellungen). Das Ziel dieser Arbeit ist es daher, ein Instrument zu entwickeln, mit welchem die professionelle Unterrichtswahrnehmung angehender Physiklehrkräfte bezüglich physikspezifischer Aspekte in bestimmten Themenbereichen valide und reliabel gemessen werden kann. Aus diesen Daten kann Kenntnis über die professionelle Unterrichtswahrnehmung abgeleitet und adaptive Förderung dieser entwickelt werden.

Beim Erfassen der professionellen Unterrichtswahrnehmung als einer latenten Größe muss beachtet werden, dass die Erhebungssituation eine Realsituation so gut wie möglich abbildet, damit die Ergebnisse valide in Bezug auf die unterrichtliche professionelle Unterrichtswahrnehmung interpretierbar sind. Blömeke, Gustafsson & Shavelson (2015) nutzen die professionelle Unterrichtswahrnehmung, um Kompetenzen von Lehrkräften als Mediator zwischen Dispositionen und Performanz von Lehrkräften definieren zu können. Verschiedenes spricht dafür, dass es bisher kein (physikspezifisches) Messinstrument gibt, welches den Ansprüchen eines validen Instruments zur Erfassung der professionellen Unterrichtswahrnehmung gerecht wurde:

- Die professionelle Unterrichtswahrnehmung als situative und handlungsnahe Fähigkeit kann nur in ebenso handlungsnahen Situationen valide erhoben werden. Das bedeutet entweder im realen Unterricht oder anhand von Videos, die Unterricht glaubhaft abbilden.
- Da die professionelle Unterrichtswahrnehmung in nicht planbaren Situationen zum Tragen kommt und Lehrkräfte während des Unterrichtens unter Zeitdruck stehen, müssen die Testpersonen während der Erhebung in einen hohen zeitlichen Druck versetzt werden, ohne dass sie über das Gesehene reflektieren können. Dies ist mit einem Paper-Pencil-Test nicht möglich.
- Wie oben beschrieben, hängt die professionelle Unterrichtswahrnehmung von den Dispositionen der Lehrkräfte ab. Für die Konstruktion des Instruments sollten diese Voraussetzungen bereits für die Erhebungssituation (durch das Darbieten von Informationen über den Unterricht) beachtet und erhoben werden, um über Zusammenhangsanalysen Aussagen über Konstruktvalidität ableiten zu können.

³Es gibt weitere Aspekte, die der professionellen Unterrichtswahrnehmung bedürfen, wie beispielsweise die Erziehung, Inklusion oder sozialer Interaktion betreffend. Diese Aspekte sind aber nicht Gegenstand der vorliegenden Arbeit.

- Als situative Fähigkeit ist die professionelle Unterrichtswahrnehmung vom Kontext der Situation, in der sie angewendet wird, abhängig. Die Stimuli, die durch eine Situation geliefert werden, müssen laut Lamme (2003) drei Filter passieren: Ein Stimulus kann bemerkt werden (perceived), dann kann Aufmerksamkeit auf ihn gerichtet werden (attend to) und anschließend kann der Stimulus ins Bewusstsein der Lehrkraft dringen (awareness). Es kann allerdings nur die professionelle Unterrichtswahrnehmung derjenigen Stimuli beachtet werden, welcher das Bewusstsein der Lehrkräfte erreichen.

Um das Ziel der Arbeit unter den eben genannten Bedingungen zu erreichen, wird folgende Forschungsfrage untersucht: „Inwiefern kann die professionelle Unterrichtswahrnehmung angehender Physiklehrkräfte bezüglich Schülervorstellungen und Erkenntnisgewinnungsprozessen in den Themenbereichen Energieumwandlung, Elektrostatik und Wärmelehre von Lernenden beispielsweise der 8./9. Klasse valide und reliabel mithilfe eines Onlineinstruments, bestehend aus Videostimuli sowie offenen und geschlossenen Fragen, erfasst werden?“

Die vorliegende Arbeit bereitet diese Forschungsfrage in Teil I „Theoretischer Hintergrund und Stand der Forschung“ vor. In Kapitel 2 werden theoretische Überlegungen zur professionellen Unterrichtswahrnehmung vorgestellt. Anschließend wird das Konzept für die Interessen der vorliegenden Arbeit modifiziert und somit in eine eigene Konzeptualisierung überführt. Das zweite Kapitel endet mit der Einordnung der professionellen Unterrichtswahrnehmung in die professionellen Kompetenzen von Lehrkräften und mit der Abgrenzung von der Diagnose- und Reflexionskompetenz.

Im dritten Kapitel wird der Stand der empirischen Forschung zur professionellen Unterrichtswahrnehmung präsentiert. Neben Studien zum Experten-Novizen-Vergleich, zur Entwicklung der professionellen Unterrichtswahrnehmung angehender sowie praktizierender Lehrkräfte und zum Zusammenhang mit anderen Konzepten, wird der Fokus auf Studien zur Instrumententwicklung zur Erfassung der professionellen Unterrichtswahrnehmung gelegt.

Teil II der vorliegenden Arbeit widmet sich der Empirie zur Entwicklung des Messinstruments. Zunächst werden die hier angeklungenen empirischen Vorüberlegungen ausgeweitet und begründet. In diesem Rahmen werden Forschungsfrage, Hypothesen, Forschungsdesign und Operationalisierung der professionellen Unterrichtswahrnehmung weiter ausdifferenziert. Überlegungen, wie sich die Validität des Messinstruments sichern lässt, werden ebenfalls diskutiert.

Die empirische Arbeit dieser Studie teilt sich in drei Teilstudien auf. In Kapitel 5 wird die Entwicklungsstudie I vorgestellt, in der die Unterrichtsvideos entwickelt und validiert wurden. Dazu werden zunächst die Drehbücher und die Methode der Validierung beschrieben. Anschließend werden die Ergebnisse der Validierung präsentiert und diskutiert.

Es folgt in Kapitel 6 die Präsentation der Entwicklungsstudie II, welche zweigeteilt ist. Gegenstand dieser Teilstudie waren Interviews mit Anteilen des lauten Denkens mit Lehramtsstudierenden und Lehrkräften, in denen sie die entwickelten Unterrichtsvideos an-

gesehen und ihre Wahrnehmung mitgeteilt haben. Aus diesem verbalen Datenmaterial wurden durch eine qualitative Inhaltsanalyse Fragen für einen geschlossenen Fragebogen entwickelt. Neben der Fragebogenentwicklung wurden auch explorative Erkenntnisse über die professionelle Unterrichtswahrnehmung durch eine weitere qualitative Inhaltsanalyse abgeleitet. In Kapitel 6.1 wird die Methode der Entwicklungsstudie II.1 beschrieben, die Ergebnisse präsentiert und abschließend diskutiert. Im Anschluss wird analog die Entwicklungsstudie II.2 in Kapitel 6.2 präsentiert.

Kapitel 7 dient der Präsentation der Hauptstudie. Zunächst wird das Instrument, die Stichprobe und die Erhebung zur abschließenden Validierung beschrieben. Anschließend folgt die Beschreibung der Kodierung und Auswertung der gesammelten Daten. Daraufhin werden die Ergebnisse der Auswertung präsentiert und diskutiert. Die Daten wurden mit einer IRT-Skalierung und Zusammenhangsanalysen analysiert.

In Kapitel 8 wird die Studie zusammengefasst, außerdem werden Limitationen aufgezeigt und Implikationen für Forschung und Lehre abgeleitet.

Teil I

Theoretischer Hintergrund und Stand der Forschung

2 Theoretischer Hintergrund zur professionellen Unterrichtswahrnehmung

Stellen wir uns eine typische Stunde im Physikunterricht vor: In Gruppen bearbeiten die Lernenden ein Experiment. An einem Gruppentisch ist nur eine Schülerin aktiv, an anderen Tischen geraten die Experimentiermaterialien durcheinander, Lernenden laufen ungefragt durch den Physikraum, Hände werden erhoben und fragende Blicke ausgetauscht. In diesem nicht abbrechenden Strom der – zum Teil ungeplanten – gleichzeitigen Ereignisse ist es für Lehrkräfte schwer möglich, alle Aktivitäten wahrzunehmen und ihnen Aufmerksamkeit zu schenken. Es ist für Lehrkräfte kaum möglich, in diesem Strom an simultan stattfindenden Ereignissen die Einhaltung aller Merkmale guten Unterrichts sicherzustellen. Bisher gibt es wenig Erkenntnis darüber, wie Lehrkräfte im Verlauf einer Unterrichtsstunde entscheiden oder auf welche Aspekte des Unterrichts sie wann und warum ihre Aufmerksamkeit lenken. Es stellt sich die Frage, *wie* eine Lehrkraft ihre Wahrnehmung – bewusst oder unbewusst – lenkt. Aus fachdidaktischer Sicht steht fest, dass Lehrkräfte unter anderem diejenigen Unterrichtssituationen wahrnehmen sollten, in denen sich Gelegenheiten für Lernende bieten, fachlich über physikalische Begriffe, Konzepte und Methoden zu lernen.

Dieses Kapitel soll aus einer theoretischen Perspektive klären, wie die professionelle Unterrichtswahrnehmung konzeptualisiert und operationalisiert werden kann, um gemessen zu werden. Es gibt im Wesentlichen zwei Zugänge zur Konzeptualisierung der professionellen Unterrichtswahrnehmung. Seit Beginn der 2000er wird aus Sicht der Expertiseforschung abgeleitet, wie Lehrkräfte kritische Unterrichtssituationen wahrnehmen. Seit 2015 wird die professionelle Unterrichtswahrnehmung als Teil von Kompetenzmodellierungen betrachtet. Da die bisherigen Konzeptualisierungen überwiegend aus der Mathematikdidaktik stammen, folgt auf den theoretischen Befunden basierend eine Erläuterung der dieser Arbeit zugrundeliegenden Konzeptualisierung und ihre Einordnung in das Forschungsfeld. In Kapitel 3 wird der Stand der Forschung zur professionellen Unterrichtswahrnehmung in der Expertiseforschung und Kompetenzmodellierung dargelegt.

2.1 Professional Vision

2.1.1 Die Anfänge

„By applying such practices to phenomena in the domain of scrutiny, participants build and contest professional vision, socially organized ways of seeing and understanding events that are answerable to the distinctive interests of a particular social group.“ (vgl. Goodwin, 1994, S. 606)

Die scheinbar ersten Überlegungen zu einem sogenannten professionellen Blick werden dem Anthropologen Goodwin (1994) zugesprochen (vgl. z. B. Dyer, 2013; Jacobs, Lamb & Philipp, 2010; Lazarevic, 2016; Sherin, Russ, Sherin & Colestock, 2008; Treisch, 2018). Er nennt den professionellen Blick *Professional Vision* und beschreibt damit die Fähigkeit von Mitgliedern einer Profession zu lernen, Ereignisse, die diese Profession betreffen, wahrzunehmen und zu verstehen. Dieser Prozess ist in einen sozialen Kontext eingebettet und wird durch eine gesellschaftliche Gruppe geprägt (ebd., S. 606). Die professionelle Wahrnehmung wird durch die Prozesse *Kodieren*, *Hervorheben* und *Produzieren* sowie *Artikulieren* physischer Repräsentationen von Ereignissen, die für das Erkenntnisinteresse der jeweiligen Profession (gesellschaftlichen Gruppe) von Bedeutung sind, definiert (ebd., S. 626). Die bloße Wahrnehmung eines relevanten Objekts wird durch ein kognitives System organisiert (ebd., S. 609). Das bedeutet, dass die Wahrnehmung bereits durch ein System gelenkt wird, das beurteilt, welche Gegenstände *relevant* sind. Goodwin erläutert diese Prozesse anhand der Professionen der Archäologie und Polizei (ebd.). Die in der Archäologie wirkenden Fachkräfte betrachten ein Stück Grund und analysieren dank ihres professionellen Blicks die vor ihnen liegende Bodenbeschaffenheit. Sie rekonstruieren ganze Städte durch die Bestimmung von Farbe, Härte und Ansammlungen von Erde, indem sie diese kodieren, hervorheben und mit Hilfe der Erfahrung der Gruppe aller in der Archäologie wirkende Fachkräfte ihren Sinn bestimmen (produzieren) und mit der Gruppe teilen (artikulieren). Goodwin zeigt außerdem wie eine Gruppe von Polizeikräften den Mitgliedern der Jury und Richtenden im Gericht eine Videoaufzeichnung eines Gewaltaktes der Polizei gegenüber eines farbigen Amerikaners mit Hilfe ihres professionellen Blicks zu ihren Gunsten auslegt.

Ist die professionelle Wahrnehmung auch auf andere Professionen übertragbar? Durch die Übersetzung des Wortes *profession* aus dem amerikanischen Englisch entsteht möglicherweise eine ungewollte Unschärfe. Das Wort *profession* kann entweder mit *Profession* oder *Beruf* ins Deutsche übersetzt werden. Es ist demnach nicht klar, ob sich Goodwin (1994) auf Professionen oder Berufe bezieht. Geht man von der Übersetzung *Profession* aus, zählen nach Terhart (2011) heute diejenigen Berufe dazu, die in einer akademischen Laufbahn erworben werden und dabei humane, komplexe, soziale, technische und/oder wirtschaftliche Probleme des Klientels behandeln (vgl. ebd., S. 204). Das bedeutet, dass Menschen, die eine nicht-akademische Ausbildung absolviert haben, wie beispielsweise Handwerkernde, landwirtschaftlich Beschäftigte, oder Flugzeugführende keinen professionellen Beruf ausführen, wenngleich sie offensichtlich Expertise in ihrem Beruf aufweisen, die andere Gruppen der Gesellschaft nicht haben. Auch die Mitwirkenden dieser Berufsgruppen

haben es gelernt, für ihren Beruf relevante Ereignisse wahrzunehmen, zu verstehen und somit ihre Tätigkeit zielführend auszuführen. Dieser Prozess geschieht beispielsweise bei landwirtschaftlich Beschäftigten, wenn sie neue Felder anlegen, durch ...

... die Auswahl geeigneter Flächen (Kodierung) und

... das Hervorheben von Besonderheiten, die zu beachten sind (Hervorheben), sowie

... das graphische Produzieren von finalen Plänen für das Feld (Produktion und Artikulation).

Dabei beziehen Beschäftigte in der Landwirtschaft ihr Wissen, worauf sie achten und was sie wahrnehmen müssen, aus einem kognitiven System an Erfahrung und Wissen, welches in der landwirtschaftlichen Gemeinschaft entstanden ist. Aus diesen Überlegungen wird geschlossen, dass nicht nur Mitglieder einer Profession im engeren Sinne (Terhart, 2011), sondern auch Mitglieder von Berufen ohne akademische Laufbahn über professionelle Wahrnehmung verfügen können.

Die Fähigkeit Relevantes wahrzunehmen liegt zwar bei den Individuen, aber die Gemeinschaft kompetenter Praktiker teilt sich Wahrnehmungsmuster (ebd., S. 626). Goodwin erklärt weiter, dass es sich bei der professionellen Wahrnehmung nicht um einen reinen kognitiven Prozess handelt. Professionelle Wahrnehmung wird auch dadurch erlangt, dass der Beruf in relevanten Umfeldern praktiziert und ausgeübt wird (ebd., S. 626). Aus diesen beiden Erklärungen schließt Goodwin (1994), dass die professionelle Wahrnehmung erlernt werden kann und muss (ebd., S. 627).

Man kann von diesen Überlegungen ausgehend Schlüsse für das Lehren im Unterricht ziehen und eine *professionellen Unterrichtswahrnehmung* definieren. Dabei muss allerdings die Verschiedenheit der von Goodwin (1994) angeführten Professionen zu der von Lehrkräften beachtet werden: (a) Die bei Goodwin betrachteten Gegenstände der professionellen Wahrnehmung sind statische Ausgrabungsstätten der Archäologie und Videoaufzeichnungen über eine Gewalttat zwischen Polizei und einem farbigen Amerikaner, welche wiederholt gesehen und angehalten werden können. Im Unterricht adressiert die professionelle Unterrichtswahrnehmung der Lehrkräfte aber keine statischen Objekte – sondern Lernenden und deren Interaktionen untereinander und mit der Lehrkraft. Daher müssen Lehrkräfte innerhalb kürzester Zeit und jeden Augenblick erneut vorliegende Situationen und Änderungen wahrnehmen (Endsley, 1995). (b) Zudem haben Lehrkräfte nicht die Gelegenheit das Gesehene anzuhalten und/oder erneut anzusehen, um für ein tieferes Verständnis der Situation darüber nachdenken zu können (vgl. Sherin, 2001, S. 76). Lehrkräfte müssen in einer laufenden Unterrichtsstunde kritische Situationen identifizieren, welche für die Profession der Lehrkräfte von Bedeutung sind (ebd., S. 77). Dies sind vor allem Situationen, in denen Lernenden lernen⁴. (c) Auch die Art und Weise, wie in den von Goodwin (1994) beschriebenen Professionen Relevantes hervorgehoben wird (als Teil der professionellen Wahrnehmung), unterscheidet sich drastisch von den Möglichkeiten,

⁴Was Lernen von Lernenden im Physikunterricht bedeutet, wird im Kapitel Unterabschnitt 4.4.2 erläutert.

die Lehrkräften im laufenden Unterricht zur Verfügung stehen (vgl. Krupa, Huey, Kesseig, Casey & Monson, 2017). Anders als in der Archäologie Tätige und juristisch Richtende können Lehrkräfte die Situationen nicht farbig markieren, mit handschriftlichen Bemerkungen oder klebenden Notizzetteln versehen (vgl. Goodwin, 1994, S. 610), um später dahin zurückzukehren und der Situation einen Sinn zu entnehmen, sondern sie müssen sofort handeln oder sich merken, dass sie in Zukunft noch handeln wollten. Es ist überhaupt fraglich, wie sich Kodieren und Hervorheben in der professionellen Wahrnehmung von Lehrkräften unterscheiden, was genau sie produzieren und wem gegenüber sie dies artikulieren.

Nichtsdestotrotz kann Goodwins Konzept der professionellen Wahrnehmung mit sorgfältigen Adaptionen auf die Profession der Lehrkräfte übertragen werden. Resultate derartiger Adaptionen in Form von Konzeptualisierungen werden im Folgenden dargestellt und beurteilt, weil sie zum einen Einblick in die bisherige Forschung zur professionellen Unterrichtswahrnehmung geben und zum anderen theoretische Grundlagen für die Konzeptualisierung der professionellen Unterrichtswahrnehmung sind, die dieser Arbeit zu Grunde liegt.

2.1.2 Selective Attention und Knowledge Based Reasoning

„Here, I examine one area of teachers’ expertise, the ways in which teachers learn to interpret classroom events.“ (Sherin, 2001, S. 76)

In dem Artikel, der durch den eben zitierten Satz eingeleitet wird, hat die Mathematikdidaktikerin Miriam G. Sherin als eine der ersten erziehungswissenschaftlichen Forschende mit Bezug auf Goodwin (1994) der Expertiseforschung von Lehrkräften einen neuen Schwerpunkt in Form der professionellen Unterrichtswahrnehmung gegeben (vgl. Sherin, 2001)⁵. Wie zuvor geschildert, ist für Klassenraumsituationen charakteristisch, dass ein nicht abreißender Fluss an Ereignissen herrscht, welchen man – anders als ein Beweismittel in Form eines Videos vor Gericht – nicht anhalten kann. In den Veröffentlichungen bezüglich einer unterrichtsbezogenen professionellen Wahrnehmung wurden verschiedene Begriffe verwendet. Wie Weger (2019) passend zusammenfasst, changierten die Bezeichnungen für das gesamte Konzept oder Dimensionen der professionellen Unterrichtswahrnehmung zwischen *Noticing*, *Professional Vision* und *Selective Attention* (ebd., S. 17). Um mit der deutschsprachigen Literatur kompatibel zu sein, wird in dieser Arbeit der deutsche Begriff „professionelle Unterrichtswahrnehmung“ als Bezeichnung für das gesamte Konzept verwendet (Seidel & Prenzel, 2007; Stürmer, 2011; Meschede, Fiebranz, Möller & Steffensky, 2017). In dem anhaltenden Fluss an Ereignissen beschreibt die

⁵Sherin (2001) ergänzt, dass nicht nur Lehrkräfte professionelle Unterrichtswahrnehmung bezüglich Unterricht entwickeln, sondern beispielsweise auch erziehungswissenschaftlich Forschende, die (angehende) Lehrkräfte im Unterricht begleiten. Dieser professionelle Blick ist dabei ein anderer als der von Lehrkräften (ebd., S. 76f.). Lefstein & Snell (2011) merken dazu an, dass die Diskrepanz zwischen der Wahrnehmung von Lehrkräften und Forschenden zu einem Dilemma in der Erforschung der professionellen Unterrichtswahrnehmung führen kann und diese Diskrepanz in Entwicklung und Durchführung von Forschung daher immer mitgedacht werden muss.

professionelle Unterrichtswahrnehmung eine zweidimensionale Fähigkeit von Lehrkräften, relevante Situationen im Unterrichtsverlauf durch selektive Aufmerksamkeit zu erkennen, *Selective Attention* (erste Dimension) genannt, und diese mit dem Wissen über Lehren und Lernen interpretierend begründen zu können, *Knowledge Based Reasoning* (zweite Dimension) genannt (vgl. Sherin, 2001). Somit kann über den Stundenverlauf auch während des Unterrichts durch zielgerichtete professionelle Unterrichtswahrnehmung entschieden werden und nicht nur in der Unterrichtsvorbereitung. Das bedeutet, dass die Planung konkreter Schritte im laufenden Unterricht von der professionellen Unterrichtswahrnehmung abhängt. Die Merkmale guten Unterrichts, die in die Planung von Unterricht mit einfließen⁶, haben immer auch eine situative, prinzipiell nicht planbare Komponente, auf die nur durch die professionelle Unterrichtswahrnehmung reagiert werden kann.

Es ist an dieser Stelle ein Einschub nötig, der erläutert, was Wahrnehmung für ein Prozess ist, weil in viele Konzeptualisierungen von professioneller Unterrichtswahrnehmung eine Bedeutung des Begriffs der Wahrnehmung eher implizit mit eingeht. Um die Unterrichtswahrnehmung theoretisch beschreiben und praktisch unterstützen zu können, müssen die psychologischen und pädagogischen Grundlagen berücksichtigt werden, wie Scheiner (2016) in einem Kommentar zum Umgang mit dem Konzept der professionellen Unterrichtswahrnehmung betont.

Aus *psychologischer Sicht* ist Wahrnehmung das implizite oder explizite Lenken von Aufmerksamkeit auf Stimuli (vgl. Scheiner, 2016, S. 231f.). Anders als es hier den Anschein hat, ist Wahrnehmung ein interaktiver Prozess zwischen Personen und Stimuli (vgl. Lamme, 2003). Das *pädagogische Verständnis* von Wahrnehmung stützt sich auf diese Erkenntnisse der Psychologie und beschreibt Wahrnehmung als eine selektive, psychophysiologische Verarbeitung von Stimuli durch Sinnesorgane und Gedächtnis, welche dem Informationsgewinn dient und Grundlage für Orientierung und Handeln in der Umwelt ist (vgl. Tenorth & Tippelt, 2007, S. 100). Den Unterricht betreffend ist das Handeln das Lehren von Lernenden und die Umwelt der Klassenraum mit allen Agierenden. Die Wahrnehmung im Unterricht wird laut Tenorth & Tippelt (2007) durch sozial vermittelte Wertvorstellungen, Erwartungen und Haltungen beeinflusst. Das bedeutet, dass nicht bloß objektive Kategorien unsere Wahrnehmung beeinflussen, sondern auch individuelle subjektive Erwartungen an die Wirklichkeit (vgl. Mausfeld, 2010)⁷. Wahrnehmung kann in pädagogischen Kontexten entweder soziale Objekte wie Personen, Gruppen oder deren Verhaltensweisen und Merkmale als Gegenstand haben (interpersonale Wahrnehmung) und auch eine Bewertung und Beurteilung dieser umfassen. Im Unterricht gibt es selbstverständlich noch weitere Gegenstände wie die Un-

⁶Merkmale guten Unterrichts wurden sowohl theoretisch normativ als auch empirisch hergeleitet von zum Beispiel Lipowsky (2006), Kunter & Trautwein (2013), Helmke (2014), Hattie (2014), Beschluss der Kultusministerkonferenz (2004), Merzyn (2015), Kircher (2015) und Wagenschein (1982).

⁷Eine objektive Kategorie eines Stuhls könnte beispielsweise seine rote Lehne sein. Wie die rote Farbe der Lehne von Individuen wahrgenommen wird, beeinflussen aber individuelle subjektive Erwartungen.

terrichtsmaterialien oder die Atmosphäre, die wahrgenommen werden können. Der wahrgenommene Gegenstand kann ein sozialer Sachverhalt sein, in dem Personen ihre Rollen im Sozialgefüge aushandeln (soziale Wahrnehmung). Im Vergleich zum psychologischen Verständnis ordnet das pädagogische Verständnis der Wahrnehmung ein Ziel (nämlich das der Bewertung sozialer Kontexte) zu. Da die pädagogische Perspektive aus der psychologischen Perspektive entwickelt wurde, ist beiden die Annahme gemein, dass die Aufmerksamkeit, beziehungsweise die Wahrnehmung, durch subjektive Erwartungen gelenkt wird.

Die beiden Teilprozesse der professionellen Unterrichtswahrnehmung *Selective Attention* und *Knowledge Based Reasoning* sind daher zusammenhängend und zyklisch (vgl. Sherin et al., 2011, S. 5): Das Erkennen einer Unterrichtssituation, die zentral für das (Fach-) Lernen ist (im Folgenden kritische Situation genannt), beeinflusst die Interpretation dieser Situation. Die erlebte Situation und ihre Interpretation können wiederum beeinflussen, auf welche kritischen Situationen in der Folge Aufmerksamkeit gerichtet wird. Nichtsdestotrotz können diese beiden Teilprozesse theoretisch getrennt definiert werden.

Die Dimension der *Selective Attention* beschreibt das (bewusste oder unbewusste) Erkennen kritischer Situationen im Unterrichtsverlauf (vgl. van Es & Sherin, 2002). Das Erkennen geschieht innerhalb der Komplexität des Unterrichts. Die *Selective Attention* ist also eine Art Filter für die Verarbeitung der Reize, mit denen Lehrkräfte konfrontiert werden. Man kann den Prozess der Aufmerksamkeitslenkung psychologisch folgendermaßen beschreiben: Lehrkräfte „wählen“ im Unterricht einen oder mehrere der dargebotenen Reize. Anders als die Wortwahl impliziert, läuft diese Wahl nicht notwendigerweise bewusst ab und wird laut Gestaltpsychologie (vgl. Mausfeld, 2006, S. 102) von Erfahrungen, Gewohnheiten, Rolle und Orientierungen beeinflusst (vgl. Schoenfeld, 2011, S. 232). Diese situative Fähigkeit liefert dabei ein Spektrum vom Richten der Aufmerksamkeit auf einfach beobachtbare Situationen einerseits bis hin zu hoch inferenten Situationen andererseits (auf Sherin beziehend: Mason, 2011).

Häufig zeigen Situationen verschiedene kritische Inhalte gleichzeitig, aus denen Lehrkräfte auswählen müssen: Beispielsweise äußern Lernende physikalisch angemessene Konzepte, während andere Lernende den Unterricht durch laute, ungebetene Kommentare stören und andere wiederum unbeteiligt aus dem Fenster blicken. Die Lehrkraft kann nicht alle Reize gleichzeitig verarbeiten oder angemessen auf sie reagieren. Um handlungsfähig zu bleiben, ist es darum sinnvoll auszuwählen, ob sie der Äußerung eines angemessenen Konzepts, den Gründen für Ruhestörung oder der vermeintlichen Unaufmerksamkeit der Lernenden Aufmerksamkeit schenkt. Die eigenen Erwartungen von Personen bezüglich bestimmter Situationen determinieren, auf welche Ereignisse die Aufmerksamkeit von der Person gerichtet und wie diese interpretiert werden (Mausfeld, 2006; Neisser, 1976; Schoenfeld, 2011).

Die Dimension des *Knowledge Based Reasoning* beinhaltet die Komponente der Interpretation. Nach Floro & Bostic (2017) ist die Interpretation der ausgemachten kritischen Situation die Koordination des Gesehenen mit den allgemeinpädagogischen oder fachdi-

daktischen Theorien, die der Lehrkraft dazu bekannt sind (ebd., S. 76).

Ohne dass eine Lehrkraft eine kritische Situation wahrgenommen hat, ist eine Interpretation dieser nicht möglich (vgl. Males, 2017, S. 93f.). Die Interpretation ist eine wichtige Voraussetzung, um lernförderliche Entscheidungen zu treffen und den Fortgang der Stunde positiv zu beeinflussen (vgl. Sherin & van Es, 2008). Es gibt Forschungsgruppen, die die Tiefe des Knowledge based Reasoning von Lehrkräften nach Sherin (2001) in drei Teilfähigkeiten einteilen (vgl. Seidel & Stürmer, 2014, 745f.):

- a) Eine kritische Situation kann *beschrieben* werden, ohne dass die Lehrkraft eine Bewertung vornimmt.
- b) Eine kritische Situation kann *erklärt* werden, wobei die Lehrkraft ihr theoretisches und praktisches Wissen für eine Bewertung der Situation miteinbezieht.
- c) Der Fortgang nach einer kritischen Situation kann *vorhergesagt* werden, indem die Lehrkraft mit Hilfe ihres Wissens Konsequenzen für den weiteren Stundenverlauf ableitet und ihre Orientierungen mit dem Lernprozess der Situation in Verbindung bringt.

Diese Kategorisierung zeigt eine Graduierung von geringerer zu höherer Vernetzung von Fähigkeiten und Wissen, die häufig auf eine Betrachtung von erfahrenen und weniger erfahrenen Lehrkräfte übertragen wird (vgl. beispielsweise Blömeke, Kaiser & Clarke, 2015; Carter, Cushing, Sabers, Stein & Berliner, 1988; Erickson, 2011). So seien Junglehrkräfte und Studierende eher selten in der Lage, kritische Situationen nach dem Erkennen und Beschreiben adäquat zu erklären oder Vorhersagen für den Fortgang der Stunde zu treffen (vgl. Erickson, 2011).

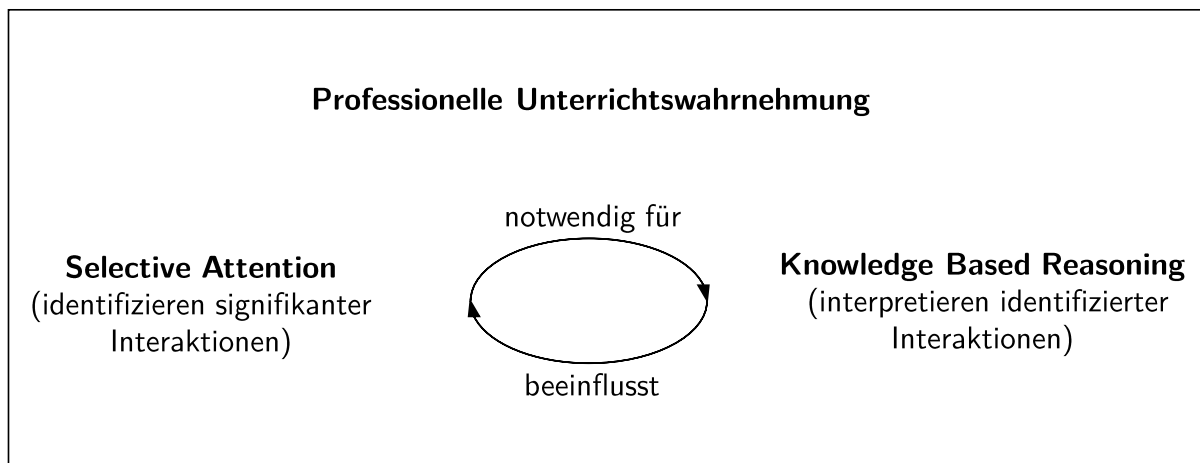


Abbildung 2.1: Die Teilprozesse der professionellen Unterrichtswahrnehmung Selective Attention und Knowledge Based Reasoning sind zusammenhängend und zyklisch (vgl. Sherin, Jacobs & Philipp, 2011).

Dies ist plausibel, da Junglehrkräfte und Studierende im Vergleich zu erfahrenen Lehrkräften zum einen durch geringere schulische Praxis einen Mangel an Erfahrung haben und zum anderen ihr theoretisches Wissen noch nicht ausreichend mit der Praxis ver-

knüpft haben. Es ist fraglich, ob Lehrkräfte – erfahrene Lehrkräfte oder Junglehrkräfte und Studierende – die erkannte Situation tatsächlich kognitiv beschreiben und erklären. Die hier beschriebenen Teilfähigkeiten bilden eher ab, was in einem Performanztest von Lehrkräften zur professionellen Unterrichtswahrnehmung erwartet würde, aber nicht was in der entsprechenden Situation im Unterricht kognitiv geschieht. Für Erickson (2011) scheint es selbstverständlich, dass die Aufmerksamkeitslenkung einen kritischen Schritt darstellt, der Junglehrkräfte und Studierenden von erfahrenen Lehrkräften trennt. Da das Ziel der professionellen Unterrichtswahrnehmung unter anderem darin besteht, Evidenz für das Lernen der Lernenden zu sammeln, dürfen die Körpersprache, der Tonfall und andere physische Zeichen des Lernens nicht missachtet werden: Sowohl auditive als auch visuelle Aspekte der kritischen Situation verdienen demnach Aufmerksamkeit und Interpretation (vgl. Wells, 2017, S. 186). Die professionelle Unterrichtswahrnehmung wird, wie die Erläuterungen gezeigt haben, von Sherin (2001) als aktiver, zweidimensionaler Prozess konzeptualisiert (siehe dazu Abbildung 2.1), der kein Teil des professionellen Wissens⁸ ist. Bei der professionellen Unterrichtswahrnehmung handelt es sich sowohl um einen Bottom-up-Prozess, in dem der Stimulus der Situation Aufmerksamkeit der Person erlangt, als auch um einen Top-down-Prozess⁹, in dem die Person aktiv die Aufmerksamkeit auf Stimuli der Situation richtet (vgl. Scholten et al., 2018; Treisch, 2018).

Man kann die professionelle Unterrichtswahrnehmung aber auch als einen Prozess modellieren, in dem sogenannte Situationstypen erkannt und die dazu passenden Handlungsalternativen ausgewählt werden (vgl. Wahl, 2001, S. 158). Diese Situationstypen sind komprimierende Repräsentationen für immer wieder auftretende, ähnliche Situationen im Unterricht, die eine adäquate und rasche Reaktion auf die Anforderungen der vielen kritischen, aber auf ähnliche Weise bereits durchlebten Situationen im Unterricht ermöglichen (ebd.). Die eigentliche Analyse von Unterrichtssituationen liegt demnach eher in dem Erkennen welcher Situationstyp vorliegt. Es muss keine von Grund auf neue Analyse stattfinden. Neben den Situationstypen, werden von Lehrkräften durch das Praktizieren auch Handlungstypen entwickelt, welche adäquates Handeln in den wiederkehrenden Situationen eines Typs erlauben (ebd., S. 159). Diese Handlungstypen beinhalten dabei situativ auszufüllende Platzhalter, wie Name, Gegenstand oder genaue Formulierung und Einbindung in den Redefluss, aber die Art der Reaktion (beispielsweise Aufforderungen, Maßregelungen oder ähnliches) steht fest (ebd.). Es konnte empirisch gezeigt werden, dass zu jedem Situationstyp mindestens ein, maximal aber sechs passende Handlungstypen wählbar sind¹⁰. Ist kein Handlungstyp für eine Ausprägung eines (neuen) Situationstypen vorhanden, so muss eine Handlungsmöglichkeit neu konstruiert werden. Dieser Prozess ist im Vergleich zur Auswahl eines Handlungstyps zeitaufwändiger (ebd., S. 160). Diese Annäherung an Wahrnehmungsprozesse und Handlungskontingenz ist gut vereinbar mit den Ergebnissen der Expertise-Forschung in Bezug auf die professionelle Unterrichtswahr-

⁸Unter dem professionellen Wissen wird die Taxonomie von Shulman (1986) verstanden.

⁹Siehe hierzu Scholten et al. (2018, S. 82f.) für eine detaillierte Darstellung der Prozesse.

¹⁰Dass die Handlungstypen wählbar sind, gibt auch einen Hinweis darauf, dass es sich nicht zwingend um eine Reiz-Reaktions-Kette im behavioristischen Sinn handelt. Die Lehrkraft wählt (unbewusst) aus diesen Handlungstypen aus.

nehmung (siehe unten Abschnitt 3.1).

Ähnlich zu dieser Modellierung integrieren auch Jacobs et al. (2010) und in Folge dessen viele weitere Forschungsgruppen (vgl. im Sammelband von Schack, Fisher & Wilhelm, 2017) die Entscheidung, wie auf die kritische Situation reagiert werden soll, *Deciding how to respond* genannt, in das Konzept der professionellen Unterrichtswahrnehmung (ebd., S. 172). Das Konzept bildet dann einen dreidimensionalen Prozess ab. Auch dieser Teilprozess ist nicht losgelöst vom Wissen, über das die Lehrkräfte verfügen, sodass die Entscheidung auch als Aspekt des Knowledge Based Reasoning verstanden werden kann (ebd., S. 173). Jacobs et al. (2010) betonen, dass es nicht *die* eine richtige Entscheidung infolge einer erkannten und interpretierten kritischen Situation gibt, es aber festzustellen ist, ob sich die (begründete) Entscheidung als konsistent mit den Theorien der aktuellen allgemeinpädagogischen und fachdidaktischen Forschung und lernförderlich für die Lernenden zeigt (ebd.). Es ist außerdem fraglich, ob die zu beobachtende Handlung¹¹ bereits Produkt des aktiven Begründungsprozesses ist, oder ob es sich dabei lediglich um ein Handlungsschema handelt, welches (meist unbewusst) durch einen erkannten Situationstyp evoziert wurde (vgl. Wahl, 2001). Zu der Frage, ob der Plan für eine (angemessene) Handlung dem Konzept der professionellen Unterrichtswahrnehmung inhärent ist, liefert Erickson (2011) das Argument, dass Lehrkräfte wahrnehmen, um Entscheidungen zu treffen, die die Lernenden beim Lernen unterstützen, sodass der Unterricht lerneffizienter gestaltet werden kann (ebd., S. 18f.). Die Art der Entscheidung hängt von der erkannten Situation ab. Handelt es sich beispielsweise um eine erkannte Schülervorstellung¹², die nicht mit dem kanonisierten Lehrbuchwissen konform ist, kann eine Lehrkraft beispielsweise die Thematik wiederholen, die Schülervorstellung zurückweisen oder andere Lernenden mit-einbeziehen (vgl. Floro & Bostic, 2017, S. 76). Selbstverständlich kann sich eine Lehrkraft auch dazu entscheiden, nicht (sichtbar) zu handeln.

Baldinger (2017) stellt außerdem heraus, dass sich die einzelnen Teilprozesse, die Sherin (2001) und Jacobs et al. (2010) vorschlagen, gegenseitig bedingen und daher die Veränderung einer der Komponenten Einfluss auf alle Komponenten der professionellen Unterrichtswahrnehmung haben kann (ebd., S. 233).

Zusammenfassend beschreibt die professionelle Unterrichtswahrnehmung einen zwei- oder dreidimensionalen Prozess, der von (professionellen) Wissen abzugrenzen ist, aber auf dieses schließen lässt und von diesem abhängig ist (vgl. Stürmer, Könings & Seidel, 2013). Sie umfasst, ...

- was eine Lehrperson während des Unterrichtens in einem Top-Down- oder Bottom-Up-Prozess erkennt: Zum einen können entweder einzelne Lernende, Gruppen oder die eigene Person wahrgenommen werden und zum anderen verschiedene Themen,

¹¹Die sich der Wahrnehmung und Interpretation anschließende Handlung der Lehrkraft wird von Jacobs et al. (2010) nicht in den Prozess der professionellen Unterrichtswahrnehmung einbezogen (ebd., S. 173).

¹²Was Schülervorstellungen sind, wird in Abschnitt 4.4 präsentiert. Selbstverständlich bezieht sich das Wort auf Vorstellungen aller Lernenden und nicht nur auf Schüler. Der Begriff „Schülervorstellungen“ ist aber ein etablierter BEgriff und wird daher auch in dieser Arbeit verwendet.

die beispielsweise pädagogischer oder fachdidaktischer Natur sein können (vgl. Lee & Choy, 2017, S. 124);

- wie eine Lehrperson das Erkannte interpretiert: Ereignisse im Unterricht können beschreibend, erklärend oder evaluativ – also mit oder ohne Evidenz – interpretiert werden (vgl. Lee & Choy, 2017, S. 124);
- welche Entscheidung eine Lehrperson trifft, um aus der erkannten und interpretierten Situation eine nutzbare und sinnhafte Lerngelegenheit für Lernenden zu schaffen, indem sie – den Lernprozess der Lernenden positiv oder negativ beeinflussend – handelt oder den Unterricht fortlaufen lässt.

Kritisch zu betrachten ist hier aus physikdidaktischer Perspektive, dass fast alle dieser Konzeptualisierungen aus der Mathematikdidaktik stammen und sich mit der professionellen Wahrnehmung vom mathematischen Denken der Lernenden im Mathematikunterricht beschäftigen. Ob und wie sich diese Konzepte auch für (angehende) Physiklehrkräfte bewährt, muss sich noch zeigen. Eine Übertragung auf den Physikunterricht wird bei Unterabschnitt 2.3.1 versucht.

Blömeke, Gustafsson & Shavelson (2015) greifen die Ergebnisse bezüglich der theoretischen Überlegungen zur professionellen Unterrichtswahrnehmung aus der Expertiseforschung auf und modellieren mit dieser Fähigkeit Kompetenzen von Lehrkräften. Im folgenden Abschnitt wird diese Modellierung erläutert.

2.2 PID

„Processes such as the perception and interpretation of a specific job situation together with decision-making [...] may mediate between disposition and performance.“ (vgl. Blömeke, Gustafsson & Shavelson, 2015, S. 7)

In dem Artikel „Beyond Dichotomies“ zeigen Blömeke, Gustafsson & Shavelson (2015) eine Diskrepanz zwischen verschiedenen Konzeptualisierungen von Kompetenzen der Lehrkräfte auf und generieren ein heuristisches Modell für die Kompetenz der Lehrkräfte. Die Diskrepanz zwischen den Konzeptualisierungen entsteht durch zwei verschiedene Perspektiven. Kompetenzen können unter anderem aus einer Verhaltensperspektive als Performanz verstanden werden oder aus einer kognitiven Perspektive als Dispositionen (ebd., S. 4f.).

Aus der Verhaltensperspektive heraus würde man Kompetenzen definieren, indem erfolgreiche Lehrkräfte beobachtet werden und ihr Verhalten als Exempel für eine kompetente Lehrkraft festgehalten wird (ebd., S. 4): „If someone wants to know who will make a good teacher, they will have to get videotapes of classrooms [...] and find out how the behaviors of good and poor teachers differ.“ (vgl. McClelland, 1973, S. 8). Statt vor und während des Studiums mithilfe von Ergebnissen von Intelligenztests über die Fähigkeit (angehender) Lehrkräfte urteilen zu lassen, würde die Eignung von Studierenden vielmehr über ihre Leistung in der tatsächlichen Handlung als Lehrkräfte festgestellt werden (ebd.,

S. 7-13). Eine Schwäche dieser Perspektive ist, dass sie es nicht erlaubt, Rückschlüsse über Gründe für das (kompetente) Handeln einzig aus der Beobachtung im Klassenraum abzuleiten (Stahnke, Schueler & Roesken-Winter, 2016). Es wäre also bekannt, was erfolgreiches Handeln ist, aber nicht welche Dispositionen dazu führen.

Aus der kognitiven Perspektive hingegen wird Kompetenz als eine personenbezogene, zugrundeliegende, kognitive Charakteristik und somit als Voraussetzung für gute Performanz definiert: „Based on top managers definitions of their competence he defined it as an underlying characteristic of a person which results in effective and or superior performance in a job (Boyatzis, 1982, p. 97).“ (vgl. Blömeke, Gustafsson & Shavelson, 2015, S. 5).

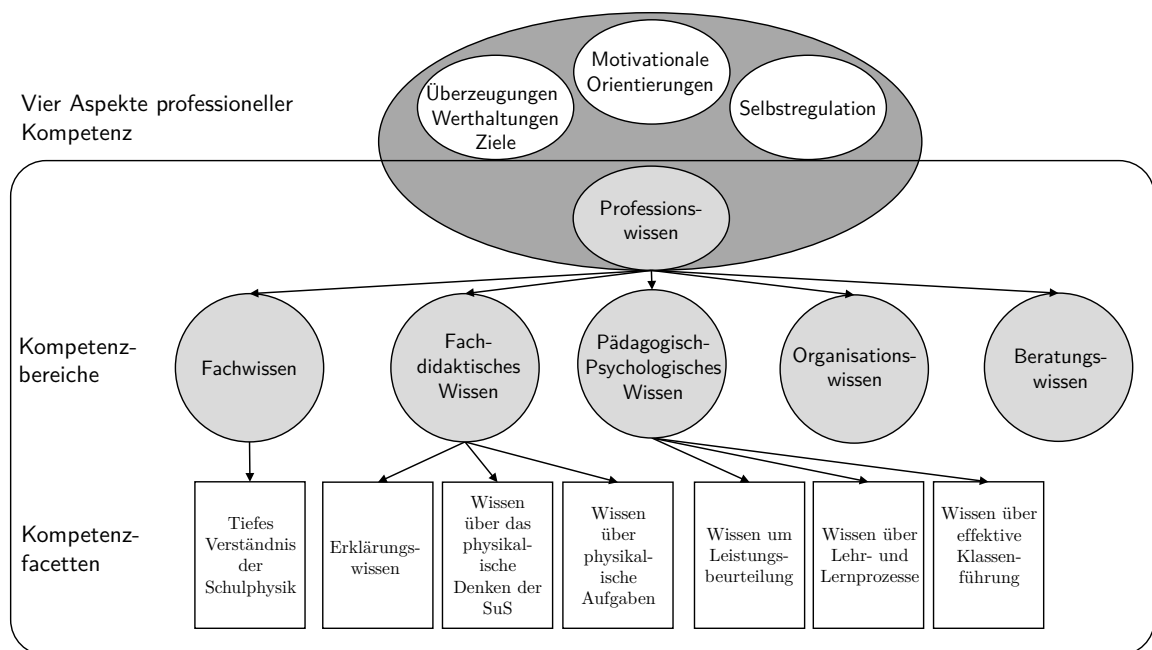


Abbildung 2.2: Modell für Kompetenzen von Lehrkräften adaptiert nach Kunter, Baumert & Blum (2011, S. 32) und für den Physikunterricht modifiziert durch Treisch (2018, S. 24).

Beispielsweise haben Kunter, Baumert & Blum (2011) im Rahmen des Projekts COACTIV ein generisches Strukturmodell der professionellen Kompetenzen von Lehrkräften entwickelt (ebd., S. 32). Es beinhaltet Überzeugungen, Werthaltungen, Ziele, motivationale Orientierungen, Selbstregulation und das Professionswissen von Lehrkräften¹³ (vgl. Abbildung 2.2) und basiert damit auf der Kompetenzdefinition von Weinert (2001a). Dieser definiert:

„[...] action competence includes all those cognitive, motivational, and social prerequisites necessary and/or available for successful learning and action.“ (ebd, S. 51)

¹³Für eine detaillierte Erklärung des Kompetenzmodells siehe Kunter et al. (2011) oder für die Ursprünge Baumert & Kunter (2006). Physikdidaktische Abwandlungen dieses Modells besonders in Bezug auf das Professionswissen sind beispielsweise bei Riese (2009) und Vogelsang (2014) zu finden.

Weinert (2001a) betont damit, dass diese sogenannte Handlungskompetenz (von Lehrkräften) intellektuelle Fähigkeiten, Fachwissen, kognitive Fähigkeiten, domänenspezifische Strategien, Routinen, motivationale Neigungen, volitionale Kontrollsysteme, persönliche Wertorientierungen und soziales Verhalten zu einem komplexen System an Dispositionen kombiniert (ebd.). Anders als die Definition von Kompetenzen der Verhaltensperspektive, gehen Weinert (2001a) und somit auch Kunter et al. (2011) davon aus, dass nicht das Trainieren der Tätigkeit alleine zu kompetentem Verhalten führt, sondern auch ein komplexes System an Dispositionen notwendig ist, um Aufgaben in speziellen professionellen Positionen erwartungsgemäß zu erfüllen (vgl. Weinert, 2001a, S. 51). Diese Perspektive erlaubt zwar Rückschlüsse auf die Kognitionen der handelnden Personen, bleibt dabei aber fern von realen Unterrichtssituationen und ignoriert somit den sozial-historischen Kontext und die Art und Weise, wie das Wissen von Lehrkräften mit der Umwelt interagiert (Stahnke et al., 2016).

Blömeke, Gustafsson & Shavelson (2015) kritisieren diese komplementären Perspektiven und schlagen vor, dass diese dichotomen, konkurrierenden Perspektiven auf Kompetenzen kombiniert werden sollten, um die jeweiligen Stärken für ein tieferes Verständnis kompetenten Handelns zu nutzen (ebd., S. 5). Dazu modellieren sie Kompetenzen mit Hilfe des Konzepts der professionellen Unterrichtswahrnehmung entlang eines sogenannten horizontalen Kontinuums (siehe Abbildung 2.3).

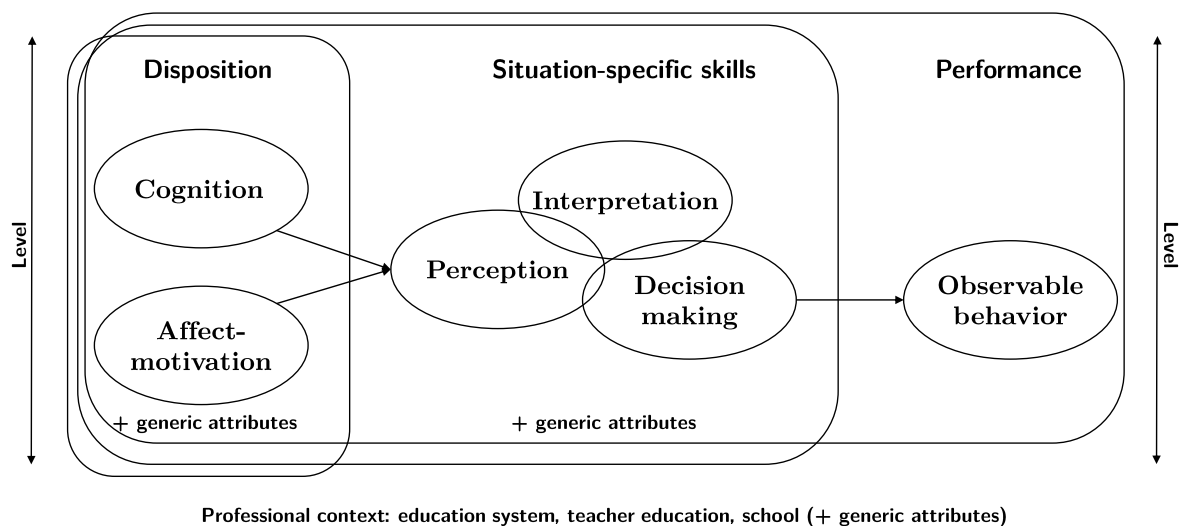


Abbildung 2.3: Das PID-Modell: Ein Kompetenzmodell mit den situationsspezifischen Fähigkeiten Perception, Interpretation, Decision Making Blömeke, Gustafsson & Shavelson (2015) (weiterentwickelt von Blömeke & Kaiser (2017)).

Das horizontale Kontinuum wird zwischen den beiden Polen der Dispositionen und der Performanz aufgespannt. Dabei umfassen die Dispositionen wie in der Definition von Weinert (2001a) kognitive und affektiv-motivationale Aspekte, sowie generische Attribu-

te (wie zum Beispiel Kreativität, ethische Verantwortung und Kommunikationsfähigkeit, vgl. Blömeke & Kaiser, 2017, S. 5). Die Performanz ist das beobachtbare Verhalten der Lehrkräfte. Blömeke, Gustafsson & Shavelson (2015) schlagen als Brücke zwischen diesen beiden Polen, die aus den verschiedenen Perspektiven auf Kompetenzen definiert wurden, situationsspezifische Fähigkeiten vor. Die oben beschriebene dreidimensionale Konzeptualisierung der professionellen Unterrichtswahrnehmung – *Perception* (Erkennen), *Interpretation* (Interpretieren) – *Decision Making* (Handlung Planen) – ist ein möglicher medierender Kandidat (vgl. Blömeke & Kaiser, 2017, S. 5), der Dispositionen in Performanz übersetzen kann (vgl. Blömeke, Gustafsson & Shavelson, 2015, S. 7). Aus dieser Modellierung geht wie in Unterabschnitt 2.1.2 zu Selective Attention und Knowledge Based Reasoning hervor, dass die situationsspezifischen Fähigkeiten Perception, Interpretation und Decision Making auf den Dispositionen der Lehrkräfte basieren und in einer beobachtbaren Handlung münden (ebd.). Die verschiedenen vertikalen Level, die Blömeke & Kaiser (2017) dem Modell hinzugefügt haben, entsprechen dabei den höheren oder niedrigeren Leistungsstufen sowie früheren und späteren Entwicklungsstadien in der Ausbildung und in der professionellen Ausübung des Lehrberufs (ebd., S. 6).

Man könnte in Anlehnung an Neuweg (2014) die beiden Pole als Wissen (Dispositionen) und Können (Performanz) bezeichnen. Die Abgrenzung des Wissensbegriff zum Können geschieht folgendermaßen: Das Professionswissen lässt sich in Wissen 1, Wissen 2 und Wissen 3 unterteilen (vgl. Neuweg, 2014, S. 584-587). Unter *Wissen 1* wird ein didaktisierbares, kodifiziertes, systematisches Wissen, welches in der Ausbildung anzueignen ist, verstanden (Wissen im objektiven Sinn). Dieses Wissen 1 lässt sich von Können, stabilen Persönlichkeitsmerkmalen, Erfahrungen, Prozessen der reflexiven Auseinandersetzung mit der eigenen Person, praktischen Erfahrungen und Angeboten des expliziten Wissens abgrenzen. Das Wissen 1 im objektiven Sinne kann distal über Prüfungsleistungen gemessen werden.

In Abgrenzung zu Wissen 1 ist *Wissen 2* ein psychologisches Konstrukt, welches kognitive Strukturen (subjektive Theorien, epistemologische Überzeugungen) in Folge von Lernen beschreibt (Wissen im subjektiven Sinn). Es ist die innere Erzeugungsgrundlage kompetenten Wissens und liegt oft in Schemata, Skripts und sogenanntem deklarativen und prozeduralen Wissen vor. Es entsteht über komplexe Prozesse der Transformation und Vernetzung von Informationen aus unterschiedlichen Wissensquellen (Erfahrungen, Reflexion, u.v.m.). Dieser Wissensbegriff bietet einen Übergang zum Denken, weil das Wissen bereits in der Handlung zum Tragen kommt. Wissen 2 ist im hohen Maße implizit, weshalb die Grenzen zum Können verschwimmen und es nur in abstrakter Form durch Befragungen oder situierte Aufgaben und deren Analyse rekonstruiert werden kann. Diese Grenze zum Können wird mit dem *Wissen 3* aufgehoben. Es definiert Wissen als Disposition und wird dem Wissenden/Handelnden von außen zugeschrieben. Das Wissen kann also demonstriert aber selten verbalisiert werden. Man kann dieses Wissen auch als das Wissen der Forschenden (Beobachtenden) zur Rekonstruktion der Logik des beobachteten Handelns verstehen. Dann wird dieses Wissen durch direkte Unterrichtsbeobachtungen oder Beobachtungen der Lösung unterrichtsnaher Problemstellungen gemessen. Hierdurch verschmiert der Unterschied zwischen Wissen 3 und Kompetenzdiagnostik.

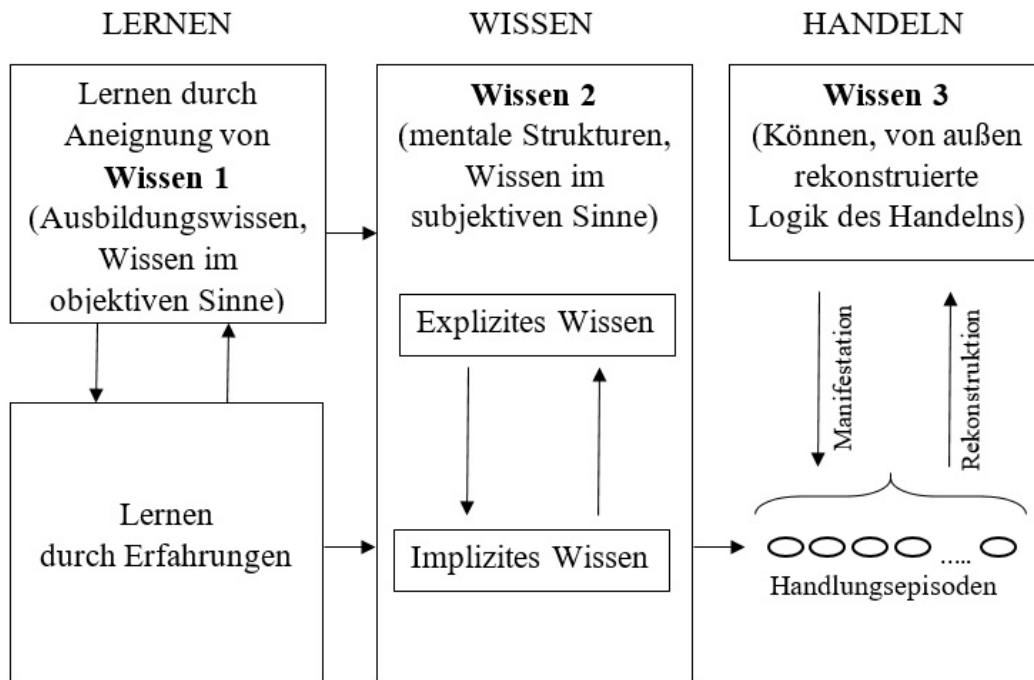


Abbildung 2.4: Konzepte des Lehrerwissen nach Neuweg (2014).

Die drei Wissensbegriffe sind nicht unabhängig voneinander (siehe Abbildung 2.4). Aneignung des Wissens 1 führt bereits zu Modifikationen der kognitiven Strukturen (Wissen 2). Die Spannung zwischen Wissen 1 und Wissen 3 wird in der Forschung unter dem Theorie-Praxis-Problem diskutiert, während der Zusammenhang von Wissen 2 und Wissen 3 das Verhältnis von Wissen und Können beschreibt (Neuweg, 2014).

Legt man dieses Verständnis von Wissen zur Überbrückung der beiden Pole zu Grunde, so umfassen Dispositionen Wissen 1 und 2, wohingegen die Performanz am ehesten dem Wissen 3 zugeordnet werden kann. Die oben beschriebenen situationsspezifischen Fähigkeiten sind nach den Ausführungen zu Wissen dem Wissen 2 zuzuordnen, da es sich sowohl beim Wissen 2 als auch bei den situationsspezifischen Fähigkeiten um Wissen handelt, das zum Teil prozedural ist, in Handlungen zur Anwendung kommt, auf Erfahrungen basiert und in Schemata sowie Skripten vorliegt (vgl. Wahl, 2001). Somit ist das von Blömeke, Gustafsson & Shavelson (2015) vorgeschlagene medierende Element wieder eine Disposition und das Problem der Überbrückung würde bestehen bleiben. Allerdings gibt es viele Prozesse, in denen ein Übergang von einem Konzept ins andere ohne kontinuierliches Medium erfolgt. Beispielsweise werden Gedanken (am ehesten Wissen) über die Hand (ein Körperteil) in Schrift (Performanz) mediert, ohne dass die Hand als Element eines Kontinuums verstanden wird.

Der These des horizontalen Kontinuums kann insofern vorgeworfen werden, dass zum einen nicht klar ist, was zwischen Dispositionen, situationsspezifischen Fähigkeiten und Performanz noch liegen soll und zum anderen, dass das Wissen ein anderes Konzept ist als das Können. Man kommt zwar auch durch das Wissen zum Können, aber entlang eines Kontinuums liegen diese beiden Aspekte nicht. Neben dieser bisher wenig diskutierten

Annahme, wurden weitere Elemente der oben dargestellten Konzeptualisierungen jüngst kritisiert und modifiziert. Das nächste Kapitel widmet sich dieser Kritik und Modifikation.

2.3 Kritik und Modifikation bisheriger Konzeptualisierungen

„Certainly, in any field as complex as teacher noticing is, it is difficult to develop deep theoretical understanding; however, we will not achieve this if we do not set our minds to it.“ (vgl. Scheiner, 2016, S. 237)

Scheiner (2016) fasst in einem Kommentar zusammen, welche Aspekte in der Debatte um die professionelle Unterrichtswahrnehmung nur zum Teil bedacht und betrachtet oder zu simplifiziert dargestellt werden und erweitert somit den Blick auf die bisherigen Konzeptualisierungen:

- Neben den individuellen Ressourcen, beeinflussen auch Ressourcen aus dem situativen Kontext das Handeln der Menschen (ebd., S. 228, 229).
- Die Entscheidungen von Lehrkräften (in Folge des Erkennens und Interpretierens) unterliegen psychologischen und sozial-kulturellen Einflüssen (ebd.).
- Es werden viele verschiedene Begriffe für mögliche Teilprozesse der professionellen Unterrichtswahrnehmung verwendet. Dabei bleiben die präzisen Bedeutungen und die Beziehungen zwischen diesen Begriffen oftmals unklar (ebd., S.230). So können die Beziehungen in den verschiedenen Studien als ...
 - ... linear und hierarchisch (Blömeke, Gustafsson & Shavelson, 2015; Bruckmaier, Krauss, Blum & Leiss, 2016),
 - ... aufeinander folgende Phasen oder Schritte (Dunekacke, Jenßen, Eilerts & Blömeke, 2016; Hoth et al., 2016; Pankow et al., 2016), oder
 - ... zyklisch (als andauernder Prozess, vgl. Dyer & Sherin, 2016; Santagata & Yeh, 2016) verstanden werden.
- Die quantitative Betrachtung, die häufig für die Beurteilung der professionellen Unterrichtswahrnehmung verwendet wird (vgl. unten Kapitel 3), erklärt das Phänomen der professionellen Unterrichtswahrnehmung nicht ausreichend (ebd.).

Folgende Studien sollten demnach besonderes Augenmerk darauf legen, dass alle Begriffe und Relationen – soweit möglich – transparent sind.

Im folgenden Absatz soll eine Konzeptualisierung der professionellen Unterrichtswahrnehmung vorgestellt werden, die verwendete Begriffe und deren Relationen genauer erklärt als bisherige und den Kontext des zu konzeptualisierenden Konstrukts vermehrt miteinbezieht. Die Konzeptualisierung modifiziert das PID-Modell (Santagata & Yeh, 2016). In

dem modifizierten PID-Modell überlappen die Dispositionen – Wissen und Überzeugungen – mit der Unterrichtspraxis. In der Mitte dieser Überlappung stehen die situationspezifischen Fähigkeiten, welche Lernende, Unterrichtsmaterialien und die Lehrkräftegemeinschaft adressieren. Wissen und Überzeugungen werden im Unterricht also erst durch die situationsspezifischen Fähigkeiten relevant, beeinflussen diese aber gleichzeitig. Durch die Anwendung der situationsspezifischen Fähigkeiten wiederum werden die Dispositionen der Lehrkräfte verändert. Es handelt sich in diesem Modell um einen zyklischen Prozess. Daraus kann geschlossen werden, dass das Praktizieren von Unterricht sowohl die situationsspezifischen Fähigkeiten als auch andere Dispositionen fördert (vgl. Scheiner, 2016, S. 229). Santagata & Yeh (2016) verwenden die Begriffe Wissen und Überzeugungen statt Dispositionen, um klar zu machen, dass es sich weniger um an Personen verhaftete, als vielmehr um sozial und kontextuell durch Lehrkräfte, Lernenden und Lehrkräfteauszubildende ausgehandelte Attribute handelt. Wissen und Überzeugungen gibt es nach Santagata & Yeh (2016) auch jenseits von der Unterrichtspraxis, weshalb dies im Modell auch so dargestellt wird.

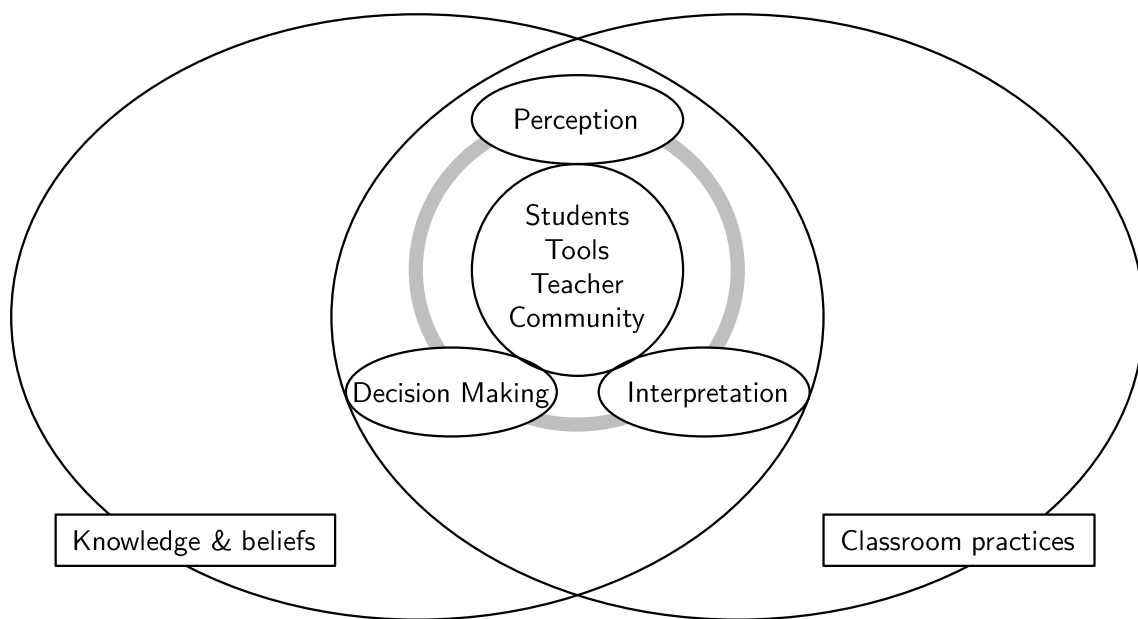


Abbildung 2.5: Überarbeitetes PID-Modell von Kompetenzen der Lehrkräfte in Anlehnung an Santagata & Yeh (2016).

Diese Aspekte von Wissen und Überzeugungen, welche in Abbildung 2.5 überlappungsfrei sind, kommen in der Unterrichtssituation nicht zum Tragen¹⁴. Nach Scheiners Kommentar (2016) ist der größte Zugewinn der Modifikation durch Santagata & Yeh (2016) die Anerkennung der Abhängigkeit von Individuum und Umwelt (ebd., S. 229). Bringt man die in

¹⁴Es ist anzumerken, dass Wissen und Überzeugungen, die in Unterrichtssituationen aber nicht zum Tragen kommen, entweder in der Tat nicht relevant sind, oder aber ihre Relevanz von den handelnden Personen ignoriert oder übersehen wird.

Unterabschnitt 2.1.2 zu Selective Attention und Knowledge Based Reasoning vorgestellte professionelle Unterrichtswahrnehmung in Verbindung mit der Kompetenzmodellierung von Lehrkräften, erhält man ein heuristisches Modell von Kompetenzen der Lehrkräfte. Dieses setzt keinen künstlichen Schwerpunkt auf Dispositionen oder Performanz und erklärt die Übersetzung zwischen diesen mit den situationsspezifischen Fähigkeiten Perception, Interpretation und Decision Making. Durch die Verbindung dieser Konzeptualisierung mit (sozialem) Kontext und Umwelt erhalten wir ein erklärungsstarkes Konzept von situativen, unterrichtlichen Handlungen von Lehrkräften, welches möglicherweise die Förderung von kompetentem Handeln der Lehrkräfte unterstützen kann.

2.3.1 Eigene Konzeptualisierung

Um den Forderungen von Scheiner (2016) nach einer begrifflichen Klarheit nachzukommen, wird im folgenden Abschnitt die Konzeptualisierung der professionellen Unterrichtswahrnehmung, welche dieser Arbeit zu Grunde liegt, vorgestellt¹⁵. Es wird besonderer Fokus auf die Klärung der verwendeten Begriffe und deren Relation zueinander gelegt. Die professionelle Unterrichtswahrnehmung kommt in sogenannten kritischen Situationen zum Tragen. Diese werden folgendermaßen konzeptualisiert:

Kritische Situationen Eine pädagogische Situation ist „die Gesamtheit der Reizfaktoren, welche ein Wesen nötigen, als eine Gesamtheit zu reagieren“ (vgl. Petersen, 1946, S. 59) und vor allem erziehende, bildende oder betreuende Aspekte betrifft. Situationen können sich dabei entweder ungeplant ereignen (funktional) oder werden (durch die Lehrkraft) bewusst herbeigeführt (intentional, vgl. Stöcker, 1975, S. 272f.). Innerhalb einer solchen Situation können demnach mehrere Ereignisse (Reize) gleichzeitig auftreten, welche verschiedene Aspekte des Unterrichts adressieren (beispielsweise fachdidaktische, allgemeindidaktische oder allgemeinpädagogische Aspekte) und sich überlagern.

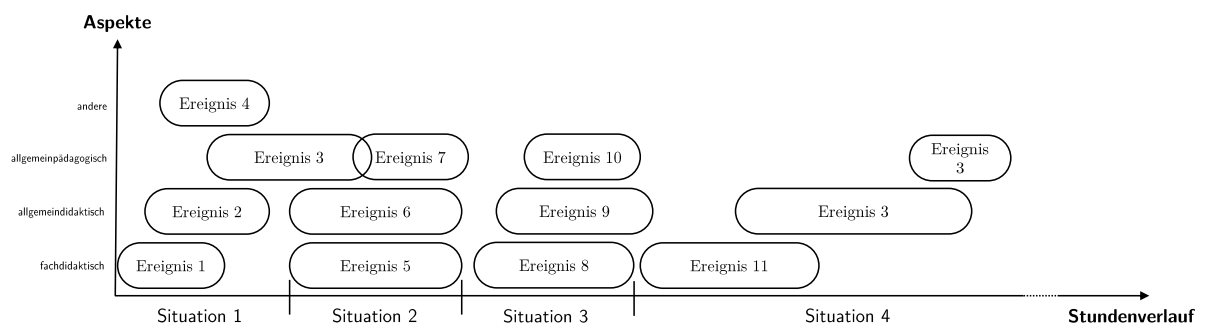


Abbildung 2.6: Mögliche Abfolge von Ereignissen und Situationen im Unterrichtsverlauf.

¹⁵In Abschnitt 4.4 folgt eine Operationalisierung dieser Konzeptualisierung für die Entwicklung eines Messinstruments.

In Abbildung 2.6 ist eine mögliche Konstellation von Ereignissen innerhalb von Situationen während einer Unterrichtsstunde skizziert. Wird das Leben nicht holistisch konzipiert, sondern als in Situationen zerstückelt, kann es scheinen, dass nach Abschluss dieser Ereignisse die Situation beendet ist. Es ist jedoch auch möglich, dass ein Ereignis einer Situation noch nicht beendet ist und trotzdem ein Übergang in eine neue Situation stattfindet. Somit gibt es die Möglichkeit, dass sich Situationen überlappen. Möglicherweise überschneiden sich Ereignisse (verschiedener Situationen) zum gleichen Aspekt.

Dimensionen der professionellen Unterrichtswahrnehmung Diese Arbeit folgt in Anlehnung an das PID-Modell der Annahme, dass die professionelle Unterrichtswahrnehmung aus drei Dimensionen besteht und eine situationsspezifische Fähigkeit ist (vgl. Blömeke, Gustafsson & Shavelson, 2015). Diese Dimensionen werden Perception, Interpretation und Decision Making genannt. Relationen dieser Dimensionen werden in Abbildung 2.7 zusammenfassend dargestellt.

Perception Ähnlich wie beim PID-Modell gehen in dieser Konzeptualisierung der *Perception* weitere Dispositionen voraus (Blömeke, Gustafsson & Shavelson, 2015). Diese beinhalten das Professionswissen und Überzeugungen, die den geplanten Unterricht betreffen. Darüber hinaus wird Perception durch Erwartungen an die Situation (zum Beispiel Zielsetzungen für die Unterrichtsstunde) und Erfahrungen mit ähnlichen Situationen geprägt (vgl. Neisser, 1976). Die Erwartungen können aus eigenen Überlegungen oder implizit sowie explizit aus Gesprächen mit anderen Lehrkräften resultieren (vgl. Santagata & Yeh, 2016). Die subjektiven Wissensstrukturen, die sich durch das Praktizieren von Unterricht aufbauen, haben in jeder Situation einen nicht zu vernachlässigenden Einfluss. So hilft einem Wissen über Schülervorstellungen zu einem konkreten Konzept, diese zu erkennen und zu interpretieren, wohingegen Erwartungen an die Klassenraumsituationen dazu beitragen können, dass gewisse Ereignisse (nicht) wahrgenommen werden. Es gehen demnach viele Annahmen, was in der kommenden Situation geschehen kann, mit in die professionelle Unterrichtswahrnehmung der Lehrkräfte ein. Die Perception der Situationen beinhaltet dabei

- das Erkennen der Situationsstimuli, welche das Auge erreichen und im Gehirn verarbeitet werden und
- das Richten von Aufmerksamkeit, das heißt die Aufmerksamkeit auf „relevante“ Ereignisse fokussieren, „irrelevante“ Ereignisse ausblenden und Ausschnitte der Situation wählen.

Diese Interaktion zwischen Stimulus der Situation und wahrnehmender Person läuft häufig unbewusst ab (vgl. Scholten et al., 2018; Treisch, 2018; Wahl, 1991). Informationen über den Kontext der Situation – wie beispielsweise die Leistungen der einzelnen Lernenden, des vorangegangenen Unterrichts, Vorkommnisse des Tages und Verhältnisse zwischen partizipierenden Personen – haben eine zentrale Rolle in der Aufmerksamkeitslenkung (vgl. Neisser, 1976). Resultat dieses Teilprozesses ist, dass einige der simultan ablaufen-

den Ereignisse fokussiert werden. Nur diese werden in den nächsten beiden Teilprozessen weiterverarbeitet. In Bezug auf Lehrkräfte ist es denkbar, dass vor allem jene Ereignisse einer Situation bemerkt werden, die in Verbindung zu ihrer Zielsetzung für den Unterricht stehen und ihrem Professionswissen, ihren Erfahrungen und Erwartungen entsprechen.

Interpretation Sowohl Perception, als auch Interpretation einer Situation werden von den Wissen 1, 2 und 3 beeinflusst. Nur Ereignisse von Situationen, die durch Perception Aufmerksamkeit erlangt haben, können auch weiter interpretiert werden. Zu Interpretation gehören die folgenden (analytisch, normativ gesetzten) Schritte (in Anlehnung an Seidel & Stürmer, 2014):

- Erfassen der relevanten Elemente eines Ereignisses,
- Zusammenhänge der erfassten Elemente erkennen und (mit Wissen) erklären,
- Ausprägung der Elemente und ihrer Zusammenhänge im Hinblick auf die Unterrichtsziele beurteilen und
- die Ereignisse zu bisherigen Erfahrungen und Erwartungen in Bezug setzen.

Diese Schritte können dazu führen, dass die Ereignisse, die Aufmerksamkeit erlangt haben, bezüglich der Dringlichkeit der Weiterverarbeitung beurteilt werden. Auch die Interpretation wird von dem Kontext, der die Situation umgibt, berührt. Die Äußerung eines physikalisch nicht angemessenen Konzepts zum Kraftbegriff würde eine Lehrkraft weniger besorgniserregend interpretieren, wenn diese zu Beginn einer Einheit über die Thematik der Newton'schen Axiome getätigt wird, als in einer abschließenden Sitzung dazu. Neben der Perception beeinflussen auch die Zielsetzung und das Professionswissen die Interpretation. Da gerade das Wissen 2 durch Praktizieren verändert wird (vgl. Neuweg, 2014), kann die Interpretation der Ereignisse einer Situation zur Modifizierung des Professionswissen beitragen. Dieses bedingt wiederum die Perception folgender Situationen, sodass man den Prozess der professionellen Unterrichtswahrnehmung schon durch die Betrachtung nur dieser beiden Dimensionen als spiralförmig verstehen kann (vgl. Abbildung 2.6).

Decision Making Der dritte Teilprozess der professionellen Unterrichtswahrnehmung ist das *Decision Making*. Dieses umfasst das Generieren, Bewerten und Wählen von Handlungsoptionen auf Grundlage der Interpretation der wahrgenommenen Situation. Diese Schritte werden wiederum unter anderem vom Kontext der Unterrichtssituation gelenkt (vgl. Santagata & Yeh, 2016). Die Dispositionen, die die Situation und deren Ereignisse betreffen, beeinflussen auch das Decision Making. Die auf Basis des Decision Making folgende Handlung wird nicht als Teil der professionellen Unterrichtswahrnehmung konzipiert, diktiert aber unter anderem den Fortlauf der Unterrichtsstunde und hat somit Einfluss auf die folgenden Gelegenheiten, um lernrelevante Situationen wahrzunehmen.

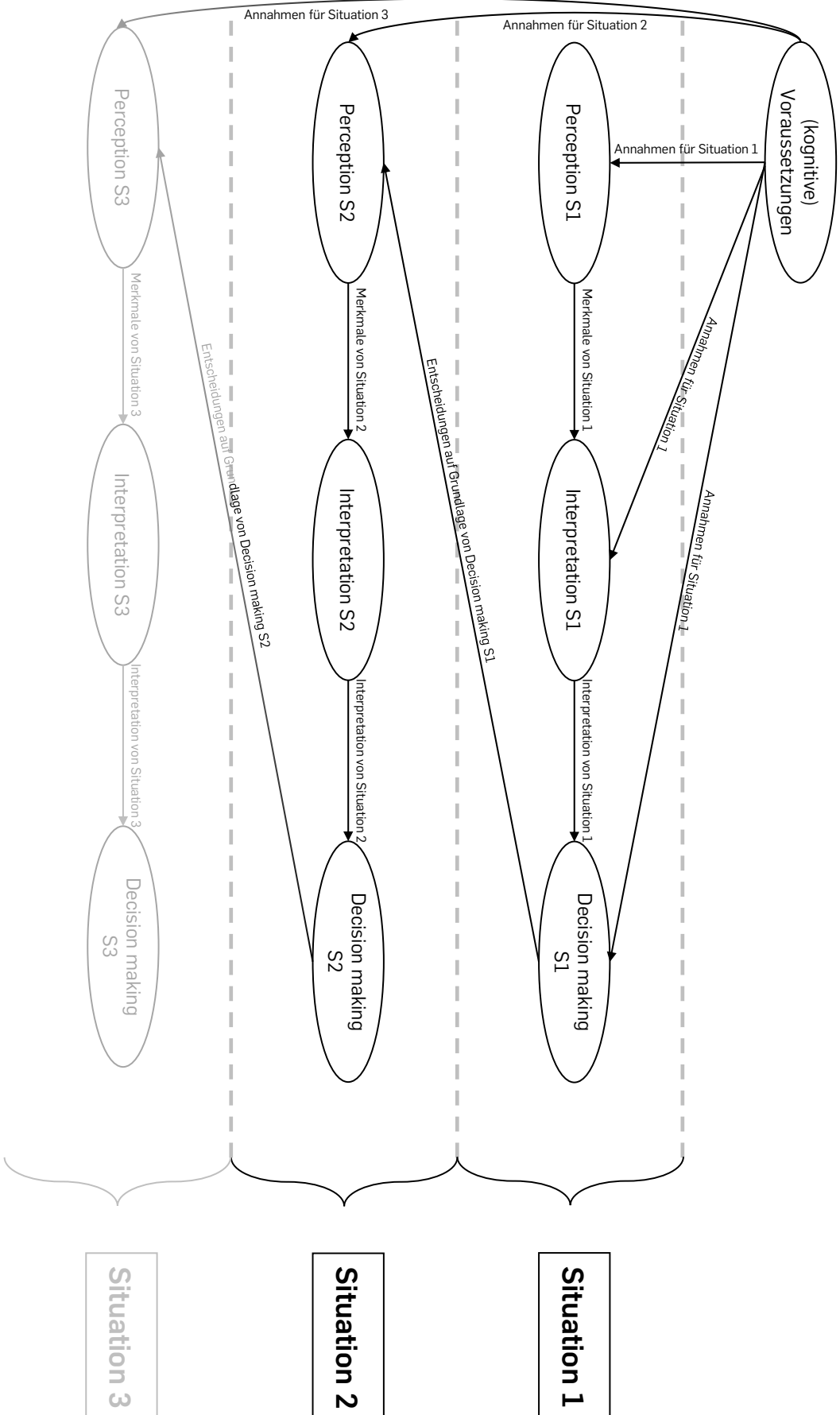


Abbildung 2.7: Eigene Konzeptualisierung der professionellen Unterrichtswahrnehmung.

Der Prozess an sich Insofern ist die professionelle Unterrichtswahrnehmung auf die Dauer einer ganzen Stunde bezogen als ein spiralförmiger Prozess konzipiert, in dem die Ergebnisse der Teilprozesse aus der einen Situation Einfluss auf die Wahrnehmung der folgenden Situation nehmen. Innerhalb einer Situation liegt – unter der Annahme der Dispositionen als einzige Einflussfaktoren auf Perception – allerdings ein hierarchischer Prozess vor, für den erst Perception, dann Interpretation und abschließend Decision Making notwendig ist.

Zusammenfassend wird die professionelle Unterrichtswahrnehmung a priori als ein dreidimensionaler Prozess konzeptualisiert, der innerhalb kürzester Zeit abläuft und kaum planbar ist. Dieser Prozess besteht aus den Teilprozessen (a) Wahrnehmen (Perception), (b) Interpretieren (Interpretation) und (c) Fassen eines Handlungsplans (Decision Making) (in Anlehnung an Sherin et al., 2011). Mit der Betonung der kontextbezogenen und situativen Fähigkeiten von Lehrkräften in jüngeren Veröffentlichungen (vgl. beispielsweise Scheiner, 2016) wie auch in dieser Arbeit, deutet sich eine Verschiebung im Experten-Paradigma an. Bezüglich der Profession von Lehrkräften vollzieht sich eine zunehmende kontextbezogene Betrachtung von Expertise in der bildungswissenschaftlichen Forschung. Im folgenden Kapitel werden die bisher durchlaufenen (und nicht absolut abgeschlossenen) Paradigmen (Persönlichkeits-Paradigma, Prozess-Produkt-Paradigma und Expertenparadigma) kurz beschrieben. In einem zweiten Schritt wird begründet, warum es zu einer Verschiebung im aktuellen Experten-Paradigma kommen und welche Folgen dies nach sich ziehen könnte.

2.3.2 Professionelle Kompetenzen unter Kontextbeachtung

„Because some teacher knowledge may, in fact, be inert, there is growing interest in studying teacher knowledge in context [...]“ (vgl. Kersting, Givvin, Sotelo & Stigler, 2010, S. 173)

Im Laufe der bildungswissenschaftlichen Forschung wurde die Profession der Lehrkraft (zumindest retrospektiv) vor dem Hintergrund verschiedener Paradigmen definiert. Um festzustellen, was eine „professionelle Lehrkraft“ ausmacht, wurden in verschiedenen Paradigmen verschiedene Merkmale von Lehrkräften erfasst. So war beispielsweise das Ziel des sogenannten *Persönlichkeits-Paradigmas* (1900-1960), die Eigenschaften und Tugenden von erfolgreichen Lehrkräften zu identifizieren, indem Unterrichtsstile systematisch beobachtet und spezifische Charaktereigenschaften von Lehrkräften festgestellt wurden (vgl. hier und in folgenden Abschnitten Reichhart, 2018, S. 10-16). Studien dieses Paradigmas zeigen, dass eine Lehrkraft dann erfolgreich unterrichtet, wenn sie weniger Zeit mit Kritik und Vorträgen verbringt, sondern mehr Zeit darauf verwendet, Fragen zu stellen und Ideen und Meinungen von Lernenden weiterzuentwickeln, zu loben und die Lernenden zu ermutigen. Außerdem haben sich die Eigenschaften Freundlichkeit, Sympathie, Tugendhaftigkeit, Enthusiasmus im Gegensatz zu Grausamkeit, Amoralität und depressiven Zügen als förderlich herausgestellt. Die einzigen Forschungsgegenstände aus diesem Paradigma, die bis heute überdauern, sind die Überzeugungen und motivationalen Ori-

entierungen von Lehrkräften. Die Hauptkritik an diesem Paradigma ist, dass die Befunde nicht konsistent sind und die Seite der Lernenden vernachlässigt wird.

Diese Kritik führte letztendlich zu einem Paradigmenwechsel hin zum *Prozess-Produkt-Paradigma* (1960-heute). Ziel dessen war/ist es, eng umgrenzte Fertigkeiten von Lehrkräften (Prozess) und deren Auswirkungen auf die Leistung der Lernenden (Produkt) auszumachen. Dazu wurde das beobachtete Verhalten der Lehrkräfte in spezifische Einheiten zergliedert und genau beobachtend ausgezählt. Dabei wurden als Prozess die Klarheit der Sprache der Lehrpersonen, die Strukturierung des Unterrichts, die Qualität und Quantität der Fragen der Lehrkräfte analysiert und als Produkt der Lernzuwachs der Lernenden verstanden. Die Ergebnisse zeigen, dass eine Veränderung im Prozess (Verhalten der Lehrkraft) das Produkt (Leistungsentwicklung der Lernenden) beeinflusst. Als förderlich stellte sich heraus, wenn Lehrkräfte in ihrem Verhalten klar und adaptiv sind, sowie angemessenes Unterrichtstempo halten und Reflexionen anbieten. Auch hier wurden Eigenschaften wie Enthusiasmus und eine positive Haltung als förderliche Merkmale festgestellt. Die Paradigmen sind demnach nicht ganz frei von Überlappungen. Die Fokussierung auf den Lernzuwachs als einziges Produkt seitens der Lernenden sowie die entkoppelte Betrachtung des Verhalten der Lehrkräfte vom fachlichen Rahmen (Abhängigkeit vom Lernstand, Motivation, Interesse und Intelligenz der Lernenden, Wissen der Lehrkraft) sind die Kritikpunkte an diesem Paradigma.

Seit den 1980er Jahren wird der Frage nach „guten Lehrkräften“ im *Experten-Paradigma* (seit 1985 und bis heute zentral) mit der Erforschung von notwendigem (bereichsspezifischem, prozeduralem und deklarativem) Wissen, Fertigkeiten und subjektiven Theorien nachgegangen. Dabei ist das Paradigma zweigeteilt: In der englischsprachigen Forschung geht man von einer Diskussion um Junglehrkräfte und Studierende sowie erfahrene Lehrkräfte aus. Ein „expert teacher“ ist demnach eine Lehrkraft, die qualifiziert ist. Sie erkennt Muster schneller, handelt einfühlsamer, löst Probleme schneller, beherrscht Automatisierungstechniken und besitzt ein großes Repertoire an Handlungsalternativen. In der deutschsprachigen Forschung hingegen zählt der Lehrberuf an sich als Profession (siehe Unterabschnitt 2.1.1 zu den Anfängen der professionellen Unterrichtswahrnehmung) und somit ist jede Lehrkraft ein_e Experte_Expertin für Lehren und Lernen. Es wird an einem an den Unterricht angepassten Modell erforscht, wie Kompetenzen der Lehrkräfte auf Erfolg seitens der Lernenden wirkt. Es resultieren Wissen, Überzeugungen und motivationale Orientierungen von Lehrkräften, die lernförderlich sind.

In den letzten Jahren zeichnet sich ein neuer Trend bezüglich der Paradigmen ab, in denen Bildung von Lehrkräften beforscht und diskutiert wird. In dieser Arbeit wird es mit „Experten-Paradigma unter Kontextbeachtung“ bezeichnet. Ziel dieser Schwerpunktsetzung innerhalb des Experten-Paradigmas ist es, die Kompetenzen von Lehrkräften noch enger verknüpft mit der Realsituation zu betrachten (vgl. Santagata & Yeh, 2016). Der Fokus liegt immer noch auf den Fähigkeiten guter Lehrkräfte, wird jedoch um eine Perspektive auf *situatives* Lernen der Lernenden erweitert. Durch eine handlungsnähere

Modellierung der Kompetenzen von Lehrkräften sollen größere Effekte auf die tatsächliche Unterrichtshandlung von Lehrkräften erzielt werden. Dieses Ziel wird durch den Einbezug von Kontexten in die Forschung verfolgt (vgl. Kersting et al., 2010, S. 173). Der Einfluss von – mehr oder weniger – realen Kontexten in Unterricht und Schule soll dabei helfen zu verstehen, welchen Zusammenhang es zwischen erlernten Kompetenzen und dem Handeln in der Schule gibt. Einige Studien der letzten Jahre konnten nachweisen, dass nicht „handlungsrelevantes Wissen“ getestet wird (vgl. Cauet, 2016). Um aber Konsequenzen für die Bildung von Lehrkräften ziehen zu können, muss das getestete Wissen handlungsrelevant sein. Wie die Ergebnisse der Forschungsbemühungen dieses (neuen) Paradigmas aussehen, gilt es in den folgenden Jahren zu verfolgen. Forschungsprojekte wie Profile-P, Profile-P+, OBSERVER und TEDS-FU integrieren in die Planung und Umsetzung von Forschung mögliche situative Kontexte, um belastbarere Informationen für die Ausbildung von Lehrkräften und Lernenden zu erlangen, also eine Handlungsvalidierung von Kompetenzen der Lehrkräfte zu sichern.

2.4 Angrenzende Konzepte: Reflexions- und Diagnosekompetenz

„Die *Reflexion* von Unterricht hängt besonders eng mit der Professionellen Unterrichtswahrnehmung zusammen, da die beiden Konstrukte einerseits einen ähnlichen Prozesscharakter aufweisen und sich andererseits gegenseitig bedingen.“ (vgl. Szogs, Kobl, Volmer & Korneck, 2019, S. 318)

Wie in diesem Zitat deutlich wird, scheinen die professionelle Unterrichtswahrnehmung und die Reflexionskompetenz ähnliche Konstrukte zu sein. Außerdem haben die professionelle Unterrichtswahrnehmung und die Diagnosekompetenz deutliche Überlappungsbereiche (vergleiche dazu beispielsweise das Projekt *DiaKom* der Pädagogischen Hochschule Freiburg). Daher soll die theoretische Betrachtung des Konzepts der professionellen Unterrichtswahrnehmung durch den Vergleich mit und mögliche Abgrenzung zu angrenzenden Konzepten wie der Reflexions- und Diagnosekompetenz ergänzt werden.

Szogs et al. (2019) schlagen Merkmale zum Vergleich für die verwandten Konzepte der Reflexionskompetenz und professionellen Unterrichtswahrnehmung vor. Die Merkmale lauten *zeitlicher Bezug, Inhalt, Tiefe, teilnehmende Personen, Modalität, Offenheit* und *Medium*. Die Autorenschaft wendet diese Merkmale auf die Reflexionskompetenz an und nennt anschließend Ähnlichkeiten und Unterschiede zum Konzept der professionellen Unterrichtswahrnehmung (ebd., S. 318, 319). Es folgt nun ein kurzer Überblick über die drei Fähigkeiten, bevor die Unterscheidung durch die von Szogs et al. (2019) vorgeschlagenen Merkmale erfolgt.

Bei der Diagnose, Reflexion und professionellen Unterrichtswahrnehmung handelt es sich jeweils um wissensbasierte Prozesse (von Aufschnaiter et al., 2015; Szogs et al., 2019; Blömeke, Gustafsson & Shavelson, 2015).

Diagnosen umfassen das Feststellen von Merkmalen oder deren Ausprägungen einer Person oder Aufgabe (vgl. von Aufschnaiter et al., 2015, S. 740) und führen über sinnvolle Anschlusshandlungen (Lernende unterstützen (Feedback geben) und fördern, Aufgaben anpassen, u.a.m.) zur Qualitätssicherung des Unterrichts. Ziel ist es, durch die Diagnose Maßnahmen zur Entwicklung eines Merkmals abzuleiten oder eine Selektion vorzunehmen. In der Physikdidaktik können vier Arten der Diagnostik unterschieden werden. Die *Statusdiagnostik* hat die Erfassung der Ausprägung einer Kompetenz, die zu einem bestimmten Zeitpunkt vorliegt, zum Gegenstand (vgl. von Aufschnaiter et al., 2015, S. 744-745). In Abgrenzung dazu fokussiert die *Prozessdiagnostik* Lösungsprozesse, emotionale und motivationale Prozesse während dieser Lösungsprozesse sowie die Nutzung der Lernmaterialien und Wissensressourcen (ebd., S. 745). Als weitere Feingliederung lassen sich die *Veränderungs-* und *Verlaufsdiagnostik* unterscheiden. Die Veränderungsdiagnostik vergleicht zwei Diagnosen zum gleichen Merkmal und zu verschiedenen Zeitpunkten (ebd., S. 745, 746). Es können damit Aussagen über Zuwachs, Stagnation und Abnahme von Leistung der Lernenden getätigt werden. Ziel ist es, die Wirksamkeit von Fördermaßnahmen einzuschätzen. Reiht man jedoch viele Prozessdiagnostiken hintereinander, wird von Verlaufsdiagnostik gesprochen (vgl. von Aufschnaiter et al., 2015, S. 746-747). Die Methodik ist daher dieselbe wie in der Prozessdiagnostik. Dieses Vorgehen wird eingesetzt, um Aussagen über die Dynamik des Lernens sowie über Voraussetzungen für die Ableitung von Fördermaßnahmen zu erhalten (ebd., S. 746).

Reflexion findet auf kognitiver Ebene statt und ist ein kritischer und wissensbasierter internaler Prozess (vgl. Szogs et al., 2019, S. 317), der auf die eigene Professionalisierung abzielt (vgl. von Aufschnaiter, Fraij & Kost, 2019, S. 148) und nicht nur bezüglich unterrichtlichen Handelns, sondern auch bezüglich der „eigenen Wissensbestände, professionellen Fähigkeiten, Einstellungen/Überzeugungen und Bereitschaften u.a. zum Fach, zum Studium oder zum eigenen Weiterlernen“ (ebd., S. 147) relevant ist. Dieser Prozess hat bei Lehrkräften das Ziel, sich mit Geschehnissen im Unterricht auseinanderzusetzen, um die eigenen professionellen Kompetenzen und die resultierende Performanz weiterzuentwickeln (von Aufschnaiter et al., 2019, vgl. Szogs et al., 2019).

Die folgenden Schritte können als – nicht zwingend hierarchisch ablaufende – Phasen eines reflexiven Prozesses identifiziert werden (vgl. von Aufschnaiter et al., 2019, S. 151f.). Der Lehrkraft muss ein Ereignis bewusst werden, sodass dieses anschließend (gedanklich) beschrieben werden kann. Gegenstand der Reflexion kann die eigene Person und das eigene Verhalten oder aber andere Aspekte des Ereignisses sein.

Die beschriebene Beobachtung wird interpretativ gedeutet, sodass erläuternd gelungene und misslungene Anteile identifiziert werden.

Außerdem werden Ursachen für die beobachtete Gegebenheit interpretativ gesucht, wobei es sich häufig um Spekulationen handelt, da sich Zusammenhänge in vielen Fällen nicht prüfen lassen. von Aufschnaiter et al. (2019) merken an, dass die Ursache nicht immer bei der reflektierenden Person liegen muss, die Reflexion dennoch zur eigenen Professionalisierung beitragen kann.

Die aus den vorherigen Phasen der Reflexion abgeleiteten Konsequenzen können entweder in Form von Umsetzung als Handlung beobachtet oder in Form von Nennung dokumentiert werden.

von Aufschnaiter et al. (2019) ordnen Reflexion als Kompetenz in das Kompetenzmodell von Blömeke, Gustafsson & Shavelson (2015) ein. Als für die Reflexion als Kompetenz notwendige Dispositionen nennen von Aufschnaiter et al. (2019) auf Reflexion bezogene Einstellungen, Überzeugungen und die Bereitschaft zur Reflexion sowie reflexionsspezifische Wissensbestände und die dazugehörenden grundlegenden Fähigkeiten (ebd., S. 152). Als situationsspezifische Fähigkeit zählt demnach die Reflexion als Prozess (siehe oben), welche aber auch performante Anteile habe (ebd.). Zusätzlich zählt laut der Autorenschaft zur Performanz das Verhalten, von dem auf Reflexion geschlossen werden kann. Reflexion führt zur eigenen Professionalisierung durch wiederholendes Durchlaufen der Reflexion und anschließenden Handlungen, sodass die professionellen Kompetenzen weiterentwickelt werden können (Szogs et al., 2019).

Professionelle Unterrichtswahrnehmung konstituiert sich aus den in Unterabschnitt 2.1.2 und Abschnitt 2.2 beschriebenen Teilprozessen (a) Erkennen, (b) Interpretieren und (c) Entscheiden. Damit sind die Prozesse der Reflexion und der professionellen Unterrichtswahrnehmung bis zu einem gewissen Grad ähnlich, allerdings legt letztere einen stärkeren Fokus auf Tätigkeiten der Lernenden, die orchestriert und unterstützt werden müssen, statt auf die eigene Person.

In Tabelle 2.1 werden die drei Konzepte vor allem hinsichtlich ihres zeitlichen Bezugs verglichen (vgl. Merkmale nach Szogs et al., 2019). Die Diagnose, unter Umständen aber auch die professionelle Unterrichtswahrnehmung, beginnt mit der Planung diagnostischer Tätigkeit vor dem Unterricht. Dies kann durch die Entwicklung schriftlicher Aufgaben oder Beobachtungen mündlicher Äußerungen geplant werden. Dabei kann unterschieden werden, dass Status- und Veränderungsdiagnostik einer expliziten Planung bedürfen, während die Prozess- und Verlaufsdiagnostik auch spontan ohne Vorabplanung ablaufen kann. Im Unterricht beginnen die Prozesse der Diagnose und professionellen Unterrichtswahrnehmung in einer Situation, die zum einen als Gelegenheit für eine geplante Informationsbeschaffung dient (Status- und Veränderungsdiagnostik) und zum anderen durch Wahrnehmung einer Lerngelegenheit Anlass zu Prozess- und Verlaufsdiagnostik sowie Teilprozesse der professionellen Unterrichtswahrnehmung gibt. In der Literatur über die Reflexion von Unterricht wird die Wahrnehmung oder Informationsbeschaffung nicht als Teil der reflexiven Prozesse aufgenommen. Reflexion beginnt demnach frühestens nach einer Situation während des Unterrichts („reflection in action“) mit den vier Prozessen Bewusstwerden, Beschreiben, Bewerten und Generieren von Handlungsoptionen. Bezüglich der Diagnose und professionellen Unterrichtswahrnehmung vollziehen sich ähnliche Prozesse (Interpretation und Generieren von Handlungsoptionen). Allerdings beinhalten Diagnose und Reflexion nicht zwingend eine sofortige Interpretation und Generierung von Handlungsoptionen. Diese Teilprozesse können auch erst nachbereitend im Anschluss an den Unterricht stattfinden. Im Sinne der Diagnose würde eine Analyse der beschaffenen Daten durchgeführt werden.

Tabelle 2.1: Angrenzende Konzepte zur professionellen Unterrichtswahrnehmung: Reflexion und Diagnose.

Konzept	Ziel	Vor dem Unterricht		Während des Unterrichts				Nach dem Unterricht			
				In der Situation	Nach der Situation						
Diagnose	Verbesserung der ... - ... Schülerleistungen - ... Aufgaben	Statusdiagnostik	Planung	Informationsbeschaffung	Interpretation (Analyse der Informationen)	Generierung von Handlungsoptionen		Interpretation (Analyse der Informationen)	Generierung von Handlungsoptionen		
		Veränderungsdiagnostik									
		Prozessdiagnostik	(Planung)								
		Verlaufsdiagnostik									
Professionelle Unterrichtswahrnehmung	Lerngelegenheiten für Schüler_innen identifizieren und nutzbar machen	(möglicherweise Schwerpunktplanung)		Wahrnehmung	Interpretation	Generierung von Handlungsoptionen					
Reflexion	Verbesserung der eigenen Kompetenzen von Lehrkräften			Reflection in action				Reflection on action			
			Bewusst werden	Beschreiben	Bewerten	Generierung von Handlungsoptionen	Bewusst werden	Beschreiben	Bewerten	Generierung von Handlungsoptionen	

Für die Reflexion einer Situation mit etwas zeitlichem Abstand zu dieser hat Schön (1983) den Begriff „reflection on action“ geprägt. Es zeigt sich, dass die professionelle Unterrichtswahrnehmung sowohl mit der Reflexion als auch mit der Diagnose Überschneidungen aufweist. Die Überschneidung mit der Reflexion betrifft reflexive Prozesse, die zu einer Regulierung des eigenen Handelns zur Förderung von Lernen der Lernenden führen. Nach den bisherigen Konzeptualisierungen muss die professionelle Unterrichtswahrnehmung als Voraussetzung für die Reflexion verstanden werden. Diagnose und professionelle Unterrichtswahrnehmung überschneiden sich bezüglich der Prozesse während des Unterrichts: Man kann zusammenfassen, dass ungeplantes Diagnostizieren (Prozess- und Verlaufsdiagnostik) und professionelle Unterrichtswahrnehmung eine hohe Ähnlichkeit aufweisen, da die Wahrnehmung in einer Situation stattfindet, die Interpretation einer Situation direkt im Anschluss während des Unterrichts vollzogen wird und eine Generierung von Handlungsoptionen nach sich zieht. In diesem Fall kann daher nicht klar zwischen den Prozessen somit zwischen den Fähigkeiten unterschieden werden. Da die Reflexion eigene Tätigkeiten und Fähigkeiten adressiert und die Diagnose Lernenden und Aufgaben fokussiert, sind diese Prozesse disjunkt.

2.5 Zusammenfassung

Die professionelle Unterrichtswahrnehmung wurde durch eine Adaption des professionellen Blicks nach Goodwin (1994) in die Diskussion der Bildung von Lehrpersonen eingebracht. Sie wird in der Literatur als ein zwei- beziehungsweise dreidimensionaler Prozess konzipiert. Diese drei Dimensionen werden vom PID-Modell aufgegriffen, welches die professionelle Unterrichtswahrnehmung als mediierendes Element zwischen Wissen und Überzeugungen (Dispositionen) sowie Unterrichtspraxis (Performanz) definiert. Diese drei Dimensionen sind die Teilprozesse Perception (Erkennen), Interpretation (Beschreiben, Erklären, Vorhersagen) und Decision Making. Diese Arbeit folgt dem PID-Modell, integriert allerdings den Kontext und psychologische Überlegungen zu Wahrnehmung und Kognition. Scheiner (2016) und Santagata & Yeh (2016) machen deutlich, dass der Einbezug von Kontext und Situation dringend nötig ist, um die Prozesse der professionellen Unterrichtswahrnehmung in der Forschung und Lehre möglichst valide und realitätsgetreu abbilden zu können. Die professionelle Unterrichtswahrnehmung ist notwendige Voraussetzung für die Konzepte der Diagnose und Reflexion, es gibt allerdings Überschneidungen und Ähnlichkeiten der Konzepte.

Im nächsten Kapitel wird der aktuelle Stand der empirischen Forschung über die professionelle Unterrichtswahrnehmung vorgestellt. Diese Informationen legen zusammen mit den theoretischen Konzeptualisierungen die Grundlage für die Operationalisierung der professionellen Unterrichtswahrnehmung in der vorliegenden Arbeit.

3 Stand der empirischen Forschung

In diesem Kapitel werden zentrale Ergebnisse der empirischen Forschung mittels Videos bezüglich professioneller Unterrichtswahrnehmung in Bezug auf Mathematik- und Naturwissenschaftsunterricht (auch mit allgemeinpädagogischen Anteilen) berichtet, um anschließend Forschungsdesiderate abzuleiten und das Vorhaben der vorliegende Arbeit daraus abzuleiten.

Gemäß der Entwicklung der Debatte um die professionelle Unterrichtswahrnehmung gibt es zahlreiche Studien, die einen Experten-Novizen-Vergleich anstellen (vgl. beispielsweise Berliner, 2001; Carter et al., 1988; Kersting, 2008).

Ein Großteil der weiteren aktuelleren Studien lassen sich unterteilen in Studien zu

- der Entwicklung der professionellen Unterrichtswahrnehmung von Lehramtsstudierenden (vgl. beispielsweise Star & Strickland, 2008; Schäfer & Seidel, 2015; Sunder, Todorova & Möller, 2016),
- der Entwicklung der professionellen Unterrichtswahrnehmung von Lehrkräften (vgl. beispielsweise van Es & Sherin, 2006; Santagata, Zannoni & Stigler, 2007; Roth et al., 2011),
- Zusammenhängen der professionellen Unterrichtswahrnehmung mit anderen Konzepten wie Fachwissen und fachdidaktischem Wissen (vgl. beispielsweise Kersting et al., 2010; Stürmer, Könings & Seidel, 2015; Meschede et al., 2017).

Im nächsten Abschnitt werden zentrale Ergebnisse dieser Studien exemplarisch¹⁶, zusammenfassend berichtet und deren Grenzen aufgezeigt. Da im Zentrum dieser Arbeit die Entwicklung eines validen und reliablen Messinstruments steht, wird im zweiten Abschnitt detaillierter auf Studien eingegangen, in denen Forschungsgruppen (standardisierte) Messverfahren zur Erfassung der professionellen Unterrichtswahrnehmung entwickelt haben (vgl. beispielsweise Kaiser et al., 2017; Seidel & Prenzel, 2007; Wolters, 2015).

¹⁶Für eine detaillierte Übersicht über Studien zur professionellen Unterrichtswahrnehmung bis einschließlich 2016 siehe Stahnke et al. (2016).

3.1 Studien zur professionellen Unterrichtswahrnehmung

Experten-Novizen-Vergleich Studien, in denen die professionelle Unterrichtswahrnehmung von erfahrenen Lehrkräften, Junglehrkräften und Studierenden und Nicht-Lehrkräften (aber mit Kenntnissen im Fach) verglichen werden, kommen zu folgenden Ergebnissen (vgl. Carter et al., 1988; Sabers, Cushing & Berliner, 1991; Seidel & Prenzel, 2007; Oser, Heinzer & Salzmann, 2010): Erfahrene Lehrkräfte ...

- bauen auf ihren Unterrichtserfahrungen und ihrem Wissen auf,
- lenken ihre Aufmerksamkeit auf eher atypische Ereignisse, die häufig eine hohe Lernrelevanz im Unterricht haben,
- interpretieren Ereignisse auf Grundlage von Erfahrung und lernförderlicher Annahmen,
- fokussieren ihre Aufmerksamkeit auf Aktivitäten der Lernenden,
- nennen häufiger Alternativen oder Handlungspläne als Studierende oder Nicht-Lehrkräfte,
- sind im Erklären der Ereignisse besser als Studierende oder Nicht-Lehrkräfte.

Berliner (2001) hinterfragt die ursprüngliche Bezeichnung von erfahrenen Lehrkräften als „Expert_innen“ kritisch und folgert, dass Expertise keine Eigenschaft einer Person, sondern vielmehr die Eigenschaft von Interaktionen zwischen Personen und Kontext ist. Im Rahmen dieser Untersuchung stellt Berliner (2001) verschiedene Aussagen über Experten-Lehrkräfte zusammen, von denen sich einige auf situative Fähigkeiten von Lehrkräften beziehen. Experten-Lehrkräfte sind (a) zwar opportunistischer dadurch aber flexibler, (b) bei Problemlösungen sensibler für die Aufgabenanforderungen und die umgebende soziale Situation, (c) verfügen über schnellere und genauere Mustererkennung, (d) nehmen in der Domäne, in der sie erfahrener sind, mehr wahr und (e) bieten reichhaltigere, persönlichere Informationsquellen zur Lösung von Problemen an als Novizen-Lehrkräfte.

Dem gegenübergestellt haben Junglehrkräfte und Studierende wenig Erfahrung. Ihr Wissen ist in Teilen nicht vernetzt, weshalb ihnen das Beschreiben von Klassenraumsituationen mit passendem Vokabular zwar leicht fällt, ihre Interpretationen aber häufig oberflächlich oder unpassend sind (Carter et al., 1988; Seidel & Prenzel, 2007; Oser et al., 2010).

In der Studie von Stürmer, Seidel, Müller, Häusler & Cortina (2017) wurde die Blickrichtung und -dauer von Studierenden aufgezeichnet mithilfe mobiler Eye-Tracking Technologie untersucht, während sie in standardisierten Mikroteaching-Sequenzen oder im realen Mathematik- oder Naturwissenschaftsunterricht Lernprozesse unterstützen. Es konnte auf einer kleinen Stichprobe ($N = 7$) basierend festgestellt werden, dass Lehramtsstudierende (Mathematik oder Naturwissenschaften) ihre Aufmerksamkeit im Mittel am häufigsten auf die Lernenden richten, sich dabei aber auf einige wenige Lernende beschränken (ebd.). Unter den Testpersonen variiert allerdings die Person, auf die der Fokus liegt. Die Ergebnisse ließen sich für vier der Testpersonen auch im realen Unterricht reproduzieren (ebd.)

Über Nicht-Lehrkräfte wird berichtet, dass sie ein hohes Fachwissen vorzeigen können, mit dessen Hilfe sie das Lernen der Lernenden fokussieren, dabei allerdings von der Unübersichtlichkeit im Klassenraum überfordert sind (Carter et al., 1988; Oser et al., 2010). In den oben zitierten Studien werden erfahrenen Lehrkräften im Wesentlichen durch zwei Kriterien charakterisiert: Sie haben erstens im Schnitt mehr als fünf Jahre Unterrichtserfahrung und werden zweitens von Schulleitungen oder Inspizierenden vorgeschlagen. Diese beiden Kriterien sind eine gute Annäherung an einen wahrhaften Vorsprung an Erfahrung. Sie sollten aber durch weitere Kriterien ergänzt werden, weil es denkbar ist, dass Lehrkräfte viele Jahre Unterrichtserfahrung und gute Kooperationen mit dem Kollegium aufweisen, aber ihre Erfahrungen unreflektiert bleiben und sich daher nicht von den Erfahrungen von Junglehrkräften und Studierenden unterscheiden.

Studien zur Entwicklung der professionellen Unterrichtswahrnehmung von Lehramtsstudierenden Wie bereits erläutert, ist die professionelle Unterrichtswahrnehmung von Studierenden in ihrer Qualität noch nicht mit der von erfahrenen Lehrkräften vergleichbar. Es gibt diverse Studien, die erforschen, ob Interventionen (wie beispielsweise Seminare mit Videos und/oder Reflexion, Unterrichtsskripten¹⁷, Unterrichtsbeobachtungen, Peer-Teaching, Learning to Learn oder Praktika) diese situative Fähigkeit fördern können. Die Ergebnisse von Tests zur professionellen Unterrichtswahrnehmung zeigen, dass Interventionen dazu führen, dass Lehramtsstudierende ...

- ... zunehmend Lernrelevantes wahrnehmen,
- ... häufiger Bezug zu ihrem professionellen Wissen herstellen, um erkannte Situationen zu interpretieren (Schäfer & Seidel, 2015),
- ... Handeln von Lehrkräften und Lernenden signifikant häufiger beschreiben (Gold, Förster & Holodynski, 2013; Santagata & Guarino, 2011),
- ... signifikant mehr interpretierende Kommentare zum Lernen von Lernenden und Lehrstrategien tätigen (Gold et al., 2013; Santagata & Guarino, 2011),
- ... sich vor allem im Vorhersagen über folgende Unterrichtsgeschehen verbessern (Kramer, König, Kaiser, Ligtvoet & Blömeke, 2017),
- ... signifikant mehr Vorschläge für alternative Lehrstrategien geben (Gold et al., 2013; Santagata & Guarino, 2011) und
- ... Aspekte wie Classroom Management, Inhalt, Aufgaben und Kommunikation vermehrt wahrnehmen (Kramer et al., 2017; Star & Strickland, 2008).

Insgesamt lässt sich erkennen, dass die professionelle Unterrichtswahrnehmung zumindest bezüglich einiger Aspekte von Unterricht gefördert werden kann. Das gilt mittlerweile nicht mehr nur für allgemeinpädagogische Themen (wie die Klassenführung bei Gold et al., 2013), oder mathematikspezifische Inhalte (wie bei Santagata & Guarino, 2011),

¹⁷Interventionen mit Unterrichtsskripten bedeutet, dass Studierende mit schriftlichen Unterrichtsbeschreibungen arbeiten, in denen Situationen lernrelevant sind und daher professioneller Unterrichtswahrnehmung bedürfen.

sondern auch für fachspezifische Themen des Sachunterrichts (Sunder et al., 2016). Die Interventionen mit Unterrichtsvideos sind zwar motivierend, zeigen als Interventionsmerkmal aber im Vergleich zu Unterrichtstranskripten keinen signifikant größeren Effekt auf die professionelle Unterrichtswahrnehmung (Kramer et al., 2017; Sunder et al., 2016). Mehrere Studien konnten auch berichten, dass sich die Fähigkeit, Unterricht professionell wahrzunehmen, nur bezüglich der Themen der Intervention verbessert (vgl. Gold et al., 2013; Star & Strickland, 2008; Sunder et al., 2016). Wie weiter unten ebenfalls berichtet wird, gibt dies Hinweise darauf, dass die professionelle Unterrichtswahrnehmung zweidimensional ist und in generische sowie themenspezifische Anteile zerfällt (Steffensky, Gold, Holodynski & Möller, 2015). Denkbar ist auch, dass mit ausreichender Praxis und Erfahrung eine generische professionelle Unterrichtswahrnehmung erreicht werden kann. Allerdings handelt es sich dabei um Vermutungen, die man in zukünftigen Studien beispielsweise über Zusammenhangsanalysen zu Fachwissenstests untersuchen könnte. Wenn ein positiver Zusammenhang bestände, wäre es notwendig herauszufinden, welche Themen des Physikunterrichts Schwierigkeiten bezüglich der professionellen Unterrichtswahrnehmung bergen.

Weitere Limitation einiger Studien sind die geringen Stichprobenzahlen, welche zugleich nicht repräsentativ für alle (Fach-)Lehrkräfte sind (vgl. Santagata & Guarino, 2011; Star & Strickland, 2008; Sunder et al., 2016). Möglicherweise ist die Arbeit mit unbearbeiteten Unterrichtsvideos, deren Komplexität durch Schwerpunktsetzung und Reduktion an simultan ablaufenden Geschehnissen nicht reduziert wurde, für Bachelorstudierende auf Grund der Komplexität der Videos eher überfordernd als förderlich (Kramer et al., 2017). Schäfer & Seidel (2015) schlagen vor, eine größere Vielfalt an Videos zu verwenden, um ein größeres Spektrum an Fähigkeitsausprägungen der professionellen Unterrichtswahrnehmung fördern zu können. Es stellt sich dabei die Frage, ob die themenspezifischen professionellen Unterrichtswahrnehmungen alle gleich und gleich gut gefördert werden können. Abschließend müssten Langzeiteffekte durch Follow-Up-Tests erforscht, sowie die Anwendung der geförderten Wahrnehmung in Form von Handlungen im Klassenraum untersucht werden (Schäfer & Seidel, 2015).

Studien zur Entwicklung der professionellen Unterrichtswahrnehmung von Lehrkräften Die professionelle Unterrichtswahrnehmung von Lehrkräften und ihre Entwicklung ist weniger gut erforscht. Bisher wurden zwei Formate der Intervention für Lehrkräfte implementiert: Die Videoclubs von Sherin und van Es und die Seminare von STeLLA¹⁸ von Roth und anderen (2011) dienten bisher als einzige Lerngelegenheit neben dem eigentlichen Unterrichten. Dennoch gibt es gerade durch die sogenannten Video Clubs von Sherin und van Es Aussagen über die Entwicklung dieser situativen Fähigkeit von Lehrenden. In den Video-Clubs sieht eine Gruppe von Lehrkräften Unterrichtsvideos von Unterricht der Anwesenden an und diskutiert über die Situationen, die sie wahrnehmen. Die Lehrkräfte richten anders als vor der Teilnahme – nun angeregt durch den Diskurs in den Video Clubs – ihre Aufmerksamkeit mehr auf die Lernenden, das mathematische Denken, sind

¹⁸Für weitere Informationen zu diesem Projekt siehe Roth et al. (2011).

dabei eher interpretierend und blicken auf Spezifika mehr als auf das Allgemeine (van Es & Sherin, 2006). Insgesamt konnten die Lehrkräfte in diesen Studien ihre Fähigkeit, Unterricht professionell wahrzunehmen, verbessern (Roth et al., 2011; Sherin & van Es, 2008). Roth et al. (2011) haben sogar gezeigt, dass in den Lerngruppen der untersuchten Lehrkräfte auch die Leistung der Lernenden bezüglich des mathematischen Wissens zugenommen hat.

Eine deutliche Limitation der hier vorgestellten Studien ist, dass sie allesamt nur die professionelle Unterrichtswahrnehmung von Mathematikunterricht untersuchen und fördern (auch wenn van Es & Sherin (2006) in ihren Clubs auch Nicht-Mathematiklehrkräfte integrierten). Zudem sind auch hier die Stichprobenzahlen gering (maximum $N = 48$), sodass verallgemeinerbare Aussagen zur Förderbarkeit der professionellen Unterrichtswahrnehmung von Lehrkräften nicht ohne Weiteres möglich sind.

Studien zum Zusammenhang der professionellen Unterrichtswahrnehmung mit anderen Konzepten

Häufig wurde in Verbindung zur Entwicklung der professionellen Unterrichtswahrnehmung untersucht, mit welchen anderen Konzepten diese Entwicklung zusammenhängt. Es hat sich ergeben, dass das (mathematische) fachdidaktische Wissen (Kersting, 2008; Meschede et al., 2017), das Interesse, die Anzahl belegter Seminare (Stürmer, 2011), die Lerngelegenheiten (Stürmer et al., 2015) und Beliefs (Meschede et al., 2017) mit der professionellen Unterrichtswahrnehmung zusammenhängen. Das (mathematische) Fachwissen steht in einem Zusammenhang mit der Teilfähigkeit zur Formulierung von Alternativen (Kersting, 2008). Kaiser et al. (2017) konnten außerdem zeigen, dass die mathematikspezifische professionelle Unterrichtswahrnehmung stark mit der pädagogischen professionellen Unterrichtswahrnehmung und dem allgemeinpädagogischen Wissen korreliert (ebd., S. 174). Der Zusammenhang von der mathematikspezifischen professionellen Unterrichtswahrnehmung zum mathematikdidaktischen Wissen sowie zwischen der pädagogischen professionellen Unterrichtswahrnehmung und dem allgemeinpädagogischen Wissen ist eher gering (ebd., S. 175). Entgegen den Erwartungen aus Tests zu anderen Kompetenzen gibt es keinen signifikanten Zusammenhang der professionellen Unterrichtswahrnehmung mit Vorerfahrung aus Praktika (Stürmer, 2011), der Abiturnote (Stürmer et al., 2015) oder dem Fachsemester (Gold et al., 2013). Transmissive Überzeugungen korrelieren signifikant negativ mit der professionellen Unterrichtswahrnehmung (Meschede et al., 2017).

Seidel & Stürmer (2014) konnten durch eine IRT-Skalierung nachweisen, dass sich die professionelle Unterrichtswahrnehmung reliabler als ein eindimensionales Modell abbilden lässt, das drei Subdimensionen (beschreiben, erklären, vorhersagen) umfasst, als ein dreidimensionales Modell, in dem die Dimensionen nicht Teil einer Fähigkeit sind.

Steffensky et al. (2015) haben außerdem die professionelle Wahrnehmung betreffend der Klassenführung und inhaltlichen Lernbegleitung (kognitive Aktivierung und Strukturierung des Inhalts) getrennt modelliert. Diese beiden Fähigkeiten sind zwar positiv assoziiert, aber distinkt. Dies liefert einen ersten Hinweis darauf, dass die professionelle

Unterrichtswahrnehmung nicht nur domänenspezifisch, sondern sogar abhängig von verschiedenen fokussierten Unterrichtsaspekten ist.

Der Großteil der Studien zur Untersuchung der professionellen Unterrichtswahrnehmung und der Zusammenhänge zu anderen Konzepten erfolgt bisher überwiegend in Laborsituationen außerhalb der Klassenräume mithilfe von Unterrichtsvideos. Um die Fähigkeit Unterricht professionell wahrnehmen zu können und somit den Unterricht zu verbessern, müssten Untersuchungen und Förderung in Zukunft auch im eigentlichen Unterricht stattfinden (vgl. Forderungen von Kaiser et al., 2017; Kersting, 2008). Es muss außerdem der Zusammenhang von professioneller Unterrichtswahrnehmung und Lernerfolg der Lernenden untersucht werden. In einem ersten Schritt wurden konkrete Fähigkeiten und ihr Zusammenwirken untersucht, jetzt ist es notwendig zu untersuchen, ob diese Fähigkeiten auch praktischen Effekt auf die Leistungen der Lernenden haben (ähnlich wie bei Roth et al., 2011). Für die Mathematikdidaktik gibt es das Projekt TEDS-Instruct, welches die Beziehungen zwischen Kompetenzen der Lehrkräfte und Leistungen der Lernenden analysiert (Schlesinger, Jentsch, Kaiser, König & Blömeke, 2018). Es ist für die Beurteilung in laufendem Unterricht durch Beobachtende zur Feststellung der professionellen Unterrichtswahrnehmung gedacht. Das bedeutet, der Unterricht wird nicht durch ein Medium wie beispielsweise Videos mediiert, sondern direkt beobachtet. Das Instrument enthält sowohl Items zu generischen als auch zu themenspezifischen Ereignissen im (Mathematik-) Unterricht, welche auf einer 4-stufigen Likertskala bewertet werden können. Das Instrument beurteilt die äußeren Prozesse im Unterricht (zum Beispiel die effektive Nutzung von Lernzeit oder den Gebrauch von verschiedenen Repräsentationsformen). Erst wenn man die Güte der professionellen Unterrichtswahrnehmung für die in-vivo beobachtete Lehrkraft kennt, können in Ergänzung mit den Beobachtungen und Bewertungen des Instruments aus TEDS-Instruct Aussagen über den Effekt der Fähigkeit auf Leistungen der Lernenden getätigt werden. Die Beziehung zwischen Kompetenzen der Lehrkräfte und Leistungen der Lernenden analysieren zu können, bleibt (gerade auch für den Physikunterricht) somit weiterhin eine Herausforderung.

3.2 Studien zur Instrumententwicklung zur Erfassung der professionellen Unterrichtswahrnehmung

Viele der oben genannten Studien basieren auf kleinen Stichprobenzahlen und erfassen die professionelle Unterrichtswahrnehmung qualitativ. Um die empirischen Ergebnisse dieser Studien zu bestätigen, sind standardisierte Testinstrumente wünschenswert. Ergebnisse, die Aussagen über größere Fallzahlen erlauben, helfen zu verstehen, wie die professionelle Unterrichtswahrnehmung konzeptualisiert und gefördert werden kann. Es wurden – vor allem für den Mathematikunterricht – bereits Messinstrumente entwickelt, die dem Anspruch gerecht werden, standardisiert einsetzbar und sowohl reliabel als auch valide zu sein. Hier werden diejenigen Instrumente vorgestellt, die auf Videos basieren und Mathematik- oder Naturwissenschaftsunterricht betreffen. Folgende Kriterien können die

Messinstrumente zur Erfassung der professionellen Unterrichtswahrnehmung charakterisieren (siehe Tabelle 3.1 und Tabelle 3.2):

- Das Fach: Fokussiert das Instrument auf ein bestimmtes Schulfach. Wenn ja, auf welches?
- Die Anzahl und Länge der Videoclips pro Person: Wie viele Videoclips von welcher Länge sehen die Testpersonen?
- Die Art der Videoclips: Sind die Videoclips aus realem Unterricht oder auf Basis von Drehbüchern entstanden?
- Der Gegenstand der Videoclips: Welche Klassenstufe und welches Thema wird in den Videoclips gezeigt?
- Der Umgang mit den Videoclips: Wie sind die Videoclips in das Instrument eingebettet, wie oft darf ein Clip angesehen, gestoppt werden?
- Die Art der Items: Sind die Items geschlossen oder offen?
- Die erfasste(n) Fähigkeit(en): Welche der Teilfähigkeiten der professionellen Unterrichtswahrnehmung werden erfasst?
- Die adressierten Unterrichtsaspekte: Welche Unterrichtsaspekte sollen wahrgenommen werden (allgemeinpädagogisch oder fachdidaktisch, generisch oder themenspezifisch)?

Das wohl erste Instrument dieser Art wurde von Seidel & Prenzel (2007) im Rahmen des Projektes „Lernen aus Unterrichtsvideos“ entwickelt. Das Instrument umfasst acht zweiminütige Videoclips pro Person, die realen Unterricht abbilden. Die Videoclips stammen aus der IPN-Videostudie und betreffen Physikunterricht. Die Clips sind so angeordnet, dass sie den Ablauf einer ganzen Unterrichtsstunde skizzieren. Die Testpersonen sehen sich die Clips an und beantworten dazu geschlossene Rating-Items (4-stufige Likertskala). Ihre Antworten sollen Aufschluss über die Analysekompetenz (beschreiben, erklären, bewerten) bezüglich des Experimentierens, der Fehlerkultur, Lernbegleitung und Zielorientierung der (angehenden) Lehrkräfte geben. Da die Analysekompetenz eine leicht andere Fähigkeit beschreibt als die professionelle Unterrichtswahrnehmung, ist die Art und Weise der Erfassung in Form mehrstufiger Likertskalen angemessen. Um die professionelle Unterrichtswahrnehmung valide messen zu können, wäre vermutlich ein Design angebracht, in dem Testpersonen selbstständig Fokusse legen. Hier werden demnach sowohl generische (Fehlerkultur, Lernbegleitung, Zielorientierung) als auch fachspezifische (Experimentieren) Aspekte betrachtet.

Zur gleichen Konzeptualisierung wie Seidel & Prenzel (2007) entwickelte Kersting (2008) ein Messinstrument, um die professionelle Unterrichtswahrnehmung im Mathematikunterricht zu untersuchen. Die Testpersonen sehen sich dazu zehn ein- bis drei-minütige Videos realen Unterrichts an und diskutieren den Umgang der Agierenden untereinander schriftlich in einem offenen Antwortformat. Die gezeigte 8. Klasse wurde im Rahmen der TIMMS 1999 Videostudie videographiert. Die Analysekompetenz (beschreiben, erklären,

bewerten) wurde durch die Fragestellungen und Videoausschnitte des Unterrichts auf die generischen Aspekte Unterstützung durch Lehrkräfte und Fehler der Lernenden gelenkt. Seidel, Blomberg & Stürmer (2010) und Stürmer (2011) entwickelten ein Instrument (Observer¹⁹), welches nicht nur den Mathematik- und Physikunterricht betrifft, sondern auch die Fächer Englisch, Französisch und Geschichte. Der Fokus für die professionelle Unterrichtswahrnehmung (beschreiben, erklären, vorhersagen) liegt daher auch auf überfachlichen, lernförderlichen Aspekten wie Lernatmosphäre, Lernbegleitung und Zielorientierung. Die sechs zwei- bis vierminütigen Videoclips zeigen realen Unterricht mit Phasen der Erarbeitung und Übung von 8. und 9. Klassen des jeweiligen Faches. Die Testpersonen durchlaufen die Clips jeweils zweimal: Im ersten Durchgang sehen sie sich die Videoclips an und schätzen nach jedem Clip ein, welchen Unterrichtsaspekt (Lernatmosphäre, Lernbegleitung und Zielorientierung) der Clip adressiert. Im zweiten Durchgang sehen die Testpersonen sich die Clips erneut an und beantworten danach geschlossene Ratingitems (4-stufige Likertskala), die auf ihre professionelle Unterrichtswahrnehmung abzielen. Dieses Vorgehen scheint relativ fern von der eigentlichen unterrichtlichen Handlung von Lehrkräften zu sein: Die Testpersonen sehen die Situationen zum einen zweimal hintereinander, was im Unterricht nicht möglich ist und beurteilen das Geschehen zum anderen erst nachdem sie den ganze Clip gesehen haben. Das führt dazu, dass sie nicht selbst ihre Aufmerksamkeit auf die Situation lenken müssen (Selective Attention) und dass sie möglicherweise eine Unterrichtswahrnehmung berichten, die durch das Sehen des ganzen Clips eine andere ist als sie es im Laufe eines realen Unterrichts wäre. Einige Situationen erscheinen im Lichte der ganzen Unterrichtsstunde oder zumindest folgender Situationen gegebenenfalls mehr oder weniger lernrelevant. Bei derartigen Messvorgängen wird die Situativität des Konstrukts der professionellen Unterrichtswahrnehmung in den Hintergrund gerückt.

Für den Grundschulunterricht der Fächer Deutsch, Englisch und Mathematik entwickelten Gold et al. (2013) einen standardisierten Videotest. Die sieben Videoclips sind ein bis drei Minuten lang und zeigen realen Grundschulunterricht mit sowohl gelingenden als auch weniger gelingenden Ereignissen. In Kleingruppen sehen die Testpersonen sich die Clips auf einer Leinwand an und beantworten anschließend in Einzelarbeit geschlossene, dichotome Items. Mit ihrem Antwortverhalten sollen Rückschlüsse auf die professionelle Unterrichtswahrnehmung (Prozessmerkmale identifizieren, theoriegeleitet interpretieren) bezüglich der generischen Unterrichtsmerkmale Allgegenwärtigkeit, Aufrechterhaltung des Gruppenfokus und Strukturierung des Unterrichtsverlaufs gezogen werden. Auch hier gilt eine ähnliche Kritik wie oben. Die Testpersonen urteilen über einzelne Situationen wieder nur mit dem Wissen über den gesamten Clip. Außerdem können nur wenig valide Rückschlüsse auf die professionelle Unterrichtswahrnehmung der Testpersonen aus dem Ankreuzen eines dichotomen Items gezogen werden. Die Problematik liegt darin, dass die Testpersonen ihre Aufmerksamkeit nicht selbst steuern und die Interpretation der Situationen durch das dichotome Item stark eingegrenzt wird.

Der Test von Wolters (2015) ergänzt den obigen Test durch den Fokus auf naturwissen-

¹⁹Es gibt eine erweiterte Version des Observers: Observer Extended Research Tool (Stürmer & Seidel, 2015).

schaftlichen Sachunterricht in der Grundschule. Die Testpersonen sehen sich neun drei- bis vierminütige Videoclips²⁰ an und beantworten im Anschluss geschlossene Ratingitems (4-stufige Likertskala). Die Items zielen dabei auf das Bemerken und Interpretieren von kognitiv aktivierenden Situationen im Unterricht ab. Die Testpersonen kennen die einzelnen Situationen wieder im Lichte des gesamten Clips. Das verändert möglicherweise den professionellen Blick auf das Geschehen und grenzt in jedem Fall die Möglichkeit der Aufmerksamkeitslenkung ein. In den letzten Jahren wurden vermehrt Videos auf Basis von Drehbüchern eingesetzt.

Wiedmann (2015) entwickelte ein Messinstrument für die professionelle Unterrichtswahrnehmung für den Mathematikunterricht der 6. und 7. Klasse. Die drei ein- bis dreiminütigen Videoclips basieren auf Drehbüchern (auch *staged* oder *scripted Videos* genannt) und zeigen – an realen Unterricht angelehnt – mehr und weniger gelingende Unterrichtsereignisse. Nach dem Sehen der einzelnen Videoclips haben die Testpersonen 60 Sekunden Zeit eine allgemeine Einschätzung durch Beantwortung geschlossener Ratingitems (4-stufige Likertskala) zu geben. Anschließend wird die professionelle Unterrichtswahrnehmung bezüglich kooperativer Lernphasen in geschlossenen Ratingitems (4-stufige Likertskala), zur Kooperation der Lernenden in Gruppenarbeit sowie zur Kognition und Metakognition erfragt. Die Testpersonen werden durch die erste Aufgaben zwar in einen Handlungsdruck versetzt, der dem im Unterricht nahe kommt, der wissensbasierte Prozess der Interpretation wird allerdings wieder erst im Anschluss gefordert. Die so gewonnenen Aussagen können der eigentlichen professionellen Unterrichtswahrnehmung daher nur wenig gleichen.

Kaiser et al. (2017) verwenden ebenfalls staged Videoclips zu Mathematikunterricht der 8. und 10. Klasse. Die Testpersonen sehen sich jeweils drei drei- bis vierminütige Clips an und antworten anschließend auf offene und geschlossene Fragen (Likertskala). Dieses Vorgehen kann auch bei dieser Studie dazu geführt haben, dass die Testpersonen eine andere Unterrichtswahrnehmung berichten als die, die sie in einem ähnlichen Unterricht als handelnde Lehrkraft gehabt hätten, da ihre Aufmerksamkeit gelenkt und ihre Interpretation eingeschränkt wird. Dem Instrument liegt die Annahme zu Grunde, dass es eine mathematikspezifische und pädagogikspezifische professionelle Unterrichtswahrnehmung gibt. Diese wird durch die Items in Hinblick auf Klassenführung, Heterogenität, mathematisches Lernen und Lehrstrategien untersucht.

Keppens, Consuegra, Goossens, Maeyer & Vanderlinde (2019) verwenden entgegen des Trends Videos realen Unterrichts aus der Grundschule, um die professionelle Unterrichtswahrnehmung angehender Grundschullehrkräfte in Bezug auf Inklusion zu messen. Die Studierenden sehen dazu 10 Videoclip-Paare aus einer Auswahl von 15 verschiedenen zweiminütigen Clips und wählen jenen der zwei Clips aus, in dem entweder die Lehrkraft-Lernende-Interaktion oder Differenzierung besser gelingt. Anschließend beurteilen die Testperson Aussagen zu den Clips entlang einer fünfstufigen Likertskala. Damit werden die Dimensionen Bemerken und Interpretieren erfasst.

²⁰Die Videoclips stammen aus dem ViU-Projekt der Universität Münster.

Tabelle 3.1: Überblick bisheriger Instrumente zur Erfassung der professionellen Unterrichtswahrnehmung mit Bezug zu Mathematik- und Physikunterricht.

Studie	Adressierte Personen	Anzahl: Länge der Clips p.P.	Art der Clips	Gegenstand der Clips	Umgang mit Clips	Art der Items	Erfasste Fähigkeit(en)	Adressierter unterrichtlicher Aspekt
Seidel & Prenzel 2007	Lehrants-studierende, Lehrkräfte, Schul-inspizierende	8; 2 min	Realer Unterricht	Physikunterricht (aus der IPN-Videostudie), Anordnung der Videoclips entspricht dem Verlauf der Unterrichtsstunde	Clips ansehen und anhand von Rating-Items einschätzen	geschlossen: 4-stufige Likertskala	Analysekompetenz: Beschreiben, Erklären, Bewerten	Experimente, Fehlerkultur, Lernbegleitung, Zielorientierung
Kersting 2008	Erfahrene und unerfahrene Lehrkräfte	10; 1-3 min	Realer Unterricht	Mathematikunterricht 8.Klasse (TIMSS 1999)	Clips ansehen und den Umgang der Agierenden untereinander schriftlich diskutieren	offen: freie Textfelder	Analysekompetenz: Beschreiben, Erklären, Bewerten	Lehrerunterstützung, Schülerfehler
Seidel et al. 2010, Stürmer 2011	Lehrants-studierende und Fachkräfte der Lehrerbildungs-forschung	6; 2-4 min	Realer Unterricht	Phasen der Erarbeitung oder Übung in der 8./9.Klasse in Englisch, Französisch, Geschichte, Mathematik, Physik	1. Durchgang: Clip ansehen und zu Unterrichtsaspekten zuordnen; 2. Durchgang: Clip ansehen, anhand von Rating-Items einschätzen	geschlossen: 4-stufige Likertskala	professionelle Unterrichtswahrnehmung: Beschreiben, Erklären, Vorhersagen	Lernatmosphäre, Lernbegleitung, Zielorientierung
Gold et al. 2013	Lehrants-studierende	7; 1-3 min	Realer Unterricht	Grundschulunterricht (Deutsch, Englisch, Mathe) in der 2./3. Klasse mit 12 weniger und 15 gut gelingenden Ereignissen	In Kleingruppen Clips ansehen und die geschlossenen Fragen in Einzelarbeit beantworten	geschlossen: dichotom	professionelle Unterrichtswahrnehmung: Prozessmerkmale identifizieren und theoriegeleitet interpretieren	Allgegenwärtigkeit, Aufrechterhaltung des Gruppenfokus, Strukturierung des Unterrichtsverlaufs

Tabelle 3.2: Überblick bisheriger Instrumente zur Erfassung der professionellen Unterrichtswahrnehmung mit Bezug zu Mathematik- und Physikunterricht-Fortsetzung.

Studie	Adressierte Personen	Anzahl; Länge der Clips p.P.	Art der Clips	Gegenstand der Clips	Umgang mit Clips	Art der Items	Erfasste Fähigkeit(en)	Adressierter unterrichtlicher Aspekt
Wolters 2014	Lehramts-studierende	9; 3-4 min	Realer Unterricht (ViU)	Naturwissenschafts-unterricht in der Grundschule	Clips ein Mal ansehen, im Anschluss die geschlossenen Fragen beantworten	geschlossen: 4-stufige Likertskala	professionelle Unterrichtswahrnehmung: Bemerken, Interpretieren	kognitive Aktivierung
Wiedemann 2015	Lehramts-studierende	3; 1-3 min	nach Drehbuch	Mathematik-unterricht in der 6./7. Klasse mit gelingenden und weniger gelingenden Ereignissen	Clip ein Mal ansehen, in 60 Sekunden eine allgemeine Einschätzung geben, anschließend geschlossene Fragen beantworten	geschlossen: 4-stufige Likertskala	professionelle Unterrichtswahrnehmung	Phasen kooperativen Lernens
Kaiser et al. 2016	Lehrkräfte	3; 3-4 min	nach Drehbuch	Mathematik-unterricht in der 8.-10. Klasse	Clips ansehen, danach geschlossene und offene Fragen beantworten	geschlossen: Likertskala; offen: freie Textfelder	Mathematische, pädagogische professionelle Unterrichtswahrnehmung	Klassenführung, Heterogenität, mathematisches Lernen, Lehrstrategien
Keppens et al. 2019	Grundschul-lehramts-studierende	15; 2 min	Realer Unterricht	Grundschul-unterricht	Je ein aus zwei Videos in Bezug auf inklusive Aspekte auswählen; Begründen über Bewerten von Aussagen	geschlossen: 5-stufige Likertskala	professionelle Unterrichtswahrnehmung: Bemerken, Interpretieren	Inklusion: Lehrkraft-Lerner-Interaktion, Differenzierung
Scholten et al. 2019	Experten-lehrkräfte	2; 9 min	nach Drehbuch	Geographie-unterricht	Clip ansehen, bei kritischen Situationen anhalten, ihre Interpretation und Handlungsplan mitteilen	lautes Denken	professionelle Unterrichtswahrnehmung: Bemerken, Interpretieren, Handlungsplan fassen	Kartenkompetenz; geographisches Denken

Das Instrument von Scholten, Höttecke & Sprenger (2019) ist in Kooperation mit dem Instrument der vorliegenden Arbeit entwickelt wurden. Es adressiert zunächst Geographieexpertenlehrkräfte, die Fachleitungen während des Referendariats inne hatten. In den zwei circa neun Minuten langen Clips wird gestellter Geographieunterricht in der 9. Klasse gezeigt, in dem zum einen die Kartenlesekompetenz und zum anderen das geographische Denken der Lernenden wahrgenommen werden soll. Die Testpersonen sehen sich dazu die Clips an und halten diese an, wenn sie kritische Situation bezüglich dieser Aspekte wahrnehmen. Nach dem Anhalten äußern die Testpersonen laut denkend, was sie wahrgenommen haben, wie sie diese Situation interpretieren und welchen Handlungsplan sie fassen, um lernförderlich zu agieren. Die Ergebnisse der Studie sind bei Scholten et al. (2019) nachzulesen.

3.2.1 Grenzen der präsentierten Instrumente

Der Großteil der hier vorgestellten Studien erfasst die professionelle Unterrichtswahrnehmung der Testpersonen anhand der Einschätzung von Unterrichtsvideos durch Likertskalen. Dabei wird häufig ein advokatorischer Ansatz gewählt, indem die Testpersonen die Handlungen der im Video zu beobachtenden Lehrkraft bewerten (Oser et al., 2010). Mit diesem Vorgehen werden die Testpersonen in eine andere Situation versetzt, als sie es selber beim Unterrichten und somit Wahrnehmen im Klassenraum wären. Statt sich eher auf die Handlungen der Lernenden zu konzentrieren, werden Handlungen der Lehrkräfte bewertet. Zudem werden die kritischen Situationen durch die kurzen Videoclips direkt oder über die einzuschätzenden Items vorgegeben, sodass das Bemerkende von Situationen beziehungsweise das Richten von Aufmerksamkeit auf die Situation nicht von den untersuchten (angehenden) Lehrkräften ausgeht, sondern vom Instrument selbst. Neben dieser Problematik besteht die Schwierigkeit, dass die Testpersonen die vom Instrument vorgegebenen kritischen Situationen häufig erst nach dem Sehen des gesamten Clips durch Ankreuzen dichotomer oder mehrstufiger Likertskalen interpretieren müssen. Das Wissen über das gesamte Geschehen im Clip beeinflusst die Beurteilung der einzelnen kritischen Situationen und verändert damit die Ausprägung der professionellen Unterrichtswahrnehmung. Im eigenen Unterricht müssen diese kritischen Situationen von den Testpersonen selbst erkannt und vor Ablauf der gesamten Stunde interpretiert werden. Die Interpretation über Bearbeitung von Likertskalen schränkt zwar den wissensbasierten Prozess der professionellen Unterrichtswahrnehmung ein, ermöglicht aber eine ökonomische Auswertung der Daten. Diese Einschränkung ist im Interesse einer Testbarkeit vermutlich notwendig, führt aber zu einer Verminderung der Konstruktvalidität.

3.2.2 Zusammenfassung und Folgerung

Es gibt bisher quantitative Messinstrumente zur Erfassung der professionellen Unterrichtswahrnehmung unter anderem für Mathematik- und Physikunterricht, die sowohl generische Unterrichtsaspekte wie Lernbegleitung, Zielorientierung und Fehlerkultur, als auch

fachspezifische Aspekte wie Experimentieren und mathematisches Denken betreffen. Unter der Annahme, dass die generische und themenspezifische professionelle Unterrichtswahrnehmung zwar stark positiv assoziiert, aber distinkt sind (Kaiser et al., 2017; Steffensky et al., 2015), fehlt es an Instrumenten, die themenspezifische professionelle Unterrichtswahrnehmung bezüglich fachdidaktisch hoch relevanter Themen wie beispielsweise Schülervorstellungen von Lernenden und Erkenntnisgewinnung im Physikunterricht erforschen. Die Erfassung der professionellen Unterrichtswahrnehmung sollte noch performanznäher erfolgen, beispielsweise dadurch, dass Testpersonen eigenständig ihre Aufmerksamkeit lenken und direkt im Anschluss an dieses Lenken eine Interpretation der erkannten Situation abgeben müssen.

Teil II

Empirie

4 Empirische Vorüberlegungen

In diesem Kapitel wird von den oben abgeleiteten Forschungsdesideraten ausgehend das Ziel und somit die Forschungsfrage der vorliegenden Arbeit vorgestellt (Abschnitt 4.1). Die Forschungsfrage wird durch Hypothesen weiter ausdifferenziert. Anschließend wird das Forschungsdesign erläutert, mit dessen Hilfe die Forschungsfrage beantwortet wird (Abschnitt 4.3). Aus dem Forschungsdesign heraus wird die Konzeptualisierung der professionellen Unterrichtswahrnehmung aus Unterabschnitt 2.3.1 in eine Operationalisierung überführt, mittels welcher das Instrument dieser Studie entwickelt werden kann (Abschnitt 4.4).

4.1 Forschungsfrage

Wie im Theorieteil dieser Arbeit dargelegt, zählt die professionelle Unterrichtswahrnehmung zu den zentralen Fähigkeiten (angehender) Lehrkräfte, um Lernen der Lernenden zu unterstützen. Studien zeigen sowohl, dass die professionelle Unterrichtswahrnehmung von erfahrenen Lehrkräften besser entwickelt ist, als die von Junglehrkräften und Studierenden, als auch, dass sich die Fähigkeit an sich durch Interventionen mit Unterrichtstranskripten oder Unterrichtsvideos fördern lässt. Instrumente, die die Fähigkeit Unterricht professionell wahrzunehmen erfassen, fokussieren bisher bezüglich Mathematik-, Physik- und Sachunterricht überwiegend auf generische Unterrichtsaspekte wie Lernbegleitung, Lernatmosphäre und Zielorientierung. Die Erfassung der professionellen Unterrichtswahrnehmung wird noch nicht ausreichend an der eigentlichen unterrichtlichen, wahrnehmenden Handlung der Lehrkräfte orientiert. Die Forschungsfrage der vorliegenden Studie lautet daher:

Forschungsfrage

Inwiefern kann die professionelle Unterrichtswahrnehmung angehender Physiklehrkräfte bezüglich Schülervorstellungen und Erkenntnisgewinnungsprozessen in den Themenbereichen Energieumwandlung, Elektrostatik und Wärmelehre von Lernenden der 8./9. Klasse valide und reliabel mithilfe eines Onlineinstruments, bestehend aus Videostimuli sowie offenen und geschlossenen Fragen, erfasst werden?

Das Instrument wird im Rahmen dieser Studie „PUPiL-Instrument“ (**P**rofessionelle **U**nterrichtswahrnehmung angehender **P**hysiklehrkräfte) abgekürzt, vor allem um Aussagen bezüglich der Daten und des Instruments selbst eindeutig zuordnen zu können. Es wurde

zwar nachgewiesen, dass die Arbeit mit Unterrichtsvideos nicht zwingend besser ist als mit Unterrichtsskripten (Kramer et al., 2017; Sunder et al., 2016), aber das im Rahmen dieser Arbeit entwickelte, standardisierte PUPiL-Instrument soll die wahrnehmende Tätigkeit so handlungsnah wie möglich abbilden, um valide Aussagen über die Fähigkeit der professionellen Unterrichtswahrnehmung ableiten zu können. Deshalb baut das PUPiL-Instrument auf Videos und offene sowie geschlossene Fragen auf.

Die Forschungsfrage fokussiert bereits die weniger generischen, aber physikspezifischen Aspekte (Schülvorstellungen und Erkenntnisgewinnung) als Gegenstand der professionellen Unterrichtswahrnehmung, um diese Aspekte detailliert erforschen zu können. Warum ausgerechnet diese Aspekte betrachtet werden, wird weiter unten erläutert (siehe unten Unterabschnitt 4.4.2).

Die Validität des PUPiL-Instruments soll unter anderem durch die Bestätigung der folgenden Hypothesen gestützt werden:

- H1: *Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften und von (angehenden) Geographielehrkräften sowie (angehenden) Sachunterrichtslehrkräften unterscheidet sich signifikant voneinander.*

Das PUPiL-Instrument sollte in der Lage sein, die professionelle Unterrichtswahrnehmung bezüglich Physikunterricht von angehenden Physiklehrkräften und Nicht-Physiklehrkräften unterscheiden zu können, um die physikspezifische Fähigkeit valide zu messen: Aus der Konzeptualisierung heraus ist deutlich geworden, dass die professionelle Unterrichtswahrnehmung von Wissen abhängt, somit auch vom fachdidaktischen und fachlichen Wissen der Lehrkräfte. Für eine Überprüfung der Konstruktvalidität (siehe unten Abschnitt 4.2) muss das PUPiL-Instrument demnach bestätigen, dass die physikspezifische professionelle Unterrichtswahrnehmung von angehenden Physiklehrkräften besser ist als von allen anderen angehenden Fachlehrkräften, weil letztere auch über weniger physikspezifisches Professionswissen verfügen müssten. Um einen solchen Vergleich anstellen zu können, sollte die professionelle Unterrichtswahrnehmung der Physiklehrkräfte mit der von Lehrkräften verglichen werden, die zumindest eine Chance haben, physikspezifische kritische Situationen wahrzunehmen. Dazu gehören alle angehenden Naturwissenschaftslehrkräfte, aber beispielsweise nicht Lehrkräfte der gesellschaftswissenschaftlichen Fächer. Letztere würden vermutlich einzig allgemeinpädagogische Aspekte in einem Physikunterricht wahrnehmen. In dieser Studie wurden als Vergleichsgruppen daher angehende Geographielehrkräfte und Sachunterrichtslehrkräfte herangezogen. Geographielehrkräfte haben zum Teil ähnliches Fachwissen wie Physiklehrkräfte (zum Beispiel bezüglich Energie und Elektrostatik). Das fachdidaktische Wissen in den beiden Disziplinen unterscheidet sich jedoch grundlegend. Auf der anderen Seite erhalten Sachunterrichtslehramtsstudierende eine umfassende didaktische und weniger fachliche Ausbildung. Es ist daher zu erwarten, dass ein valides Instrument alle diese drei Gruppen differenziert, ohne dass Boden- oder Deckeneffekte auftreten, was bei naturwissenschaftsfremden Fächern der Fall sein könnte.

H2: *Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften verändert sich im Studienverlauf signifikant.*

Da sich fachliches und fachdidaktisches Wissen im Laufe des Studiums weiterentwickeln und Lernmöglichkeiten zum Unterrichten angeboten werden, ist zu erwarten, dass die auf Wissen und Erfahrung aufbauende professionelle Unterrichtswahrnehmung ebenso im Laufe des Studiums zunimmt und fachspezifisch ist. Es bleibt zu klären, ob dies eher am wachsenden Wissen oder an wahrgenommenen Lernmöglichkeiten wie Schulpraktika liegt.

Insgesamt ist zu erwarten, dass die Geographielehramtsstudierenden die niedrigste physikspezifische professionelle Unterrichtswahrnehmung zeigen, gefolgt von angehenden Sachunterrichtslehrkräften und angeführt von angehenden Physiklehrkräften. Außerdem ist zu erwarten, dass die Masterstudierenden unabhängig von ihrem Lehrfach durch ihre theoretische und praktische Ausbildung eine ausgeprägtere professionelle Unterrichtswahrnehmung zeigen als Bachelorstudierende. Es ist denkbar, dass fachfremde Lehramtsstudierende einen größeren Fokus auf generische Aspekte des Unterrichts legen (z.B. Classroom Management), während Physiklehramtsstudierende und Sachunterrichtslehramtsstudierende mit einem naturwissenschaftlichen Fokus sich mehr mit dem Fachlernen der Lernenden befassen. Für eine hohe Konstruktvalidität sollte das PUPiL-Instrument diese Annahmen bestätigen.

Um den Zusammenhang der professionellen Unterrichtswahrnehmung mit weiteren Einflussfaktoren ermitteln zu können, werden Variablen wie Semesterzahl, Erfahrung mit Unterrichten, fachliche Selbsteinschätzung, Fach, Universität und das physikdidaktische Wissen (Gramzow, 2015) erhoben. Für die Überprüfung der H1 werden das Fach und das physikdidaktische Wissen als unabhängige Variablen herangezogen. Bezüglich H2 werden die Semesterzahl und die Erfahrung mit Unterrichten als unabhängige Variable erhoben. Einen positiven Zusammenhang mit der professionellen Unterrichtswahrnehmung wird bezüglich der Semesterzahl, des Fachs, der Universität und des physikdidaktischen Wissens erwartet. Die Erfahrungen und Selbsteinschätzung der Testpersonen werden vermutlich keinen eindeutig positiven Zusammenhang zur professionellen Unterrichtswahrnehmung zeigen, da es distale Kovariaten sind, die auf subjektiver Wahrnehmung basieren. Sie werden aber erhoben, um Rauschen in den Daten zu minimieren.

Bevor das Forschungsdesign, das zur Beantwortung der Forschungsfrage verfolgt wurde, vorgestellt wird, wird geklärt, was Validität ist und welche Arten der Validität in der Entwicklung des Tests zur Erfassung der professionellen Unterrichtswahrnehmung zum Tragen gekommen sind.

4.2 Validität (Gültigkeit)

Das Ziel dieser Arbeit ist die Entwicklung eines *validen* Tests. Valide bedeutet dabei zunächst, dass der Test die Fähigkeit, welche durch ihn gemessen werden soll, auch misst. Insbesondere heißt das, dass die Interpretation der Testergebnisse gerechtfertigt ist (vgl.

Hartig, Frey & Jude, 2012, S. 169). Damit ist die Validität eher eine Eigenschaft der Ergebnisinterpretation, die aber durch Eigenschaften des Tests bedingt sind. Validität ist erstrebenswert, weil die Ergebnisse und Interpretationen eines validen Tests belastbar und vertrauenswürdig sind²¹. Gäbe es bereits einen etablierten und validen Test, der beispielsweise eine generische professionelle Unterrichtswahrnehmung misst, könnte man die Validität eines Tests zur Erhebung einer fachspezifischen professionellen Unterrichtswahrnehmung leicht durch die Beurteilung des Zusammenhangs der beiden Konstrukte bestätigen. Da ein solcher Test aber noch nicht besteht²², musste der Test zur Erfassung der fachspezifischen professionellen Unterrichtswahrnehmung mit Hilfe theoretischer Überlegungen validiert werden. Welche Überlegungen dies sind, wird in diesem Abschnitt vorgestellt.

Für eine gültige Interpretation der Daten muss geklärt werden, welche Arten von Interpretation die Daten des Tests zulassen. Man unterscheidet Interpretationen, die Ergebnisse *bewerten*, *verallgemeinern*, *extrapolieren*, *erklären* oder auf deren Basis *Entscheidungen* getroffen werden (vgl. Hartig et al., 2012, S. 144-145).

Das PUPiL-Instrument liefert Daten, die es erlauben, bewertende, erklärende und verallgemeinernde Interpretationen anzustellen. Zum einen sollen Interpretationen durch den Test legitimiert werden, die die Leistungen der Testpersonen erklären. Erst wenn die Erklärung der Daten valide ist, können Leistungen der Testpersonen verglichen und somit bewertet werden. Die Erklärungen und Bewertungen der Leistungen werden vor allem anhand von Zusammenhängen mit Konstrukten und Variablen wie fachdidaktisches Wissen, Semesterzahl oder Studiengang der Testpersonen validiert, deren Existenz auf theoretischen Überlegungen basieren (vgl. Abschnitt 2.2 zum PID-Modell). Diese Betrachtung findet in Kapitel 7 zur Hauptstudie statt.

Die Items des Tests mussten zudem so konstruiert werden, dass auf Grundlage der Daten eine Verallgemeinerung auf ähnliche Situationen wie in den verwendeten Items ermöglicht wurde. Wenn diese Voraussetzung nicht gegeben wäre, würde der Test eine sehr spezielle und nicht auf die alltägliche Handlung von Lehrkräften übertragbare Fähigkeit messen. Für die Möglichkeit dieser Interpretation wird zum einem in diesem Kapitel durch eine präzise Operationalisierung (Kapitel 4) und zum anderen in Kapitel 5 bei der Entwicklung von Drehbüchern, auf deren Grundlage die Videos für das PUPiL-Instrument entstanden sind, die Basis gelegt.

Zur Bewertung der Validität von Interpretationen, können verschiedene Arten der Validierung vorgenommen werden. Aktuell werden im Wesentlichen drei Arten unterschieden:

²¹Weil es schon an vielen anderen Stellen ausführlich dargestellt wurde, soll hier nur auf das Reliabilitäts-Validitäts-Dilemma hingewiesen werden. Es beschreibt, dass sehr reliable Tests häufig nur einen kleinen Ausschnitt eines Konstrukts sehr präzise misst. Damit ist der Test aber im schlimmsten Fall nur schwach valide und die Ergebnisse damit nicht belastbar zu interpretieren (vgl. beispielsweise Gramzow, 2015; Rost, 2004).

²²Wie in Kapitel 3 zum Stand der empirischen Forschung vorgestellt, gibt es bereits Tests, die eine generische professionelle Unterrichtswahrnehmung messen. Jedoch tun sie dies nicht valide genug, weil sie die Fähigkeit beispielsweise über Prozesse misst, die der realen Fähigkeit kaum ähneln.

Inhaltsvalidität, Konstruktvalidität und Kriteriumsvalidität (Hartig et al., 2012).

Unter der *Inhaltsvalidität* (ebd., S. 148-152) versteht man, ob und inwieweit die Inhalte eines Tests oder dessen Bestandteile in der Lage sind, Merkmale des definierten Konstrukts zu erfassen. Dazu gehört auch zu beurteilen, ob das Aufgabenformat des Tests zu dem zu erfassenden Konstrukt passt, ob die Testpersonen also eine adäquate Gelegenheit erhalten, ihre Fähigkeit unter Beweis zu stellen. Entscheidungen bezüglich Inhalt und Format der Items basieren auf theoretischen Argumenten und Urteilen von physikdidaktisch versierten Personen. Als speziellen Unterpunkt der Inhaltsvalidität kann außerdem betrachtet werden, ob und inwiefern Testinhalte mit relevanten Lehrplänen übereinstimmen. Diese Validität wird *curriculare Validität* genannt (ebd., S. 150). Eng damit verbunden ist die Frage danach, ob die Inhalte auch auf anderen Ebenen (Verhaltensebene, Beziehungsebene) repräsentativ für den verwendeten Kontext sind (in dieser Arbeit *ökologische Validität* genannt). Sind diese beiden Validitätsarten nicht gegeben, misst das Instrument wieder eine Fähigkeit, die im Alltag der Lehrkräfte kaum relevant ist.

Die Inhaltsvalidität wird vor allem in der Konstruktion der Videos (vergleiche Kapitel 5) und des Fragebogens (vergleiche Kapitel 6) gesichert. Die genannten Kapitel sollen demnach folgende Fragen beantworten (in Anlehnung an Hartig et al., 2012, S. 147-152):

- Sind die einzelnen Items Teil der interessierenden Gesamtheit möglicher Items, um die Fähigkeit der professionellen Unterrichtswahrnehmung zu messen?
- Stellen die Items eine repräsentative Auswahl der interessierenden Gesamtheit möglicher Items dar, anhand der die professionelle Unterrichtswahrnehmung gemessen werden kann? Das heißt insbesondere (vgl. Hartig et al., 2012, S. 152):
 - Sind alle relevanten Inhalte vorhanden?
 - Stehen die Inhalte in einem angemessenen Verhältnis zueinander?
 - Sind keine Inhalte enthalten, die sich auf etwas Irrelevantes beziehen?
- Erlaubt die theoretische Fundierung des Konstrukts der professionellen Unterrichtswahrnehmung eine Erklärung der möglicherweise beobachteten Unterschiede in den Antworten der Testpersonen?

Im Gegensatz zur Inhaltsvalidität, wird für die sogenannten *Konstruktvalidität* das Konstrukt und alle angrenzenden Konzepte betrachtet (vgl. zu Konstruktvalidität Hartig et al., 2012, S. 152-162). Die Zuverlässigkeit der Interpretation von den erhobenen Daten wird durch empirische Befunde und Argumente gesichert (ebd., S. 153). Diese Befunde können testinvariant sein oder Zusammenhänge der Daten mit anderen Variablen ausdrücken. Chan (2014) schlägt als Quellen für Validitätsargumente (a) den Inhalt des Instruments, (b) die Beziehung der Daten zu anderen Größen, (c) die innere Struktur (Dimensionalität) des Instruments, (d) die notwendigen kognitiven Prozesse zum Bearbeiten des Instruments und (e) die Konsequenzen für den Gebrauch des Instruments vor. Für eine Konstruktvalidität werden die Daten vor dem Hintergrund eines theoretischen Konstrukts interpretiert (Hartig et al., 2012). Werden Zusammenhänge oder Effekte theoretisch aufgestellt, müssen

diese für eine erfolgreiche Validierung auch empirisch nachgewiesen werden können. Für diesen Nachweis gibt es zwei verschiedene Vorgehensweisen (ebd., S. 156-158):

- Mit einem *experimentellen Ansatz* werden zunächst auf Basis theoretischer Überlegungen Hypothesen über das zu erfassende Konstrukt abgeleitet. Es wird festgehalten, welche Größen welchen hypothetischen Effekt auf das zu messende Konstrukt haben. Die Hypothesen werden dann in einem Experiment überprüft. Wenn die Ergebnisse des Experiments diese theoretischen Annahmen bestätigen, ist das ein Indiz für die Konstruktvalidität des Tests. Die Interpretationen der Daten als individuelle Ausprägung des theoretischen Konstrukts würden unterstützt.
- Im *korrelativen Ansatz* werden ebenfalls theoretische Annahmen über Zusammenhänge von Variablen aufgestellt. Die Interpretation, dass von den Daten auf das angenommene theoretische Konstrukt geschlossen werden darf, wird dann unterstützt, wenn die empirisch ermittelten Korrelationen mit den theoretisch hergeleiteten Zusammenhängen übereinstimmen.

Die Konstruktvalidierung in der vorliegenden Studie setzt sich aus beide Ansätzen zusammen. Zum einen werden theoretische Annahmen über Effekte von unabhängigen Variablen (wie Fach, Studienverlauf und Studiengang) auf die professionelle Unterrichtswahrnehmung empirisch überprüft. Zum anderen legt die Theorie über die situationsspezifischen Fähigkeiten einen Zusammenhang beispielsweise zwischen dem fachdidaktischen Wissen und der professionellen Unterrichtswahrnehmung der Testpersonen nahe.

Betrachtet man die Zusammenhänge zwischen verschiedenen Konstrukten, spricht man bezüglich eines hohen Zusammenhangs von *konvergenter Validität*. Sind die Zusammenhänge eher niedrig, spricht man von einer *diskriminanten Validität*. Für das PUPiL-Instrument muss zum einen betrachtet werden, ob die angelegten Dimensionen (Selective Attention und Knowledge Based Reasoning) durch die erhobene Daten bestätigt werden und zum anderen, ob der Zusammenhang mit dem in der Operationalisierung (vergleiche Unterabschnitt 4.4.3) integrierten Konstrukt (physikdidaktisches Wissen) auch empirisch nachweisbar bestehen.

Die Überprüfung der Konstruktvalidität findet hauptsächlich auf Grundlage der Daten der Entwicklungsstudie II und der Hauptstudie statt (vergleiche Kapitel 6 und 7). Fragen die dort beantwortet werden sollen, lauten (in Anlehnung an Hartig et al., 2012, S. 161):

- Beziehen sich die Antworten der Testpersonen auf das Konstrukt der professionellen Unterrichtswahrnehmung oder führen auch unerwartete, irrelevante Prozesse zur Lösung?
- Inwiefern können die Hypothesen zu Zusammenhängen der professionellen Unterrichtswahrnehmung mit Variablen wie physikdidaktisches Wissen, Studienverlauf und Studiengang der Testpersonen bestätigt werden?

Es kann zwar geschlossen werden, dass die Bestätigung der theoretischen Vorhersagen durch empirische Beobachtungen bedeutet, dass die Interpretation der Testwerte als individuelle Ausprägungen in dem theoretischen Konstrukt valide sind. Ist dies jedoch nicht

der Fall, muss das noch nicht heißen, dass der Test keine validen Interpretationen zulässt. Möglichkeiten dennoch einen validen Test zu erlangen sind die Überprüfung der theoretischen Annahmen oder die Überarbeitung des Tests.

Existieren bereits Kriterien, die außerhalb der Testsituation auf die zu untersuchende Fähigkeit schließen lassen, ist es möglich die sogenannten *Kriteriumsvalidität* festzustellen (vgl. Hartig et al., 2012, S. 162-168). Für die Validierung werden Zusammenhänge der Daten mit Kriterien außerhalb des Tests empirisch untersucht. In dieser Arbeit findet die Kriteriumsvalidität keine Anwendung, weil noch keine praktisch relevanten Kriterien zur Verfügung stehen, anhand derer sie überprüft werden kann. Denkbare Kriterium wären Daten zur Performanz im realen Unterricht der Testpersonen.

Die Validität ist ein Maß zur Beurteilung der Gültigkeit von Interpretationen der erhobenen Daten eines Tests. Hartig et al. (2012) merken an, dass für die Entwicklung von Tests und deren Dokumentation transparent sein sollte, welche Interpretationen möglich sind, welche naheliegend und wünschenswert sind, aber zum aktuellen Zeitpunkt nicht als gerechtfertigt betrachtet werden können. Jede gerechtfertigte Interpretation stützt die Validität eines Tests, beziehungsweise seiner Interpretation. Allerdings gibt es in jedem Fall Argumente und Befunde, welche einer Verbesserung bedürfen, weshalb die Validierung eines Tests nie vollständig abgeschlossen ist (ebd., S. 169). Für die angestrebten Interpretationen ist eine umfassende Definition des Konstrukts der professionellen Unterrichtswahrnehmung unabdingbar (ebd. S. 147). Daher folgt neben dem Forschungsdesign, anhand dessen die oben gestellte Forschungsfrage beantwortet werden soll, die Beschreibung einer Operationalisierung aller Elemente, die für die Erfassung der professionellen Unterrichtswahrnehmung notwendig scheinen.

4.3 Forschungsdesign

Um die Forschungsfrage zu beantworten und die Hypothesen H1 und H2 zu testen, wurde ein dreischrittiges Forschungsdesign angelegt (siehe Abbildung 4.1). In der Entwicklungsstudie I wurden Drehbücher für die gestellten Unterrichtsvideos entwickelt und mithilfe von physikdidaktisch Forschenden ($N = 6$) und praktizierenden Physiklehrkräften ($N = 5$) bezüglich der inhaltlichen, curricularen und ökologischen Validität anhand eines Kodierschemas und Fragebogens beurteilt. Anschließend wurden die validen Drehbücher verfilmt und für die weiteren Studien mit Text und Ton aufbereitet.

In der Entwicklungsstudie II wurden mit diesen Videos Interviews mit Lehramtsstudierenden ($N = 13$), Physiklehrkräften im Vorbereitungsdienst ($N = 5$) und erfahrenen Lehrkräften ($N = 4$) geführt, um zum einen die professionelle Unterrichtswahrnehmung der Zielgruppe des PUPiL-Instruments explorativ zu untersuchen und zum anderen geschlossene Fragen für das Instrument in der Hauptstudie zu entwickeln. Die Testpersonen haben zu diesem Zweck die Aufforderung erhalten, zu Vorkommnissen bezüglich Schülervorstellungen und Erkenntnisgewinnung in den Videos hinsichtlich Interpretation und

	Ziel	Erhebung und Analyse	Mit wem?
Entwicklungsstudie I	Inhaltlich, curricular, ökologisch valide Videos	Entwicklung und Rating (durch Kodierung) von fünf Drehbüchern	Inhaltlich: physikdidaktisch Forschenden (N=6); curricular und ökologisch: Physiklehrkräfte (N=5)
Entwicklungsstudie II	Explorative Untersuchung der prof. Unterrichtswahrnehmung Entwicklung eines geschlossenen Fragebogens	Angeleitetes lautes Denken über die Videos; qualitative Inhaltsanalyse	Lehramtsstudierende (N=13), Physiklehrkräfte im Vorbereitungsdienst (N=5), im Dienst (N=4)
Hauptstudie	Explorative Untersuchung der prof. Unterrichtswahrnehmung Validierung des Instruments (Inhalts-, Konstruktvalidität)	Online Survey bestehen aus drei Videos und Fragebogen; IRT-Skalierung und Zusammenhangsanalysen	Lehramtsstudierende (Physik, Geographie, Sachunterricht; N=160)

Abbildung 4.1: Die drei Teilstudien des Forschungsdesigns.

Handlungsplan Fassen laut zu denken. Das Datenmaterial wurden anschließend qualitativ inhaltlich analysiert. Die Erkenntnisse der Analysen sowie die geschlossenen Fragen führen zu einem Online-Survey für das PUPiL-Instrument.

Dieses Instrument wurde in einer Hauptstudie von Fragen zu demographischen Daten und dem physikdidaktischen Wissen flankiert und von $N = 160$ Testpersonen bearbeitet. Diese Erhebung hatte ebenfalls eine explorative Untersuchung der professionellen Unterrichtswahrnehmung und zusätzlich eine Konstruktvalidierung des Instruments zum Ziel. Insgesamt dienen die drei Teilstudien demnach der Validierung und Entwicklung eines Instruments, mit dem die professionelle Unterrichtswahrnehmung (angehender) Physiklehrkräfte gemessen werden kann, liefert darüber hinaus aber erste explorative Erkenntnisse über eine physikspezifische professionelle Unterrichtswahrnehmung.

Beschreibung des PUPiL-Instruments Ziel dieser Studie ist die Entwicklung eines validen Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von angehenden Physiklehrkräften. Dieses soll aus Videos und einem Fragebogen (offene und geschlossene Fragen) bestehen (siehe unten Abbildung 4.2). Als ein Item des Instruments zählt eine Situation im Video begleitet durch eine entsprechende geschlossene Frage bezüglich der professionellen Unterrichtswahrnehmung. Die teilweise Schließung des Fragebogens führt zwar zu einer verminderten Konstruktvalidität, weil das Konstrukt der professionellen Unterrichtswahrnehmung möglicherweise an sich besser durch Aufgaben des lauten Denkens²³ erfasst werden kann, sie dient aber einer leichteren Administration der Erfassung und somit einem Zugewinn an Erkenntnis über die professionelle Unterrichtswahrnehmung und ihre Förderbarkeit. Das PUPiL-Instrument soll durch das Erheben zusätzlicher Variablen wie dem fachdidaktischen Wissen und durch die Prüfung konstruktmananter Annahmen validiert werden (H1 und H2). Jede Testperson bearbeitet zwei Videos mit dem in Entwicklungsstudie II entwickelten geschlossenen Fragebogen. Nebenher werden Fragen zu demographischen Daten erhoben (unter anderem Informationen

²³Zur Methode des lauten Denkens siehe Abschnitt 6.1.

zu Lehrerfahrungen und Studium). Zudem wird ein Übungsvideo zur Verkehrssicherheit eingesetzt. Dieses beinhaltet zum eigentlichen Test analoge, aber konstruktunabhängige Situationen, mit welchen die Testpersonen den Umgang mit dem Test üben können. Für die Messung der professionellen Unterrichtswahrnehmung werden die Testpersonen dazu aufgefordert, das Video an Stellen anzuhalten, in denen es um Schülervorstellungen und Erkenntnisgewinnung geht, um anschließend ihre Wahrnehmung anhand von Distraktoren und einem Attraktor als Ordered Multiple-Choice Single Select zu benennen. Wenn keine der Antwortoptionen passend erscheint, können die Testpersonen ihre Beobachtung als offene Antwort geben. Distraktoren des Fragebogens sind entweder falsch in Bezug auf die Aufgabenstellung, behandeln also statt Schülervorstellungen und Erkenntnisgewinn Aspekte wie Classroom Management oder andere didaktische Aspekte wie Sprache im Physikunterricht. Oder die Distraktoren sind auf inhaltlicher Ebene falsch, betreffen zwar Schülervorstellungen oder Erkenntnisgewinnungsprozesse, diese sind aber nicht in der Lernsituation zu beobachten.

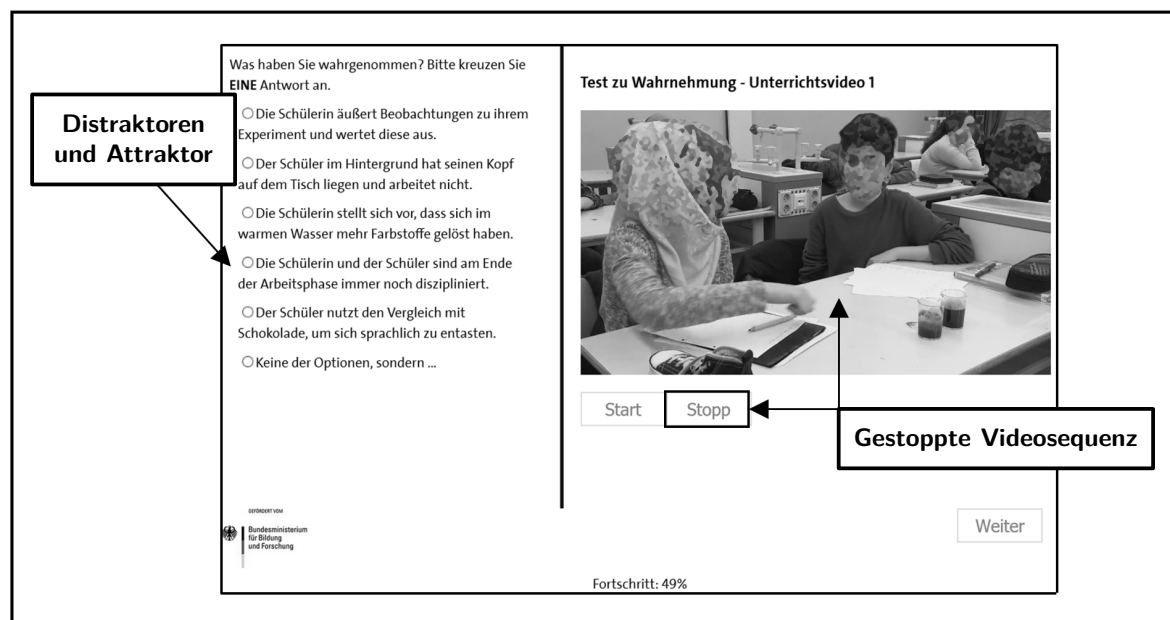


Abbildung 4.2: Beispiel für die Online-Oberfläche für das PUPiL-Instrument.

In der Hauptstudie wurde das PUPiL-Instrument zur Validierung $N = 160$ Studierenden vorgelegt, die jeweils zwei von drei scripted Unterrichtsvideos sahen und vom Instrument geleitet mittels Beantwortung offener und geschlossener Fragen (Ordered Multiple-Choice) ihre professionelle Unterrichtswahrnehmung zeigten.

Diese Fragen und somit die Distraktoren wurden in der Entwicklungsstudie II unter Einsatz von drei Videos als Stimuli in Interviews mittels lautem Erinnern mit $N = 22$ Studierenden und Lehrenden durch eine qualitative Inhaltsanalyse des Interviewmaterials entwickelt. Es wurden außerdem Übungsvideos eingesetzt, damit die Testpersonen den Umgang mit dem PUPiL-Instrument lernen können.

In der Entwicklungsstudie I wurden die Videos entwickelt, die in der Entwicklungsstudie

II und in der Hauptstudie zum Einsatz kamen. Fünf Drehbücher wurden ausformuliert und ihre inhaltliche, ökologische sowie curriculare Validität abgeprüft. Anschließend wurden drei von fünf Drehbüchern ausgewählt und verfilmt.

Die folgenden drei Kapitel stellen die Methode, Ergebnisse und Diskussion der Ergebnisse jeder Teilstudie detailliert vor. Es wird dabei einer chronologischen Reihenfolge – von Entwicklungsstudie 1, über Entwicklungsstudie 2 bis zur Hauptstudie – gefolgt. Vor der Betrachtung der Teilstudien wird im nächsten Abschnitt die Operationalisierung der professionellen Unterrichtswahrnehmung vorgestellt, die der Entwicklung des PUPiL-Instruments zugrunde liegt.

4.4 Operationalisierung

Für die Entwicklung eines Instruments zur validen Erfassung der professionellen Unterrichtswahrnehmung werden folgende Elemente als einer Operationalisierung bedürftig angesehen: Klassenraumsituation, Gegenstand der Situation sowie Prozess der professionellen Unterrichtswahrnehmung und das Konzept an sich. Nur wenn alle diese Elemente operationalisiert sind, können Entscheidungen über Inhalte für das und Umgang mit dem PUPiL-Instrument getroffen werden.

4.4.1 Kritische Klassenraumsituation

Eine kritische Situation, die von Lehrkräften im Physikunterricht wahrgenommen werden sollte, um das Physiklernen der Lernenden zu unterstützen und ermöglichen, muss folgende (nicht disjunkte) charakteristische Merkmale aufweisen (in Anlehnung an Scholten et al., 2018, S. 84):

- Die Situation findet im Physikunterricht statt.
- Die Situation liefert Hinweise auf das Lernen der Lernenden.
- Die Situation betrifft Konzepte, Begriffe und Methoden der Physik.
- Es werden physikdidaktisch relevante Aspekte in der Situation berührt.
- Die Geschehnisse der Situation sind zentral für den Physikunterricht und somit für das Physiklernen.
- Die Situation erfordert eine Entscheidung (handeln oder nicht handeln) der Lehrkraft.

In dieser Studie wurde als Bezeichnung des Gegenstandes der professionellen Unterrichtswahrnehmung nicht der Begriff „incident“ – also Vorfall oder Ereignis – gewählt, sondern der Begriff „Situation“. Diese Setzung rührt aus der Überlegung, dass die professionelle Unterrichtswahrnehmung nicht nur „Vorfälle“ oder „Ereignisse“ betrifft, in denen Neuwertiges, *Vorfallendes* geschieht, sondern auch diejenigen Zustände, in denen alles seinen

erwarteten und gewohnten Gang geht. Bisher gibt es noch keine Ergebnisse darüber, wie häufig ein durch diese Merkmale hervorgebrachter Stimulus erscheinen muss, bevor die professionelle Unterrichtswahrnehmung diesbezüglich in das Bewusstsein einer Testperson gelangen kann. Für einen Test zur Erhebung dieser Fähigkeit bietet es sich daher an, mehrere Stimuli zu gleichen lernrelevanten kritischen Situationen zu beinhalten. Welche physikdidaktischen Aspekte in dieser Studie als dringend wahrzunehmen (normativ) gesetzt werden, wird im nächsten Unterabschnitt hergeleitet.

4.4.2 Gegenstand der kritischen Situation

Der Gegenstand der kritischen Situation, die die professionelle Unterrichtswahrnehmung betrifft, lässt sich aus den Ausführungen von Goodwin (1994) und Überlegungen zum Physikunterricht ableiten. Nach Goodwin werden innerhalb einer Profession die Gegenstände wahrgenommen, die für diese Profession von (kollektiver) Bedeutung sind. Welche Gegenstände dies für den Physikunterricht sind, wird aus dem Lernbegriff abgeleitet.

Bei der Entwicklung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung ist der alltägliche Physikunterricht als Referenzrahmen von hoher Bedeutung. Möchte man eine Fähigkeit von (angehenden) Lehrkräften handlungsnah abbilden, sollten die für die Messung notwendigen Aufgaben auch möglichst nah an der eigentlichen Tätigkeit der Lehrkraft orientiert sein. Andernfalls kann von der Bearbeitung der Aufgaben nicht valide auf die Fähigkeit der angehenden Lehrkräfte geschlossen werden. Sowohl der Inhalt der gezeigten Situationen, als auch die Art der Interaktionen der Lehrkräfte mit diesen Inhalten bestimmt, wie übertragbar die gemessene Fähigkeit auf die reale Fähigkeit ist. Der Gegenstand der kritischen Situationen in dieser Arbeit ist das fachliche Lernen der Lernenden.

Im Angebots-Nutzungs-Modell von Lipowsky (2006) (siehe Abbildung 4.3) sind Bedingungsfaktoren für den Schulerfolg (Leistung, Interesse und Motivation, Lernfreude und Selbstkonzept) der Lernenden dargestellt. Neben globaleren Einflussfaktoren wie den kulturellen, ökonomischen und schulischen Bedingungen, sowie familiären und peerbezogenen Voraussetzungen der Lernenden haben die Kompetenzen der Lehrkräfte und die Unterrichtsqualität einen großen Einfluss auf das (fachliche) Lernen von Lernenden (Lipowsky, 2006). Um den Gegenstand der hier erhobenen professionellen Unterrichtswahrnehmung operationalisieren zu können, wird aus physikdidaktischer und lernpsychologischer Perspektive definiert, was Lernen (im Physikunterricht) ist. Diese Perspektiven auf Lernen sind für diese Arbeit von Bedeutung, weil für eine inhaltsvalide Messung der physikspezifischen professionellen Unterrichtswahrnehmung wichtig ist, dass der Gegenstand im PUPiL-Instrument dem Gegenstand der professionellen Unterrichtswahrnehmung in der Realität entspricht. Sind die gezeigten Lerngelegenheiten nicht physikspezifisch und nicht authentisch, könnte es dazu kommen, dass eine Fach- oder Themenspezifität nicht nachgewiesen werden kann. Somit wäre das Instrument nicht ausreichend valide.

Allgemein kann der Physikunterricht durch das Lernen von Konzepten, Begriffen und Methoden, Kommunikation sowie das Beurteilen innerfachlicher und überfachlicher Frage-

stellungen charakterisiert werden. Im Physikunterricht sollen vor allem Kompetenzen (in Anlehnung an Weinert, 2001b, S. 27, 28) zu Fachwissen, Erkenntnisgewinnung, Kommunikation und Beurteilen erlernt werden (vgl. Neumann, Kauertz, Lau, Notarp & Fischer, 2007; Schecker & Parchmann, 2006). Welche Kompetenzen Lernenden bezüglich dieser kognitiven Aspekte von Physikunterricht erwerben sollen, kann aus den normativen Anforderungen in den Bildungsstandards abgeleitet werden (Beschluss der Kultusministerkonferenz, 2004). Lernprozesse im Physikunterricht werden im Folgenden entlang der durch die Kultusministerkonferenz geforderten Kompetenzen Fachwissen und Erkenntnisgewinnung im Physikunterricht erläutert. Die Einschränkung auf diese beiden Kompetenzen erfolgt, weil sie die grundlegenden Aspekte des Physiklernens direkt betreffen (Konzepte, Begriffe und Methoden der Physik). Es hätten auch alle Kompetenzen abgedeckt werden können, dafür wären aber mehr Videos vonnöten. Je mehr Videos in das Instrument mit aufgenommen werden, desto länger wird entweder die Testung (wenn alle Testpersonen alle Videos sehen) oder desto größer müsste die Stichprobe für das Rotationsdesigns sein. Es wird im Folgenden auf der konstruktivistischen Lern-Theorie basierend geklärt (siehe dazu beispielsweise Renkl, 2008; Widodo & Duit, 2004; Gerstenmaier & Mandl, 1995), *was* gelernt wird, also was Lernen im Physikunterricht bedeutet und somit den Gegenstand der professionellen Unterrichtswahrnehmung spezifisch macht.

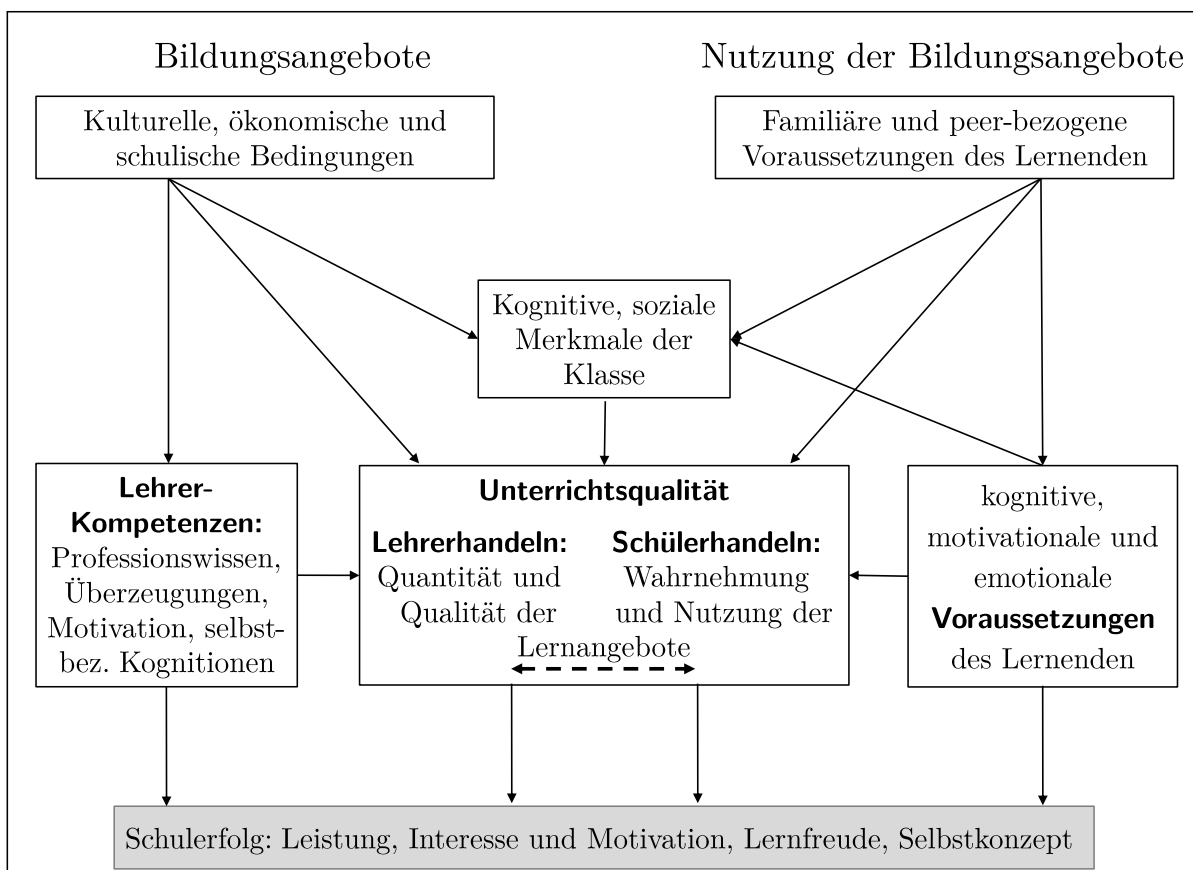


Abbildung 4.3: Angebots-Nutzungs-Modell nach Lipowsky (2006, S. 48).

Fachwissen Die Bildungsstandards (vgl. hier und im ganzen Absatz: Beschluss der Kultusministerkonferenz, 2004) sehen für die Lernenden vor, dass sie physikalisches Fachwissen in Bezug auf vier Basiskonzepte – Materie, Wechselwirkung, System, Energie – erwerben. Das Wissen beinhaltet dabei „Wissen über Phänomene, Begriffe, Bilder, Modelle und deren Gültigkeitsbereiche sowie über funktionale Zusammenhänge und Strukturen“ (ebd., S. 8). Dieses Wissen soll den Lernenden helfen, physikalische Aufgaben und Probleme zu bearbeiten.

Synthetisierend können diese Anforderungen (vgl. Beschluss der Kultusministerkonferenz, 2004, S. 11) zusammengefasst werden zu den Forderungen, dass Lernenden Konzepte und Begriffe zu physikalischen Phänomenen bilden und diese Konzepte zur Anwendung bringen. Konzepte werden in dieser Arbeit als ein Netz aus Wissens-elementen bezüglich der physikalischen Welt verstanden. Hierbei muss anerkannt werden, dass Lernenden Informationen häufig auf Grundlage vorhandener Vorerfahrungen verarbeiten (Niebert, 2010). Das bedeutet, dass sie diese Konzepte und Begriffe über die physikalische Welt immer vor dem Hintergrund ihrer Vorerfahrungen herstellen (vgl. Duit, Hopf, Schecker & Wilhelm, 2018, S. 2).

Im Physikunterricht lernen Lernenden durch die eigenständige Auseinandersetzung mit ihrer Umwelt, genauer mit der physikalischen Beschaffenheit und den physikalischen Zusammenhängen dieser Umwelt, die Phänomene hervorbringen. Einer der Schwerpunkte für den Gegenstand der professionellen Unterrichtswahrnehmung sind die Schülervorstellungen zu Begriffen und Konzepten der Physik.

Schülervorstellungen In der Auseinandersetzung mit Phänomenen aus der Umwelt werden die Vorerfahrungen, die Lernenden bereits mit dem Gegenstand, an dem sie gerade lernen, virulent. Diese kognitiven Voraussetzungen (Vorerfahrungen) haben ihren Ursprung häufig in vorangegangenen Lernprozessen bezüglich der physikalischen Welt durch Beobachtung, Einsicht oder Konstruktion von Vorstellungen und werden als *Schülervorstellungen* bezeichnet (Duit et al., 2018). Schülervorstellungen können als von der Fachdidaktik diagnostizierte Vorstellungen, die von physikalischen Vorstellungen abweichen (Duit et al., 2018) verstanden werden. Eine Schärfung erfahren diese ersten Gedanken von Lernenden zu physikalischen Phänomenen in der Alltagssprache, in Veranschaulichungen (Modelle und Graphiken), Medien und Alltagserfahrungen (vgl. Duit et al., 2018, S. 13, 14). Da die Alltagssprache breiter und unschärfer als die Wissenschaftssprache ist, teilweise veraltete Begriffe gebraucht (beispielsweise den Begriff Kraftwerk) und Darstellungen in den Medien wie Nachrichten und Fernsehen häufig physikalisch unangemessen sind, weisen viele Schülervorstellungen Aspekte auf, die aus physikalischer Perspektive inadäquat sind (ebd., S. 13). Dennoch folgen die Konzepte, welche auf Schülervorstellungen basieren, einer eigenen, inneren Logik und können Alltagsphänomene auf einer basalen Ebene erklären (ebd., S. 12). Auch aus diesem Grund sind Schülervorstellungen widerstandsfähig, vor allem gegenüber Veränderungen im Physikunterricht. Werden Schülervorstellungen im Unterricht ignoriert und vermeintlich neue Inhalte gelehrt, kommt es für die Lernenden (meist unbewusst) zu Lernschwierigkeiten. Bloßes deklaratives Lernen kann ohne

Beachtung der Vorerfahrungen der Lernenden zum Beispiel dazu führen, dass die tiefer liegenden, konstruierend erworbenen Konzepte nicht angerührt und nicht mit dem neuen Wissen in Einklang gebracht werden (Niebert, 2010; Duit et al., 2018). Die Folge dessen ist häufig ein in sich nicht schlüssiges Konzept bei den Lernenden, welches das Verstehen oder Anerkennen von physikalisch angemessenen Konzepten hemmt.

Dass Lernen auch als Ko-Konstruktion von Sinn verstanden werden kann, zeigt sich darin, dass Lehrkräfte und Lernende ihre Vorstellungen abgleichen müssen (vgl. Duit et al., 2018, S. 8). Dieser Abgleich erfolgt aktiv konstruktiv. Es wird also nur deutlich, ob Lehrende und Lernende das Gleiche meinen, wenn sie sich hinreichend intensiv darüber verständigen. Dass Lernenden ihre Vorstellungen häufig nicht verbalisieren können²⁴, spricht dafür, dass Schülervorstellungen wenigstens anteilig implizit vorliegen, die über Reflexion (angeleitet durch Lehrkräfte) in deklaratives Wissen überführt werden können. Lernenden sind in der Lage, neben ihren Schülervorstellungen in ihrem Netz aus Vorstellungen die physikalisch angemessenen Vorstellungen koexistieren zu lassen (vgl. Duit et al., 2018, S. 10). Die physikalischen Konzepte erhalten ihren Sinn sogar erst durch die Vernetzung mit dem Wissenssystem der Lernenden. Häufig werden Aspekte der eigenen, vorunterrichtlichen mit der physikalisch adäquaten Vorstellung verbunden (vgl. Widodo & Duit, 2004, S. 132). Man spricht dann von einer sogenannten „Hybridvorstellung“. Schülervorstellungen entstehen auch und sind deshalb so stabil, weil sie unzählige Male im Alltag durch Beobachtung und Wiederholung gefestigt wurden (vgl. Hopf & Wilhelm, 2018, S. 28). Der Unterricht sollte den Lernenden helfen, in variablen Situationen zu entscheiden, ob „die wissenschaftlichen Vorstellungen angemessener und fruchtbarer sind als ihre bisherigen Vorstellungen“ (Widodo & Duit, 2004, S. 132).

Grundlegender Aspekt des Lernens im Physikunterricht bezüglich der Kompetenz Fachwissen ist also das Umstrukturieren von bestehenden Wissenssystemen²⁵. Man nennt diesen Prozess in der Forschung *Konzeptwechsel*. Mögliche Änderungen von Schülervorstellungen hin zu einem physikalischen Konzept bieten die Prozesse der Assimilation und Akkommodation. Eine solche Änderung hat zum Ziel, dass die Lernenden in der Lage sind, zwischen den in ihnen wirkenden Konzepten zu unterscheiden, die Stärken und Schwächen ihrer Konzepte zu kennen und je Situation ein adäquates Konzept zu wählen. Mit Schülervorstellungen angemessen umzugehen bedeutet also, Lernenden beim Lernen zu unterstützen. Mittels didaktischer Rekonstruktion kann Unterricht auf dem Wissen über Schülervorstellungen basierend lernwirksam gestaltet werden²⁶.

Zusammenfassens kann vorerst festgehalten werden, dass Lernenden bezüglich der Kompetenz Fachwissen aus den Bildungsstandards unter anderem angemessene Schülervorstellungen ausbilden sollen und dabei von Lehrkräften unterstützt werden müssen.

Lernenden haben nicht nur Vorstellungen zu den Konzepten und Begriffen der Physik,

²⁴Der Abgleich der Vorstellungen von Lehrkräften und Lernenden gestaltet sich daher auch eher schwierig.

²⁵Man beachte, dass fachliches Lernen mehr als das Umstrukturieren von Vorstellungen ist.

²⁶Die ausführliche Literatur zu Konzeptwechsel im naturwissenschaftlichen Unterricht kann im Rahmen dieser Arbeit nicht zusammengefasst werden. Es empfiehlt sich die Lektüre von diSessa (1988), Gropengießer, Harms & Kattmann (2013), Hopf & Wilhelm (2018), Piaget (1976), Posner, Strike, Hewson & Gertzog (1982), Stark (2003), Vosniadou (1994).

sondern auch zu den wissenschaftlichen Arbeits- und Denkweisen. Daher fordern die Bildungsstandards die Förderung der Erkenntnisgewinnung von Lernenden (Beschluss der Kultusministerkonferenz, 2004). Die Kompetenzen des Experimentierens und der Erkenntnisgewinnung als naturwissenschaftliche Methode werden im Folgenden skizziert.

Erkenntnisgewinnung Nach dem Beschluss der Kultusministerkonferenz (2004) ist die Erkenntnisgewinnung ein Prozess, den Lernenden erlernen müssen, um die Entstehung physikalischen Wissens zu verstehen, dieses anwenden und zur Problemlösung nutzen zu können (ebd., S. 14). Man kann diesen Prozess in fünf Tätigkeiten ausdrücken. Die erste Tätigkeit ist das *Wahrnehmen*, welches das Beobachten und Beschreiben eines Phänomens, das Erkennen einer Problemstellung und das Vergewärtigen der eigenen Wissensbasis beinhaltet (vgl. hier und im ganzen Abschnitt: Beschluss der Kultusministerkonferenz, 2004, S. 14). Als zweite Tätigkeit zählt das *Ordnen*, also das Zurückführen auf und Einordnen in Bekanntes sowie das Systematisieren von neuen Lerninhalten. Drittens dient das *Erklären* laut der nationalen Bildungsstandards dazu, die Realität zu modellieren und Hypothesen über die physikalischen Zusammenhänge der Realität aufzustellen. Dabei werden nötige Randbedingungen genutzt, um eine logische Verknüpfung zwischen einem allgemeingültigen Gesetz und dem zu erklärenden Phänomen herzustellen (vgl. Bartels, Kulgemeyer & Schecker, 2016, S. 1). Das *Prüfen* wird beim Experimentieren, Auswerten, Beurteilen sowie dem kritischen Reflektieren von Hypothesen erlernt. Als letzte Tätigkeit gibt es das *Bilden von Modellen*, verstanden als Idealisieren, Beschreiben von Zusammenhängen, Verallgemeinern, Abstrahieren, Begriffe bilden, Formalisieren, Aufstellen einfacher Theorien und Transferieren. Bezüglich der Erkenntnisgewinnung werden unter anderem motorische Kompetenzen in Verbindung zu kognitiven Kompetenzen, wie den Umgang mit Untersuchungsmethoden (Beobachten, Experimentieren), und rein kognitiven Kompetenzen, wie das Fällen rationaler Schlussfolgerungen aus Experimentierergebnissen gefördert (vgl. Gebhard, Höttecke & Rehm, 2017, S. 55). Dabei wird immer „auf bereits vorhandenes Wissen zurückgegriffen“ (ebd., S. 15).

Weitere Kompetenzen der KMK-Bildungsstandards Neben dem Erlernen von Konzepten und Methoden der Naturwissenschaft ist grundlegend, dass Lernenden Fähigkeiten zu adressatengerechter und sachbezogener Kommunikation erlangen (vgl. Beschluss der Kultusministerkonferenz, 2004; Kulgemeyer & Schecker, 2009) und innerfachliche sowie überfachliche Bewertungen von Sachverhalten mit Naturwissenschaftsbezug durchführen können (Bögeholz, Höhle, Höttecke & Menthe, 2018; Beschluss der Kultusministerkonferenz, 2004; Schecker & Höttecke, 2007; Thode & Schecker, 2017).

Zusammenfassend wird festgehalten, dass Lernen im Physikunterricht derzeit aus konstruktivistischer Sicht definiert wird. Lernen bedeutet unter anderem die Anreicherung von Schülervorstellungen (zu Begriffen, Konzepten und Methoden) hin zu Vorstellungen, die die physikalische Welt erklären können. Gegenstand des Lernens sind die Kompetenzen Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung.

Zu einem Wahrnehmungsprozess (Bottom-Up oder Top-Down) gehören immer mindestens ein_e Akteur_in mit den jeweiligen Dispositionen wie Wissen und Erfahrungen sowie die Situation mit ihren Merkmalen. Wie der Prozess operationalisiert werden kann, wird im nächsten Unterabschnitt dargelegt.

4.4.3 Der Wahrnehmungs-Prozess

Die professionelle Unterrichtswahrnehmung einer kritischen Physikunterrichtssituation von Lehrkräften kann nicht direkt erfasst werden. Daher wird der Prozess über die Situation und deren Stimuli bis hin zur Disposition der Lehrkraft operationalisiert (vgl. Abbildung 4.4).

Die professionelle Unterrichtswahrnehmung ist ein kognitiver Prozess und betrifft, wie in Unterabschnitt 2.1.2 zu Selective Attention und Knowledge Based Reasoning bereits beschrieben, Situationen, die sich durch Mehrdimensionalität, Unmittelbarkeit und Unvorhersehbarkeit auszeichnen (vgl. Scholten et al., 2018, S. 82). Dieser kognitive Prozess unterscheidet sich daher von anderen kognitiven Prozessen wie Unterrichtsplanung oder Reflexionen nach dem Unterricht, weil ersterer einem Zeitdruck ausgesetzt ist und die anderen nicht (ebd.).

Die Operationalisierung des Prozesses beginnt bei der kritischen Situation (siehe Abbildung 4.4).

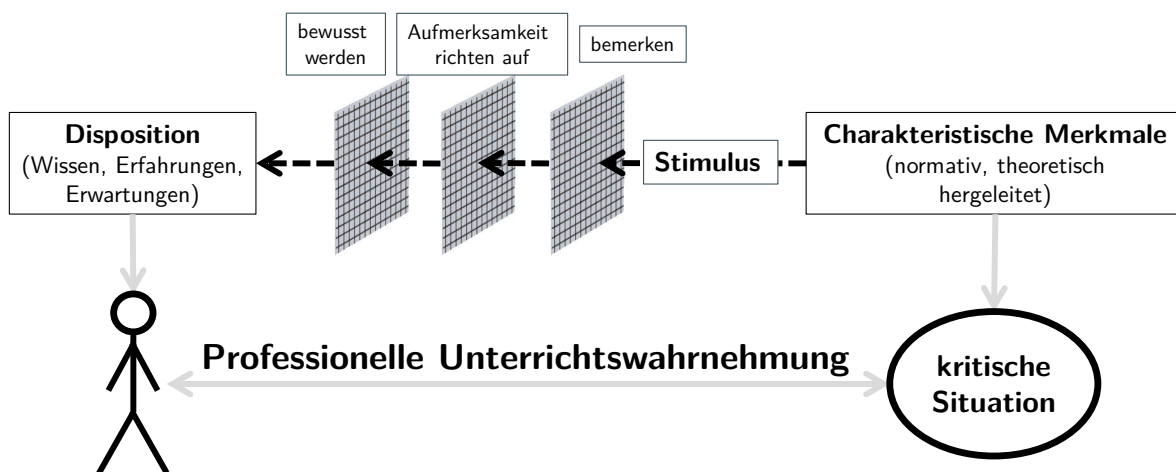


Abbildung 4.4: Operationalisierter Prozess der professionellen Unterrichtswahrnehmung.

Eine Situation ist dann kritisch, wenn sie die in Unterabschnitt 4.4.1 genannten charakteristischen Merkmale erfüllt. Die professionelle Unterrichtswahrnehmung ist ein situativer und kontextuierter Prozess und umfasst daher nicht nur den kognitiven Prozess der Lehrkräfte, sondern auch den umgebenden spezifischen Kontext (ebd., S. 84). Die Merkmale dieses Kontexts bilden Stimuli, die von Lehrkräften bemerkt werden können (Lamme,

2003). Auf bemerkte Stimuli kann Aufmerksamkeit gerichtet werden (ebd.). Wenn die Aufmerksamkeit bewusst auf den Stimulus gerichtet wird, kann die professionelle Unterrichtswahrnehmung über ein Instrument erfasst werden. Bei der professionellen Unterrichtswahrnehmung handelt es sich wie bereits berichtet aber nicht nur um einen Bottom-up-Prozess, in dem der Stimulus der Situation Aufmerksamkeit der Person erlangt, sondern auch um einen Top-down-Prozess, in dem die Person aktiv Aufmerksamkeit auf einen Stimuli der Situation richtet. Für das PUPiL-Instrument sind nur diejenigen Prozesse messbar, die einer bewussten Wahrnehmung von kritischen Situation entsprechen, diese kann aber sowohl durch einen Bottom-Up-Prozess als auch durch einen Top-Down-Prozess initiiert werden. Die kritische Situation wird dann mithilfe von Dispositionen (beispielsweise Wissen, Erfahrungen und Erwartungen) interpretiert und möglicherweise eine (passende) Handlungsoptionen generiert (vgl. Kapitel 2 zu theoretischen Überlegungen zur professionellen Unterrichtswahrnehmung).

Da es sich hier um ein Instrument handelt, dass die physikspezifische professionelle Unterrichtswahrnehmung erfassen soll, muss das Konzept weiter eingegrenzt werden, als es bisher in Kapitel 2 geschehen ist.

4.4.4 Operationalisierung des Konzepts

Damit eine Validierung möglich ist, bedarf es einer konkreten Konstrukt-Beschreibung. Diese Arbeit erhebt nicht den Anspruch die professionelle Unterrichtswahrnehmung im Allgemeinen erfassen zu können. Wie in Abbildung 4.5 ersichtlich konzentriert sich diese Studie erstens auf die professionelle Unterrichtswahrnehmung von Lehramtsstudierenden (erste Eingrenzung).

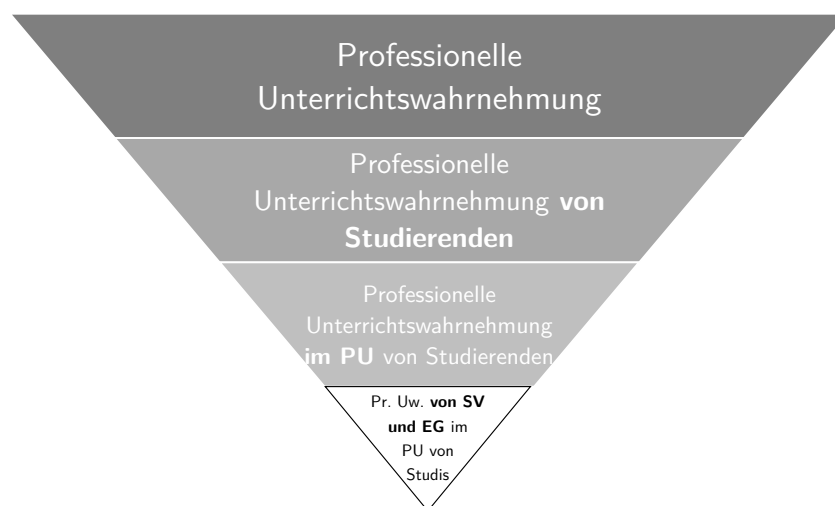


Abbildung 4.5: Operationalisierung des Konzepts der professionellen Unterrichtswahrnehmung; *Pr. Uw.* steht für professionelle Unterrichtswahrnehmung, *SV* steht für Schülervorstellungen, *EG* steht für Erkenntnisgewinnung, *PU* steht für Physikunterricht, *Studis* steht für Studierende.

Es handelt sich zweitens um professionelle Unterrichtswahrnehmung bezüglich Physikunterricht (zweite Eingrenzung) und drittens um die Wahrnehmung physikspezifischer Aspekte der Schülervorstellung und Erkenntnisgewinnung (dritte Eingrenzung). Das PUPiL-Instrument lässt daher nur Interpretationen zu, die sich auf eine physikspezifische Unterrichtswahrnehmung der Aspekte Schülervorstellungen und Erkenntnisgewinnung beziehen. Interpretationen einer generischen Unterrichtswahrnehmung oder eine physikspezifischen Wahrnehmung bezüglich anderer (fachdidaktisch relevanter) Aspekte wie beispielsweise die Kommunikationskompetenz sind dafür nicht valide. Im nächsten Kapitel wird auf diesen empirischen Vorüberlegungen aufbauend die erste Entwicklungsstudie vorgestellt. Sie beinhaltet die Entwicklung der Videos für das PUPiL-Instrument der Hauptstudie.

5 Entwicklungsstudie I

In diesem Kapitel wird die Entwicklung der Videos für den Einsatz im PUPiL-Instrument für die Hauptstudie beschrieben. In Abschnitt 5.1 wird berichtet, wie die Drehbücher entwickelt und anschließend inhaltlich, curricular und ökologisch validiert wurden. Die Ergebnisse der Validierung werden vorgestellt und somit wird die Forschungsfrage bezüglich der Validität zum Teil beantwortet (Unterabschnitt 5.1.4). Abschließend wird die endgültige Auswahl an Videos begründet (Abschnitt 5.2).

5.1 Methode

In einem mehrstufigen Verfahren wurden Videographien von gescriptetem Physikunterricht angefertigt, die inhaltlich, curricular und ökologisch valide sein mussten, um die Forschungsfrage zu beantworten (siehe dazu auch Tabelle 5.1):

- 1.) Ein erstes Drehbuch von Unterricht (bezüglich Weg-Zeit-Diagrammen) wurde in einer Vorstudie entwickelt, validiert und gefilmt.
Dadurch wurden Erfahrungen mit Drehbuchschreiben, Filmdrehen, dem Effekt der Videos auf die Testpersonen sowie der Qualität und der Effekte der Aufgabenstellung zur Erfassung der professionellen Unterrichtswahrnehmung gesammelt.
- 2.) Nachdem die professionelle Unterrichtswahrnehmung genau konzeptualisiert war, wurden fünf Drehbücher von Unterricht nach dieser Konzeptualisierung entwickelt. Die Drehbücher werden in Unterabschnitt 5.1.2 genauer beschrieben, um die Ergebnisse der Validierung darauf beziehen zu können. Außerdem sind Teile der Drehbücher Grundlage für die Entwicklungsstudie II und eine inhaltliche Vorstellung der Drehbücher hilft beim Verständnis aller Folgeergebnisse.
- 3.) Jedes dieser fünf Drehbücher beinhaltet bereits eine Anzahl vorher festgelegter, intendiert kritischer Situationen in Anlehnung an die charakteristischen Merkmale der operationalisierten Situationen (vergleiche Tabelle 5.12).
- 4.) Die Konzeptualisierung der professionellen Unterrichtswahrnehmung führte zu einem Kodierschema (siehe Unterabschnitt 5.1.3), entlang dessen sechs in Bezug auf professionelle Unterrichtswahrnehmung geschulte physikdidaktisch Tätige die Inhaltsvalidität der Drehbücher mit einem Ratingverfahren sicher stellten.
Das Vorgehen wird in Unterabschnitt 5.1.3 erläutert.

- 5.) Fünf Lehrkräfte sicherten durch einen Fragebogen die curriculare und ökologische Validität der inhaltlich validierten Drehbücher.
- 6.) Drei Videos wurden ausgewählt und anschließend verfilmt.

Tabelle 5.1: Schritte der Entwicklungsstudie I zur Sicherung der inhaltlichen, curricularen und ökologischen Validität.

Schritt	Funktion
1 – Entwicklung, Verfilmung eines ersten Drehbuchs	Erfahrungen mit Drehbuchschreiben, Filmdrehen, dem Effekt der Videos auf die Testpersonen und der Qualität sowie der Effekte der Aufgabenstellung zur Erfassung der professionellen Unterrichtswahrnehmung wurden gesammelt.
2 – Entwicklung fünf weiterer Drehbücher	Die (verfilmten) Drehbücher dienen als Stimuli im Instrument.
3 – Einbezug der Operationalisierung	Die Drehbücher enthalten kritische Situationen, die von Testpersonen wahrgenommen werden sollen.
4 – Kodierschema geleitetes Rating der Drehbücher	Die Drehbücher wurden einer inhaltlichen Validierung durch ein Ratingverfahren unterzogen.
5 – Fragebogen-Befragung von Lehrkräften	Die Drehbücher wurden einer curricularen und ökologischen Validierung durch einen Fragebogen unterzogen.
6 – Auswahl von drei Videos	Um die Testzeit zu optimieren wurden drei validierte Videos für eine Verfilmung und den Einsatz im Instrument ausgewählt.

In den folgenden Abschnitten werden die Erfahrungen mit einem ersten Video berichtet, Inhalte der Drehbücher beschrieben, wobei nur auf die intendiert kritischen Situationen fokussiert wird. Außerdem werden sowohl das Kodierschema, der Fragebogen als auch das Vorgehen der Validierung vorgestellt.

5.1.1 Schritt 1 – Pilotierung mit einem ersten Video

Bevor die Elemente des PUPiL-Instruments (Drehbücher, Videos und Fragebogen) entwickelt werden konnten, wurden zunächst erste Erfahrungen mit einem Probevideo gesammelt, um die Qualität der finalen Videos zu erhöhen. In einem Drehbuch wurden Situationen beschrieben, in denen Lernende den Umgang mit Weg-Zeit-Diagrammen lernen und dabei Fehler machen oder erfolgreich Aufgaben lösen. Es werden Interaktionen von Lernpaaren, aber auch Diskurse im Plenum beschrieben.

Dieses Drehbuch wurde mithilfe von Lernenden eines Hamburger Gymnasiums und dem Team des Medienzentrums in der Fakultät für Erziehungswissenschaft an der Universität Hamburg verfilmt. Die Kameraführung wurde an die Bewegung einer Lehrkraft im Klassenraum angelehnt, wenn Partnerarbeit gezeigt wurde, um die Testpersonen die Situationen aus der Perspektive der Lehrkraft betrachten zu lassen.

Während dieses Prozesses konnten Erfahrungen darüber gesammelt werden, wie ein Dreh-

buch geschrieben und mit einem Storyboard und Drehplan ergänzt werden muss, damit das anvisierte Video entstehen kann. Diese Schritte waren wichtig, um herauszufinden inwiefern die Drehbücher möglichst genau in Videos überführt werden können.

Anschließend wurde das Video von $N = 5$ Testpersonen (drei Physiklehramtsstudierende, eine in der Physikdidaktik tätige Person und eine Person im BWL-Studium) mit verschiedenen Aufforderungen bearbeitet. Jede Testperson bekam nur eine der folgenden Aufforderungen. Die Ergebnisse dieses Schrittes dienten dazu eine Aufforderung zu wählen, die bei Testpersonen zu den gewünschten Handlungen führt. Evoziert die Aufforderung keine Handlung, die der professionellen Unterrichtswahrnehmung nahe kommt, kann von ihr aus nicht auf diese latente Fähigkeit geschlossen werden. Die Aufforderungen gingen dabei zunächst von Situationen aus, die eine Handlungsanforderung an die Lehrkraft bieten und entwickeln sich zu Fragen nach dem Lernen der gezeigten Lernenden:

1. Wann nehmen Sie etwas wahr, das eine mögliche Lehrerintervention nach sich zieht?
2. An welchen Stellen sollte Ihrer Meinung nach die Lehrkraft Schülerdenken / -handeln bemerken?
3. Wann fallen Ihnen lernrelevante Situationen auf, die Sie als Lehrkraft wahrnehmen sollten?

Die Sichtung des transkribierten Interviewmaterials zeigte für die einzelnen Aufforderungen verschiedene Schwierigkeiten.

Der Begriff „Intervention“ in der ersten Aufforderung ist in verschiedenen fachlichen Kontexten verschieden belegt und demnach nicht als selbsterklärende, eindeutige Arbeitsanweisung geeignet.

Bei der zweiten Aufforderung zeigte sich, dass die Testperson nicht sicher war, ob nur das Handeln der Lernenden im Mittelpunkt steht oder auch die Lehrkraft in den Blick genommen werden soll. Es war zudem nicht klar, was Schülerhandeln und Schülerdenken ist.

Die Abänderung in der dritten Aufforderung zeigt sich als erfolgreich. Allerdings muss ergänzt werden, dass das mögliche Handeln der Lehrkraft im Anschluss beschrieben werden soll.

Die Testpersonen empfanden die Kameraführung als irritierend, da nicht klar war, dass die Lehrkraftperspektive gezeigt wurde. Es wurde für den weiteren Test die dritte Aufforderung in umformulierter Form verwendet und um die Aufforderung ergänzt, mögliches lernförderliches Handeln der Lehrkraft zu äußern:

1. Bitte stoppen Sie das jeweilige Video immer dann, wenn Sie etwas in Bezug auf ...
 - ... Vorstellungen über fachliche Begriffe und Konzepte der Schüler_innen,
 - ... Erkenntnisgewinnungsprozesse der Schüler_innen,
 - ... Vorstellungen über Erkenntnisgewinnung der Schüler_innen wahrnehmen²⁷.

²⁷Es wurde zu diesem Zeitpunkt noch unterschieden, ob die Testpersonen Vorstellungen über Konzepte und Begriffe oder Erkenntnisgewinnung wahrnahmen. Über die Analyse der Interviews wurde aber deutlich, dass letzterer Aspekt immer als Teil der Erkenntnisgewinnung behandelt wurde.

(Das Stoppen des Videos muss unmittelbar und in sehr kurzer Zeit nach dem Prozess des Erkennens geschehen, damit dem unmittelbaren Handlungs- und Ereignisdruck realen Unterrichts Rechnung getragen werden kann.)

2. Begründen Sie lautdenkend, warum ihnen diese Situation aufgefallen ist.
3. Stellen Sie sich zu jeder Situation vor, Sie stünden neben den agierenden Lernenden, äußern Sie lautdenkend, wie Sie handeln würden.

Außerdem wurde für die weiteren Videos auf die spezielle Kameraführung verzichtet, weil sie zu mehr Irritationen geführt hat, als dass sie den Testpersonen geholfen hat, die Perspektive der Lehrkraft zu übernehmen.

5.1.2 Schritt 2 und 3 – Die Drehbücher

Das standardisierte PUPiL-Instrument basiert auf Videostimuli, welche zu validen Ergebnissen in der Erfassung der professionellen Unterrichtswahrnehmung führen sollen. Bevor allerdings valide scripted Videos vorliegen konnten, musste sichergestellt werden, dass die Drehbücher inhaltsvalide sind. Eine Validierung erst nach dem Filmdrehen hätte dazu führen können, dass großer zeitlicher, personeller und finanzieller Aufwand in einem nur kleinen Outcome geendet hätte, weil nicht sicher gestellt worden wäre, dass inhaltsvalide Videos produziert werden. Ein Risiko bei der Validierung vor der filmischen Umsetzung war hingegen, dass die Drehbücher nicht vollständig umgesetzt werden und dadurch neue kritische Situationen entstehen oder aber validierte, kritische Situationen nicht als solche erkennbar sein könnten. Ob dies der Fall war, wurde nach der Verfilmung der Drehbücher eingeschätzt, was im gegebenen Fall zur Überarbeitung der Videos führte. Eine Schwierigkeit bei der Entwicklung von Drehbüchern zur Erfassung der professionellen Unterrichtswahrnehmung ist es, alle charakteristischen Merkmal einer kritischen Situation ohne visuelle Unterstützung einzubinden. Außerdem ist bisher nicht klar, wie viele Indikatoren für eine kritische Situation in einem Video (oder Drehbuch) ausreichen, um die Aufmerksamkeit der Lehrkräfte zu erregen. Reicht es aus, wenn eine Schülervorstellung einmal genannt wird, oder müssen mehrere Aktionen im Unterricht auf diese hinweisen, bevor sie das erste Mal wahrgenommen werden kann? Um eine Idee diesbezüglich zu bekommen, geben die Drehbücher mehr als nur ein Indiz für (adäquate und inadäquate) Schülervorstellungen und Erkenntnisgewinnungsprozesse seitens der Lernenden. Das bedeutet, dass für die Drehbücher bemerkenswerte Unterrichtssituationen ausgewählt und nachträglich zu einer Unterrichtsstunde gebündelt wurden. Die Drehbücher zeigen auch bewusst weniger gelungenen Unterricht. Sie basieren alle entweder auf eigener Unterrichtserfahrung, eigenen Unterrichtsbeobachtungen oder wurden aus der physikdidaktischen Forschung zu Schülervorstellungen oder Schulbuchkapiteln abgeleitet. Dies sollte dazu führen, die geforderte erhöhte Kontextualisierung zu erfüllen. Dabei werden in den Drehbüchern auch eher ungewöhnliche und in der Literatur kaum oder gar nicht dokumentierte Schülervorstellungen thematisiert. Dadurch wird die Aufgabenschwierigkeit für die Testpersonen erhöht: Sie müssen nicht bloß die aus der Lehre bekannten Schülervorstellungen erkennen,

sondern ihre Erfahrungen und ihr professionelles Wissen nutzen, um lernrelevante Situationen – also auch unbekannte Schülervorstellungen – physikspezifisch wahrzunehmen. Jedem Drehbuch gingen daher Informationen über die Klasse, das Vorwissen, die Unterrichtseinheit und das Ziel der Stunde voraus. Die Beschreibung der Drehbücher beginnt mit diesen Informationen und anschließend werden kritische Situationen präsentiert.

Drehbuch 1 Das erste Drehbuch handelt von einer neunten Klasse einer Stadtteilschule (Gesamtschulen in Hamburg), die eine Unterrichtsstunde zum Lösen von (Farb- und Aromastoffen von) Tee in Wasser im Rahmen der Unterrichtseinheit „Wärmelehre“ von der Lehrkraft erhält²⁸. Mit einem Fokus auf das Teilchenmodell und die thermischen Eigenschaften von Teilchen (Lösen von Tee im Wasser) zählt dieser Themenbereich in das Basiskonzept „Materie“ (vgl. Hertel & Großmann, 2016, S. B4). Damit ist das Thema dieser Unterrichtsstunde nach dem Bildungsstandards in der neunten oder zehnten Klasse zu unterrichten. In dem Unterthema zum Lösen von Teilchen in Lösemitteln gibt es verschiedene Schülervorstellungen, welche auf unterschiedliche Problemstellungen deuten (vgl. Peetz, 2019, S. 48-50). Eine dieser Schülervorstellungen konnte in einer Unterrichtsbeobachtung im integrierten Schulpraktikum der Universität Hamburg beobachtet werden und erhält aufgrund von Authentizität Einzug in das Drehbuch. Die Schülervorstellung besteht darin, dass sich die Lernenden vorstellen, der zu lösende Stoff schmelze, wenn er in das Lösemittel gelange (ebd., S. 49). Teilweise bezieht sich diese Vorstellung nicht nur auf den Stoff, sondern auch auf die den Stoff konstruierenden Teilchen (ebd.).

Man erfährt zu Beginn des Videos, dass die Lernenden Vorwissen im Bereich des Energiebegriffs, des Teilchenmodells sowie im Formulieren und Testen von Hypothesen haben. Das Ziel der im Drehbuch skizzierten Unterrichtsstunde ist, dass die Lernenden in der Lage sind, Hypothesen in Bezug auf die Brown'sche Molekularbewegung von Teilchen selbst zu formulieren und in einem teils vorgegebenen Experiment zu testen. Daher erzählt die Lehrkraft als Einstieg eine (erfundene) Anekdote aus ihrem Leben: Sie wollte wegen ihres kaputten Wasserkochers Tee mit kaltem Wasser zubereiten und fragt ihre Lernenden, ob das möglich sei. Die Erarbeitung der Antwort erfolgt über eigenständiges Experimentieren, sodass die Lernenden das Aufstellen und Überprüfen von Hypothesen im Sinne der Erkenntnisgewinnung üben können. Das Experiment besteht darin, in je ein Wasserglas mit heißem und kaltem Wasser einen Teebeutel zu hängen und fünf Minuten zu beobachten, was mit dem Tee in den verschiedenen Gläsern geschieht. Dieses Experiment eignet sich dafür, die Wirkung der Wassertemperatur auf den Lösungsprozess zu untersuchen, weil alle weiteren Parameter (zum Beispiel die Ziehzeit des Tees oder die Wasserfüllhöhe) konstant gehalten werden können. Außerdem war das Experiment den Lernenden vorgegeben, sodass sie sich ausschließlich auf das Aufstellen und Prüfen der Hypothese konzentrieren können, ohne über das Experiment nachdenken zu müssen. Die Aufgaben für die Lernenden lauten:

²⁸Die Idee zu diesem Drehbuch entstammt Unterrichtsbeobachtungen im integrierten Schulpraktikum der Universität Hamburg.

1. Formuliert zu zweit eine Hypothese: Löst sich Tee besser in warmem oder in kaltem Wasser?
2. Prüft eure Hypothese mit einem geeigneten Versuch.
3. Überlegt euch Gründe für euer Ergebnis.

Zur Sicherung werden Stichpunkte der Lernenden an der Tafel gesammelt. Die Interpretation der Ergebnisse führt zur Erkenntnis über die molekulare Bewegung von Teilchen. Die beiden zentralen Situationstypen der Unterrichtsstunde sind (a) die beharrliche Schülervorstellung zweier zusammen arbeitender Lernender, dass Teeteilchen im Wasser schmelzen (in Anlehnung an Peetz, 2019), statt sich zu lösen, und (b) dass trotz Übung in der vorangegangenen Stunde nicht allen Lernenden klar ist, was eine Hypothese ist und welche Funktion sie hat. Es werden absichtlich sowohl gelingende als auch nicht gelingende Situationen zu diesen Typen beschrieben. Um möglichst nah am Kontext realen Unterrichts zu bleiben, sitzen circa 20 Lernende im Physikraum. Wirklich aktiv verfolgt das Drehbuch abwechselnd zwei Paare von Lernenden (mit erkennbar diversen kulturellen Hintergründen) und Plenumsphasen. Wie in allen anderen Drehbüchern auch wird versucht zu vermeiden, einen Fokus auf die Lehrkraft zu setzen, sodass die Lehrkraft im Wesentlichen nur die Einführung in das Thema übernimmt und die Lernenden im Anschluss eigenständig in Partnerarbeit experimentieren. Ziel der Untersuchung ist es nicht, dass die Testpersonen die Handlungen der Lehrkraft wie im advokatorischen Ansatz (Oser et al., 2010) bewerten, sondern das Lernen der Lernenden in Form von Handlungen und Äußerungen wahrnehmen. Wie im realen Unterricht wird auch in keinem der Drehbücher Unterricht skizziert, der durchweg den Kriterien guten Unterrichts (vgl. beispielsweise Meyer, 2003; Lipowsky, 2007) entspricht.

Folgende kritische Situationen enthält das Drehbuch zu den zwei Problemen:

- Situation 1: In einer Plenumsphase zum Generieren der Hypothesen vermutet Schülerin 1, dass der Tee im kalten Wasser langsamer schmilzt. (Betrifft inadäquate Schülervorstellung: „Tee schmilzt.“ statt „Tee löst sich.“)
- Situation 2: In einer Partnerarbeit werden Hypothesen als Aussagen ohne Indikatoren formuliert. (Betrifft inadäquate Erkenntnisgewinnung: Funktion von Hypothesen wird noch nicht verstanden.)
- Situation 3: Schülerin 3 vermutet im Partnergespräch, dass sich Tee im warmen Wasser besser löst, weil sich das heiße Wasser stärker bewegt. (Betrifft adäquate Schülervorstellung: „Warmes Wasser bewegt sich mehr, darum löst sich der Tee gut.“)
- Situation 4: In einer Partnerarbeit erklärt Schülerin 1 ihrem Sitznachbarn Schüler 2, dass man bei Hypothesen aufschreiben muss, was beim Experiment passiert ist. (Betrifft inadäquate Erkenntnisgewinnung: Funktion von Hypothesen wird noch nicht verstanden.)

- Situation 5: Schüler 2 nimmt die Vermutung seiner Sitznachbarin auf und äußert, dass er denkt, Teeteilchen schmelzen besser im warmen Wasser. (Betrifft inadäquate Schülervorstellung: „Tee schmilzt.“ statt „Tee löst sich.“)
- Situation 6: Zur Überprüfung der aufgestellten Vermutungen bräuchten die Lernenden vor dem Durchführen des Experiments weitere Materialien wie Stoppuhren und Thermometer. (Betrifft inadäquate Erkenntnisgewinnung: Eigenständiges Durchführen von Experimenten ist nicht reflektiert genug.)
- Situation 7: Der Schüler 2 wiederholt seine Vermutung kurz vor dem Experiment erneut, dass Teeteilchen im Wasser schmelzen. (Betrifft inadäquate Schülervorstellung: „Tee schmilzt.“ statt „Tee löst sich.“)
- Situation 8: Schüler 4 erklärt am Ende des Experiments, warum sich Tee im warmen Wasser besser löst als im kalten Wasser. (Betrifft adäquate Schülervorstellung: „Warmes Wasser bewegt sich mehr, darum löst sich der Tee gut.“)
- Situation 9: Nach Beenden des Experiments äußert Schülerin 1 ihre Beobachtungen und wertet diese aus. (Betrifft adäquate Erkenntnisgewinnung: Schritte der Erkenntnisgewinnung werden ordnungsgemäß durchgeführt.)
- Situation 10: Schülerin 1 und Schüler 2 bestätigen mit ihren Beobachtungen ihre falsche Hypothese und verallgemeinern ihre Beobachtungen ohne adäquate Begründung. (Betrifft inadäquate Erkenntnisgewinnung: Inadäquate Verallgemeinerung der Beobachtung.)

Drehbuch 2 Im zweiten Drehbuch geht es ebenfalls um eine neunte Klasse in einer Stadtteilschule in Hamburg. Die Unterrichtsreihe, die den Kontext für die gezeigte Stunde bietet, behandelt Magnetismus und ist damit im Basiskonzept der Wechselwirkungen zu verorten (Beschluss der Kultusministerkonferenz, 2004). Magnete (hier nur Permanentmagnete) werden als Dipol (gleichnamige Pole stoßen sich ab, ungleichnamige Pole ziehen sich an) mit einem Kraftfeld zwischen diesen Polen elementarisiert (Nienhaus, 2018). Das Kraftfeld wird durch Feldlinien visualisiert. Modellhaft verlaufen außerhalb und innerhalb des Magneten Magnetfeldlinien von Nord- nach Südpol. Innerhalb des Magneten ist das Magnetfeld und die Feldliniendichte viel höher als außerhalb.

Aus dem vorherigen Unterricht können die Lernenden auf Vorwissen bezüglich elektrischer Anziehung und Abstoßung sowie magnetischer Anziehung und Abstoßung einfacher Stabmagneten bauen und kennen daher diese Elementarisierung. Am Ende der beschriebenen Stunde sollen die Lernenden das Magnetfeld zweier benachbarter Hufeisenmagneten beschreiben können. Dazu sollen Hypothesen aufgestellt und mittels angeleiteten Experiments verifiziert werden. Dieses Experiment besteht darin, zwei Hufeisenmagneten anzunähern und mithilfe von Eisenspäne die Magnetfeldlinien sichtbar zu machen. Dabei können die Lernenden das generelle Muster ihrer Magnetfeldlinien mit den durch die Eisenspäne angezeigten Feldlinien vergleichen, die Richtung der Magnetfeldlinien können

sie daraus allerdings nicht ableiten. Stattdessen hätten die Lernenden einen Aufbau mit vielen (farblich markierten) Kompassnadeln benötigt. Um die Stunde phänomenologisch einzuleiten, wird eine magnetisierte Nadel, die in einem Behältnis mit Wasser treibt, durch einen angenäherten Magneten von der Lehrkraft in Bewegung versetzt. In der Unterrichtsstunde geht es also mehr um die Prozesse der Erkenntnisgewinnung wie das Aufstellen und Testen einer Hypothese als um die Schülervorstellungen zum Magnetismus. Nichtsdestotrotz wurden in den vergangenen Jahren Schülervorstellungen über Kraftfelder im Allgemeinen und spezifisch für Magnetfelder erforscht (siehe dazu Duit et al., 2018, S. 187-195).

Zur Erarbeitung des Stundenziels bearbeiten die Lernenden folgende Aufgabe:

Stelle eine Hypothese auf, wie das Magnetfeld zweier benachbarter Hufeisenmagneten aussieht. Arbeitet zu zweit und testet eure Hypothese mit dem angegebenen Experiment (siehe im Anhang Abschnitt A.1).

Als Sicherung werden die Hypothesen und Ergebnisse an der Tafel gesammelt und im Plenum diskutiert.

Zentrales Problem in dieser Unterrichtsstunde ist, dass die Lernpaare unvollständige Hypothesen aufstellen, dabei aber fachlich angemessene Überlegungen äußern. Im Drehbuch arbeiten zwei Lernpaare am gleichen Experiment, die Lehrkraft und ihr Verhalten sind nur am Anfang der Arbeitsphase beschrieben. Auch in der hier beschriebenen Unterrichtsstunde soll die Lehrkraft möglichst wenig im Fokus stehen, um dem advokatorischen Ansatz zu umgehen. Daher arbeiten die Lernenden in Partnerarbeit und ohne weitere Intervention der Lehrkraft.

Die folgenden kritischen Situationen müssten von Lehrkräften wahrgenommen werden, um Lernende darin zu unterstützen, das Konzept der Feldlinien bei Stabmagneten auf das Konzept von Feldlinien bei benachbarten Hufeisenmagneten zu übertragen:

- Situation 1: Schülerin 1 beschreibt in der Partnerarbeit das Magnetfeld der beiden Hufeisenmagneten in Bezug auf das bekannte Feld eines Stabmagneten. (Betrifft angemessene Schülervorstellung: Vorstellungen über Magnetfeldlinien.)
- Situation 2: Schülerin 1 stellt eine Hypothese auf, indem sie eine Analogie zum Stabmagneten bildet. (Betrifft Erkenntnisgewinnung: Phänomene werden mit Analogien beschrieben, aber es resultiert keine vollständige Hypothese.)
- Situation 3: Schüler 2 benutzt in einem anderen Partnergespräch eine ihm bekannte Analogie (Wassertropfen), um sich das resultierende Feld zweier sich überlagernder Felder vorzustellen. (Betrifft adäquate Schülervorstellung: Die Überlagerung von Magnetfeldern ist analog zu der von Wellen.)
- Situation 4: Schüler 2 wählt zur Generierung einer Hypothese ein Gedankenexperiment. (Betrifft Erkenntnisgewinnung: Zur Generierung einer Hypothese wird bereits vorhandenes Wissen genutzt, es führt aber nicht zu einer überprüfbaren Hypothese.)

- Situation 5: Schüler 2 stellt eine unvollständige Hypothese auf. (Betrifft Erkenntnisgewinnung: Zur Generierung einer Hypothese wird bereits vorhandenes Wissen genutzt, es führt aber nicht zu einer überprüfbaren Hypothese.)
- Situation 6: Schülerin 1 und Schüler 3 führen ein Experiment durch und erhalten so Informationen über die Güte ihrer Hypothese. (Betrifft Erkenntnisgewinnung: Das Ergebnis des Experiments führt zu einer Anpassung der Hypothese.)
- Situation 7: Schülerin 1 verallgemeinert nach dem Experiment ihre Ergebnisse ohne weitere Überlegungen oder Experimente und lässt eine Teilaufgabe aus. (Betrifft inadäquate Erkenntnisgewinnung: Inadäquate Verallgemeinerung der Beobachtung.)

Drehbuch 3 Das dritte Drehbuch präsentiert eine neunte Klasse eines Gymnasiums aus Hamburg. Die Unterrichtsreihe behandelt die Elektrostatik mit Fokus auf die Einführung elektrischer Ladung, sodass die beschriebene Stunde im Basiskonzept Materie zu verorten ist²⁹. Elektrische Ladung wird als Eigenschaft von Materie elementarisiert, die sich im Aufbau der Atome zeigt (vgl. Nienhaus, 2018, S. 14). In der Mittelstufe wird die elektrische Ladung durch die Betrachtung von positiv, negativ und neutral geladenen Teilchen erklärt (vgl. beispielsweise Barmeier, 2013). Zwischen geladenen Teilchen wirkt ein elektrisches Feld, welches mit Feldlinien hinzu einer positiven Ladung visualisiert wird. Je dichter diese Feldlinien sind, desto höher ist die Stärke des Feldes (vgl. Nienhaus, 2018, S. 22). Die Lernenden haben zu der Thematik kein schulisch vermitteltes Vorwissen. Es gibt einige wissenschaftliche Erkenntnisse über Schülervorstellungen zu elektrischen Feldern und Stromkreisen (Duit et al., 2018), aber bisher sind Schülervorstellungen zur Elektrostatik nicht systematisch untersucht worden. In Hospitationen im integrierten Schulpraktikum der Universität Hamburg konnte aber die Schülervorstellung, dass elektrische Anziehung mit einem Klebstoff zu tun haben müsse, beobachtet werden. Diese Vorstellung bietet einen Ausgangspunkt für einen Konzeptwechsel hinzu einem Konzept über elektrische geladene Körpern und das elektrische Feld zwischen diesen. Daher bildet diese Vorstellung die Grundlage für das Drehbuch dieser Unterrichtsstunde. In diesem Drehbuch sollten (angehende) Lehrkräfte wahrnehmen, dass Lernende (a) Prozesse der Erkenntnisgewinnung überwiegend angemessen anwenden, (b) elektrostatische Anziehung als Kleben interpretieren, (c) Blitze mit Kälte in Verbindung bringen und (d) Vorwissen aus Fernsehsendungen miteinbeziehen.

Ziel der Stunde ist es, dass die Lernenden das Phänomen der elektrischen Anziehung und Abstoßung beschreiben können. Sie sollen dazu Hypothesen aufstellen und diese mittels angeleiteten Experiments verifizieren. Das Experiment besteht darin, zwei aneinander geriebene Plastikstreifen und anschließend wieder zu trennen. Das Verhalten der geladenen Streifen soll mit eine Hypothese vorausgesagt werden, nachdem die Lernenden zunächst phänomenologisch und anschließend über einen Text an ein Ladungskonzept herangeführt

²⁹Die Idee für diese Unterrichtsstunde entstammt aus Einheiten aus einem Physikschulbuch (vgl. Barmeier, 2013) und eigenen Erfahrungen im Unterrichten von Elektrostatik.

worden.. Als Einstieg zeigt die Lehrkraft einen Versuch, bei dem Funken durch Reibung erzeugt werden. Dieser Versuch soll das Interesse der Lernenden wecken und ihnen Anhaltspunkte für das Aufstellen der Hypothesen geben. Die Lernenden erarbeiten sich das angestrebte Wissen mit der Bearbeitung folgender Aufgabe:

1. Stelle eine Hypothese auf, wie geladene Objekte auf andere geladene Objekte reagieren.
2. Arbeitet zu zweit und testet eure Hypothese mit dem angegebenen Experiment (siehe im Anhang Abschnitt A.2).

Ergebnisse des Experimentierens werden zur Sicherung an der Tafel festgehalten und diskutiert. Trotz wiederholter Plenumsphasen, bleibt die Lehrkraft im Hintergrund, sodass dieses Unterrichtsstunde die Aufmerksamkeit der Lehramtsstudierenden auf die geäußerten Konzepte und Prozesse der Erkenntnisgewinnung der Lernenden gelenkt werden kann. Bei der beschriebenen Klasse handelt es sich um eine Klasse von circa 30 Lernenden, von denen durch das Drehbuch drei Lernpaare und Plenumsphasen fokussiert werden. Auch in diesem Drehbuch wird die Lehrkraft nur an nötigen Stellen der Unterrichtsführung gezeigt.

Folgende Vorkommnisse sollten von Lehrkräften während des Unterrichtens zur Aufrechterhaltung von Lernprozessen wahrgenommen werden:

- Situation 1: In einer einführenden Plenumsphase vergleicht Schüler 1 das beobachtete Phänomen mit Phänomenen, die er aus dem Alltag kennt. (Betrifft adäquate Erkenntnisgewinnung: Angemessenes Beschreiben von Phänomenen mit Hilfe von Analogien.)
- Situation 2: Die elektrostatische Anziehung zwischen einem Luftballon und elektrisierten Haaren wird von den Lernenden mit „Kleben“ bezeichnet. (Betrifft inadäquate Schülervorstellung: Elektrostatische Anziehung wird als „Kleben“ verstanden.)
- Situation 3: Die elektrostatische Anziehung zwischen einem Luftballon und elektrisierten Haaren wird von Schülerin 2 auf Nachfrage der Lehrkraft ebenfalls mit „Kleben“ bezeichnet. (Betrifft inadäquate Schülervorstellung: Elektrostatische Anziehung wird als „Kleben“ verstanden.)
- Situation 4: Schüler 3 erklärt im Plenum das Phänomen der elektrischen Anziehung mithilfe von Informationen, die er aus dem Fernsehen hat. (Betrifft adäquate Erkenntnisgewinnung: Informationen werden aus Quellen ausgewählt und bewertet.)
- Situation 5: Schüler 1 beschreibt im Plenum seine Beobachtung zu einem Experiment. (Betrifft adäquate Erkenntnisgewinnung: Adäquate Beschreibung von Beobachtungen.)
- Situation 6: Schüler 4 ergänzt die Beobachtungen von Schüler 1. (Betrifft adäquate Erkenntnisgewinnung: Adäquate Beschreibung von Beobachtungen.)

- Situation 7: In einer Partnerarbeit vergleicht Schüler 4 das Experiment mit Gewittern. (Betrifft adäquate Erkenntnisgewinnung: Angemessenes Beschreiben von Phänomenen mit Hilfe von Analogien.)
- Situation 8: Schülerin 2 vergleicht das Experiment in einem Partnergespräch mit einem Phänomen aus ihrem Alltag. (Betrifft adäquate Erkenntnisgewinnung: Angemessenes Beschreiben von Phänomenen mit Hilfe von Analogien.)
- Situation 9: Schüler 5 äußert in einem Partnergespräch die Vorstellung, dass Blitze etwas mit Kälte zu tun haben. (Betrifft inadäquate Schülervorstellung: Blitze entstehen bei Kälte.)
- Situation 10: Schüler 3 nimmt in seinem Partnergespräch wieder Bezug auf sein Wissen aus dem Fernsehen. (Betrifft adäquate Erkenntnisgewinnung: Informationen werden aus Quellen ausgewählt und bewertet.)
- Situation 11: Schüler 5 erklärt das Phänomen des Experiments im Plenum mit seinem (nicht angemessenen) Erfahrungswissen. (Betrifft inadäquate Schülervorstellung: Blitze entstehen bei Kälte.)
- Situation 12: Schüler 3 korrigiert Schüler 5 im Plenum und äußert sein Wissen zum Zusammenhang von Blitzen und Ladungen. (Betrifft adäquate Schülervorstellung: Blitze stehen in einem Zusammenhang mit Ladung.)
- Situation 13: Schülerin 2 und Schüler 5 wissen noch nicht, warum die Streifen in ihrem Schülerversuch verschieden geladen sind, lassen das aber als Fakt in ihre Hypothese einfließen. (Betrifft inadäquate Erkenntnisgewinnung: Funktion von Hypothese wird noch nicht erkannt.)
- Situation 14: Schüler 3 und sein Sitznachbar Schüler 5 entwickeln aus eigenem Interesse eine neue Hypothese und planen dazu ein Experiment. (Betrifft adäquate Erkenntnisgewinnung: Experimente werden zum Testen von Hypothesen eigenständig geplant.)
- Situation 15: Die Schüler 3 und 5 führen ihr eigenständig geplantes Experiment durch, machen dabei aber einen Fehler. (Betrifft Erkenntnisgewinnung: Eigenständiges Durchführen von Experimenten, allerdings unkorrektes Handeln.)

Drehbuch 4 Das vierte Drehbuch thematisiert eine neunte Klasse eines Hamburger Gymnasiums in der Unterrichtsreihe Energie. Dieses Thema ist zugleich ein Basiskonzept und soll in Klasse 9 und 10 unterrichtet werden (Hertel & Großmann, 2016). Energie wird häufig als die Möglichkeit eines Systems, Arbeit zu verrichten, elementarisiert und in verschiedenen Energieformen (potentielle Energie, kinetische Energie und Wärme) repräsentiert (Nienhaus, 2018). Bekannte Schülervorstellungen zu Energie sind, dass (a) Energie benötigt wird, um etwas zu bewirken, (b) Energie ein speicherbarer Treibstoff ist

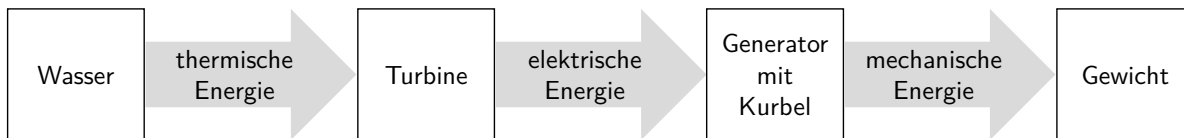


Abbildung 5.1: Eine mögliche Energieumwandlungskette für die Aufgabe der Lernenden in dem Video.

oder (c) dass Energie durch materielle Träger transportiert wird (vgl. Duit et al., 2018, S. 164-166). Außerdem stellen sich einige Lernende vor, dass Energie verbraucht wird, verloren geht und nur unter idealen Bedingungen erhalten bleibt (vgl. Duit et al., 2018, S. 166-171).

Die hier skizzierte Stunde steht in der Mitte einer Reihe zum physikalischen Energiebegriff. In den vorherigen Doppelstunden wurde der Energiebegriff eingeführt, verschiedene Energieträger und die vorkommenden Energieformen (kinetische Energie, potentielle Energie, elektrische Energie, Strahlungsenergie, chemische Energie und Wärmeenergie) wurden aus dem Alltag und aus physikalisch-technischen Bereichen abgeleitet und mathematisch beschrieben. Mit den Energieträgern wurde die Energiequadrige begonnen. Die Lernenden haben gelernt, dass Energie zwar nicht stofflich ist, aber beispielsweise in Nahrung und Brennstoffen gespeichert werden kann. Von dort kann sie über Umwandlung in verschiedene Formen überführt werden.

In der hier skizzierten Unterrichtsstunde sollen die Lernenden die weiteren drei Bestandteile der Energiequadrige (Energieübertragung, Energieerhaltung und Energieentwertung) kennenlernen und unterscheiden können. Als Anlass dient ein springender Gummiball, der ohne äußere Einwirkung nie seine Ausgangshöhe erreicht. Ziel ist es, dass die Lernenden Phänomene wie dieses durch das Aufstellen von Energieumwandlungsketten erklären können. Das Stundenziel lautet daher: Die Lernenden sollen durch das Erarbeiten von Inhalten eines Textes zur Energieumwandlung und das Aufstellen von Energieumwandlungsketten erklären können, dass Energie erhalten bleibt, aber bei der Energieumwandlung durch irreversible Prozesse entwertet werden kann. Dazu analysieren die Lernenden verschiedene Energieumwandlungen mit Hilfe von Energieumwandlungsketten. Das sind grafische Darstellungen, die bei der Energie als Treibstoff ansetzen und damit eine „tragfähige Basis zum Verständnis der Idee der Energie-Erhaltung“ (Duit, 1986) bieten. Des Weiteren helfen visuelle Darstellungen den Lernenden, die komplexen Prozesse der Energieumwandlung vereinfacht darzustellen und zu erklären. Diese Darstellungen bestehen aus Energiespeichern (z.B. die Sonne) oder Energiewandlern (z.B. ein Motor), verbunden durch die verschiedenen Energieformen. Ein Beispiel für eine Energieumwandlungskette aus dem Drehbuch ist in der Abbildung 5.1 abgebildet. Die Aufgaben für die Lernenden lauten:

1. Bild 1 zeigt ein Modell eines Wärmekraftwerks. Erstelle die Energieumwandlungskette eines Wärmekraftwerks (siehe im Anhang Abschnitt A.3).

2. Beim Inlineskaten auf der Half-Pipe wandeln sich Höhen- und Bewegungsenergie ineinander um (siehe Bild 2 im Anhang Abschnitt A.3). Würde sich der Skatende am oberen Rand nicht immer leicht abstoßen, würde er nicht mehr ganz den oberen Rand erreichen. Erkläre diesen Sachverhalt.

Die Erarbeitung erfolgt über Partnergespräche, damit die Lernenden ihre Überlegungen versprachlichen können und durch die Ko-Konstruktion lernen. Vermutungen über Energieerhaltung und Energieumwandlung werden außerdem vor und nach der Bearbeitung der Aufgaben in der Klasse diskutiert. Es wurden für diese Unterrichtsstunde auf Experimente verzichtet, damit die Lernenden ihren Fokus auf das konzeptuelle Lernen richten können, ohne parallel Anforderungen durch das Planen, Durchführen und Auswerten eines Experiments erfüllen zu müssen.

Zentral im gesamten Drehbuch ist die Problematik, dass die Lernenden zunächst von der Möglichkeit eines Energieverlustes ausgehen, später – auf Anweisung der Lehrkraft hin – Energieumwandlung sagen, aber immer noch Energieverlust meinen. Dieses Problem muss eine Lehrkraft während des Unterrichtens wahrnehmen, um das Unterrichtsziel zu erreichen. Die beschriebene Klasse zeigt sechs aktiv Lernende, die sowohl im Plenum als auch in Partnerarbeit fokussiert werden. Die Handlungen und Äußerungen der Lehrkraft werden so selten wie möglich beschrieben, stattdessen werden die Gespräche in der Partnerarbeit fokussiert.

In 14 kritischen Situationen werden Schülervorstellungen zu Energie und deren Umwandlung offenkundig:

- Situation 1: Schüler 1 erklärt im Plenum sein Verständnis des Begriffs Energieumwandlung mit der Umwandlung von Energie. (Betrifft Schülervorstellungen: Redundante Erklärung von Energieumwandlung.)
- Situation 2: Schüler 2 nennt ein Beispiel der Energieumwandlung im Plenum. (Betrifft adäquate Schülervorstellung: Das Beispiel lässt auf ein fachlich angemessenes Verständnis von Energieumwandlung schließen.)
- Situation 3: Schülerin 3 nimmt an, dass Energie überall im Alltag verbraucht wird. (Betrifft inadäquate Schülervorstellung: Energie wird verbraucht, nicht umgewandelt.)
- Situation 4: Schülerin 4 stellt sich im Plenumsgespräch vor, dass Energie auf Objekte übertragen werden kann. (Betrifft inadäquate Schülervorstellung: Energie ist etwas Stoffliches, das übergeben werden kann.)
- Situation 5: Schüler 5 hat schon in der ersten Plenumsphase die Vorstellung, dass Energie umgewandelt wird. (Betrifft adäquate Schülervorstellung: Energie wird umgewandelt.)
- Situation 6: Schülerin 6 fasst nach dem Lesen des Textes zusammen, dass Energie nach der Umwandlung manchmal weg ist. (Betrifft Schülervorstellung: Ein Übergang von der Verbrauchs-Vorstellung zur Umwandlungs-Vorstellung ist zu erahnen.)

- Situation 7: Schüler 1 macht durch die Arbeit mit dem Text und die Plenumsdiskussion die Erfahrung, dass sein bisheriges Konzept nicht ausreicht. (Betrifft Schülervorstellung: Ein Übergang von der Verbrauchs-Vorstellung zur Umwandlungs-Vorstellung ist zu erahnen.)
- Situation 8: Schülerin 4 fragt sich bei der Bearbeitung der ersten Aufgabe, ob die Energie am Ende der Kette verloren ist. (Betrifft inadäquate Schülervorstellung: Energie geht verloren, wird nicht umgewandelt.)
- Situation 9: Schülerin 3 stellt bei der Bearbeitung von Aufgabe 1 fachlich richtig klar, dass Energie nicht verbraucht, sondern umgewandelt wird. (Betrifft adäquate Schülervorstellung: Energie wird umgewandelt.)
- Situation 10: Schüler 5 erweitert im Partnergespräch seine Vorstellung durch das Konzept nicht nutzbarer Energie. (Betrifft adäquate Schülervorstellung: Energie bleibt zwar erhalten, kann aber nicht zwingend direkt in nutzbare Energie überführt werden.)
- Situation 11: Schüler 5 stellt sich vor, dass Energie in Form von Kalorien verbraucht wird. (Betrifft inadäquate Schülervorstellung: Zwar können Kalorien als Energie verstanden werden, aber der Verbrauch dieser, lässt wieder auf eine physikalisch unangemessene Vorstellung schließen.)
- Situation 12: Schülerin 6 argumentiert in der Diskussion mit ihrem Sitznachbarn um nicht nutzbare Energie mit dem Kraftbegriff. (Betrifft inadäquate Schülervorstellung: Kraft und Energie werden als das gleiche Konzept verstanden.)
- Situation 13: Schüler 2 kann den Verbrauch und die Umwandlung von Energie am Ende der Stunde immer noch nicht konzeptionell voneinander trennen. (Betrifft inadäquate Schülervorstellung: Energie wird verbraucht, nicht umgewandelt.)
- Situation 14: Schüler 1 meint, dass Reibungskraft Wärme ist. (Betrifft Schülervorstellung: Kraft und Energie werden als das gleiche Konzept verstanden. Aber Wärme wird schon als (nicht nutzbare) Energie verstanden.)

Drehbuch 5 Drehbuch 5 beschreibt eine neunte Klasse eines Hamburger Gymnasiums in der Unterrichtsreihe zu den Newton'schen Axiomen. Kräfte werden als Ursache für Bewegungsänderungen³⁰ elementarisiert (vgl. Nienhaus, 2018, S. 47). In den drei Newton'schen Axiomen werden mechanische Bewegungen und die Wirkung von Kräften mathematisch beschrieben. Bekannte Schülervorstellungen betreffen das Konzept der Kraft an sich, aber auch alle drei Newton'schen Axiome (vgl. Duit et al., 2018, S. 70-81). In der durch dieses Drehbuch beschriebenen Unterrichtsstunde sollen die Lernende anhand von Tauziehen auf

³⁰Zur Beschreibung der Bewegung wird der Impuls p als Produkt der Geschwindigkeit und Masse verwendet. Wirkt eine Kraft, ändert sich der Betrag oder die Richtung des Impulses (vgl. Nienhaus, 2018, S. 48).

Skateboards erklären können, dass jede Kraft eine Gegenkraft hervorruft, die vom Betrag gleich groß, aber entgegengesetzt gerichtet ist (drittes Newton'sches Axiom). Zum dritten Newton'schen Axiom gibt es die Schülervorstellung, dass (a) Kraft und Gegenkraft am gleichen Körper angreifen, (b) erst die Kraft und dann die Gegenkraft auftritt und (c) dass Körper als Widerstand eine Gegenkraft entwickeln (vgl. Duit et al., 2018, S. 77, 78). Die Lernenden in dieser skizzierten Unterrichtsstunde haben bereits Vorwissen zu Kräften als Ursache für Beschleunigung, der Wirkung von Kräften, Kraftmessern, Kräfteaddition und dem Hooke'schen Gesetz. Als Einstieg in die Thematik des dritten Newton'schen Axiom wählt die Lehrkraft einen Grundschulversuch „Die Luftballon-Rakete“ als Wiederholung des subjektiven Wissens durch das Lernen durch Eigenerfahrung. Für das Erreichen des Unterrichtsziels erhalten die Lernenden folgende Aufgabe:

Spielt zu zweit Tauziehen auf Skateboards. Beobachtet und erklärt, wer auf wen eine Kraft ausübt.

Dieser Versuch dient dazu, dass die Lernenden durch eine handelnde Auseinandersetzung das Phänomen beobachten und sich Ideen über das Phänomen schaffen können.

Am Stundenende werden die Ergebnisse in Stichpunkten an der Tafel gesammelt.

Eine Lehrkraft muss in dieser Unterrichtsstunde wahrnehmen, dass das Vorwissen der Lernenden fundiert ist, es ihnen aber schwerfällt, das dritte Newton'sche Axiom zu verstehen. Dass eine Gegenkraft unabhängig von der subjektiv empfundenen Stärke eines Objektes ist, scheint von den Lernenden nicht verarbeitet zu sein. Wie in allen Drehbüchern wird die Lehrkraft auch in diesem Drehbuch nur an notwendigen Stellen beschrieben, wie beispielsweise bei der Anleitung zur Aufgabe.

Folgende kritische Situationen müssen von Lehrkräften bei der Beobachtung des Unterrichts wahrgenommen werden, um die Lernenden beim Verstehen des dritten Newton'schen Axioms zu unterstützen:

- Situation 1: Schülerin 1 vergleicht nach dem Einstiegs-Versuch die Luftballon-Rakete mit einer Flugzeug-Turbine. (Betrifft adäquate Erkenntnisgewinnung: Zur Beschreibung eines Phänomens wird eine Analogie hinzugezogen.)
- Situation 2: Schüler 2 äußert im Plenum die Vorstellung, dass es beim Versuch mit der Luftballon-Rakete keine Gegenkraft geben kann, weil nur ein Objekt zu sehen ist. (Betrifft inadäquate Schülervorstellung: Das Konzept von Gegenkraft wird nur auf Phänomene mit mehreren sichtbaren Objekten bezogen.)
- Situation 3: Schülerin 3 gibt anhand eines Beispiels (Startblock) eine fachlich angemessene Vorstellung von Gegenkraft im Plenum wieder. (Betrifft adäquate Schülervorstellung: Wenn ein Körper auf einen anderen eine Kraft ausübt, gibt es eine Gegenkraft.)
- Situation 4: Schüler 4 wiederholt im Plenum den Zusammenhang Kraft und Beschleunigung fachlich angemessen, um zu beschreiben, was er beim Tauziehen

beobachtet hat. (Betrifft adäquate Schülervorstellung: Kräfte bewirken Beschleunigung.)

Situation 5: Schüler 2 äußert bei der Beobachtung des Tauziehens seiner anderen Gruppenmitglieder, dass der eine stärker ist und der andere darum weiter gerollt ist. (Betrifft inadäquate Schülervorstellung: Die Gegenkraft ist nicht unabhängig von der subjektiv empfundenen Stärke des Mitschülers.)

Situation 6: Schüler 5 versteht nicht, warum beim Tauziehen auf Skateboards beide rollen, auch wenn nur einer aktiv zieht. (Betrifft inadäquate Schülervorstellung: Eine Gegenkraft muss aktiv aufgebracht werden.)

Diese fünf Drehbücher wurden von sechs physikdidaktisch forschenden Personen in Hinblick auf ihre inhaltliche Validität (bezüglich des Konzepts der professionellen Unterrichtswahrnehmung) und von Physiklehrkräften bezüglich der curricularen und ökologischen Validität beurteilt. In den nächsten beiden Abschnitten werden die dafür gebrauchten Instrumente (Kodierschema und Fragebogen) vorgestellt.

5.1.3 Schritt 4 – Das Kodierschema und die inhaltliche Validierung

Um sicherzustellen, dass die Drehbücher und somit die Videos kritische Situationen aufweisen, in denen die professionelle Unterrichtswahrnehmung von Lehrkräften notwendig ist, wurden drei Analysen durchgeführt: Zuerst markierten sechs physikdidaktisch Tätige in MaxQDA diejenigen kritischen Situationen, die sie nach der Definition der professionellen Unterrichtswahrnehmung für kritisch hielten. Dabei sollten sie unterscheiden, ob es sich in der kritischen Situation primär um lernrelevante Schülervorstellungen oder Erkenntnisgewinnung handelt. Waren sich mindestens fünf der sechs Rater_innen einig, wurde die intendiert kritische Situation als inhaltlich valide verstanden und überprüft, ob deren Einschätzung deckungsgleich mit den Intentionen in den Drehbüchern waren. Nur diejenigen kritischen Situationen aus den Drehbüchern wurden in den Test mitaufgenommen, die diesen drei Analysen standhielten. Damit sollte die inhaltliche Validität der Drehbücher gesichert werden. Für den Zweck dieser Studie ist es dabei relevant, dass kritische Situationen mit mindestens einem der Aspekte Erkenntnisgewinnung und Schülervorstellungen die professionelle Unterrichtswahrnehmung von Lehrkräften bedürfen, was anhand des Ratings festgestellt werden sollte. Nach diesem Kriterium wurden die kritischen Situationen in den Drehbüchern ausgewählt, die anschließend Teil des Instruments werden sollten. Die Überprüfung der curricularen und ökologischen Validierung folgt im Anschluss an die Inhaltsvalidierung.

Für die Beurteilung der Inhaltsvalidität (vgl. für die Betrachtung der Inhaltsvalidität Schmiemann & Lücken, 2014) muss beantwortet werden, inwieweit die verwendeten Items des PUPiL-Instruments das theoretische Konstrukt (der professionellen Unterrichtswahrnehmung) abbilden. Dies geschah in diesem Fall über Plausibilitätsüberlegungen der physikdidaktisch Tätigen. Dank der konzentrierten Operationalisierung der professionellen

Unterrichtswahrnehmung ist überprüfbar, ob die Items alle relevanten Inhaltsbereiche des operationalisierten Konstrukts repräsentieren. In diesem Fall sind diese Inhaltsbereiche Schülervorstellungen und Erkenntnisgewinnung (in den Themenbereichen Elektrostatik, Energieumwandlung und Teilchenmodell). Für jedes Drehbuch ordneten die Rater_innen den kritischen Situationen Kategorien des Kodierschemas (siehe Tabelle 5.2 bis Tabelle 5.6) zu. Die Operationalisierung der professionellen Unterrichtswahrnehmung wurde zu dem Zeitpunkt folgendermaßen in eine Aufgabenstellung zur Kodierung übersetzt: „Eine Situation ist dann Noticing-würdig³¹, wenn Du sie wahrnimmst und eine Vorstellung oder Prozessfähigkeit der Erkenntnisgewinnung relevant für fachliches Lernen werden könnte.“

Die Rater_innen erhielten für die Kodierung folgende Hinweise:

- Jede Lernsituation darf mit jedem Code codiert werden.
- In einer Aussage können mehrere Aspekte vorkommen, alle dürfen codiert werden.
- Es soll für jedes einzelne Drehbuch zunächst für alle Codes codiert werden, bevor das nächste Drehbuch codiert wird.
- Für die Codes „2. Lernrelevante Situation“, „2.2 Bezüglich Vorstellungen“, „2.3 Bezüglich Prozesse der Erkenntnisgewinnung“ sollen nur die Subcodes verwendet werden.

In Tabelle 5.2 bis 5.6 sind die Codes für die Kodierung benannt (linke Spalte), in Hinblick auf die professionelle Unterrichtswahrnehmung definiert (mittlere Spalte) und mit einem Ankerbeispiel versehen (rechte Spalte).

Die Codes sind dabei folgendermaßen aufgebaut und unterteilt: Code 1 wird vergeben, wenn eine Situation nach der Operationalisierung der professionellen Unterrichtswahrnehmung kritisch ist. Code 2 unterscheidet zum einen, ob Vorstellungen (bezüglich Begriffen, Konzepten und naturwissenschaftlichen Arbeitsweisen) in der kritischen Situation lernrelevant werden (Code 2.1 und 2.2) oder ob Erkenntnisgewinnungsprozesse (orientiert an den Hamburger Bildungsstandards für den Physikunterricht in der Sekundarstufe 1 (vgl. Beschluss der Kultusministerkonferenz, 2004, S. 11)) als lernrelevant wahrgenommen werden müssen (Code 2.3 und Subcodes).

Die Ankerbeispiele dienen den Rater_innen als Orientierung zum Einordnen der Textstellen in den Drehbüchern in die jeweiligen Codes. Die Beispiele entstammen früheren Drehbuchideen.

Für die Auswertung wurde die Kodierung der physikdidaktisch Tätigen bezüglich jeder der intendiert kritischen Situation betrachtet. Die Daten wurden in eine Tabelle überführt, in der pro Drehbuch die Codes in den Zeilen und in den Spalten die intendiert kritischen Situationen dargestellt werden. Die Zellen wurden dann mit der Anzahl der Rater_innen gefüllt, die den entsprechenden Code für die jeweilige kritische Situation vergeben haben.

³¹Im frühen Stadium dieser Arbeit wurde die professionelle Unterrichtswahrnehmung noch mit „Noticing“ bezeichnet.

Tabelle 5.2: Kodierschema für die Beurteilung der Inhaltsvalidität – Teil 1.

Titel	Code-Definition	Ankerbeispiel
1. Verdient Aufmerksamkeit der Lehrkraft	Die Situation verdient die Aufmerksamkeit der Lehrkraft, weil es um fachliche Lernprozesse von Lernenden geht. <i>"Aufmerksamkeit verdienen" bedeutet hier nicht, dass eine sofortige Reaktion notwendig ist.</i>	Die Lernenden sollen ein Weg-Zeit-Diagramm zu ihrem Schulweg anfertigen. Max fragt Lena, wozu sie diese Aufgaben machen sollen. „Lena: Ich glaube die Form von den Linien zeigt, ob man anhalten muss oder nicht.“
2. Lernrelevante Situation	<i>In diesen Situationen können Maßnahmen zur Förderung von Fähigkeiten mehr oder weniger erfolgreich sein.</i>	
2.1 Für einzelne Lernende / Paare / Gruppen	Die Situation könnte für einzelne Lernende, Paare oder Gruppen von Lernenden relevant für das fachliche Lernen sein. <i>Kritik am Handeln der Lehrkraft fällt nicht in diesen Code.</i>	Die Lernenden sollen ein Weg-Zeit-Diagramm zu ihrem Schulweg anfertigen. Tim hat eine Frage bezüglich der Gestaltung des Koordinatensystems „Tim: Hallo?? Wie viele Striche machst du?“ Sven: Och man, weiß ich doch nicht!“

Tabelle 5.3: Kodierschema für die Beurteilung der Inhaltsvalidität – Teil 2.

Titel	Code-Definition	Ankerbeispiel
2.2 Bezüglich Vorstellungen		
2.2.1 Vorstellungen bzgl. physikalischer Konzepte und Begriffe	– Die Situation könnte für Lernende lernrelevant in Bezug auf Vorstellungen von physikalischen Konzepten und Begriffen sein. – Vorstellungen sind kognitiv und betreffen sowohl fachliche anerkannte Begriffe und Konzepte, als auch Schülervorstellungen, die von Ersteren abweichen oder zu ihnen im Widerspruch stehen können. – Schülervorstellungen werden als kognitive Konstrukte verstanden, die auf bisherige – nicht zwingend schulische – (Lern-)Erfahrungen zurückgehen. Sie sind dem Bewusstsein des Lernenden zugänglich und beziehen sich auf spezifische Themenfelder des Faches, wie zum Beispiel die Mechanik, den Sehvorgang oder chemische Reaktionen. (nach Krüger 2016)	Nele zeichnet das Weg-Zeit-Diagramm ihres Schulweges, dabei entstehen einige waagerechte Linien. „Caren: Das ist aber gar keine Bewegung. Wir sollen uns doch unsere Bewegung anschauen!“

Tabelle 5.4: Kodierschema für die Beurteilung der Inhaltsvalidität – Teil 3.

Titel	Code-Definition	Ankerbeispiel
2.2.2 Vorstellungen bzgl. phy. Erkenntnismethoden, Arbeitsweisen	– Die Situation könnte für Schüler_innen lernrelevant in Bezug auf Vorstellungen von physikalischen Erkenntnismethoden und Arbeitsweisen sein. – Vorstellungen sind kognitiv und umfassen sowohl die fachlich akzeptierte Vorstellung von Erkenntnismethoden und Arbeitsweisen, als auch Schülervorstellungen, die von Ersteren abweichen oder zu ihnen im Widerspruch stehen können. – Schülervorstellungen werden als kognitive Konstrukte verstanden, die auf bisherige – nicht zwingend schulische – (Lern-)Erfahrungen zurückgehen. Sie sind dem Bewusstsein des Lernenden zugänglich und beziehen sich auf spezifische Themenfelder des Faches, wie zum Beispiel die Mechanik, den Schvorvorgang oder chemische Reaktionen. (nach Krüger 2016)	Die Schüler_innen sollen ein Weg-Zeit-Diagramm zu ihrem Schulweg anfertigen. Lena sagt zu Max, dass ihr Weg langweilig sei. „Max: Ja, du wohnst ja auch direkt gegenüber! Ich bin aber auch fast fertig. Und was bringen uns diese Dinger jetzt hier?“

Tabelle 5.5: Kodierschema für die Beurteilung der Inhaltsvalidität – Teil 4.

Titel	Code-Definition	Ankerbeispiel
2.3 Bezüglich Prozessfähigkeit der Erkenntnisgewinnung	<i>In diesen Situationen können Maßnahmen zur Förderung von Fähigkeiten mehr oder weniger erfolgreich sein.</i>	
2.3.1 Beschreiben von Phänomenen	Phänomene zu beschreiben und sie eventuell auf bekannte physikalische Zusammenhänge zurückzuführen.	„Lara: Der Nagel wird von dem Magneten angezogen, das ist wie bei der Pinnwand, die zieht meine Magnete auch an.“
2.3.2 Daten, Informationen, Quellen (Relevanz und Ordnung)	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit Daten und Informationen aus verschiedenen Quellen zur Bearbeitung von Aufgaben und Problemen auszuwählen, auf Relevanz zu prüfen und sie zu ordnen.	„Rieke: Ich habe im Internet gelesen, dass LEDs weniger Strom verbrauchen, also brauche ich das auch nicht noch mal testen.“
2.3.3 Analogien und Modellvorstellungen	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit Analogien und Modellvorstellungen zur Wissensgenerierung zu verwenden.	„Sven: Das ist wie ein Klebstoff zwischen dem Nagel und dem Magneten.“
2.3.4 Formeln, Mathematisierung	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit einfache Formen der Mathematisierung anzuwenden und einfache Idealisierungen vorzunehmen.	„Tom: Naja, hier ist ein Widerstand und da ist noch einer, müssen wir dann nicht einfach $R1 + R2$ rechnen?“

Tabelle 5.6: Kodierschema für die Beurteilung der Inhaltsvalidität – Teil 5.

Titel	Code-Definition	Ankerbeispiel
2.3.5 Hypothesenbildung	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit an einfachen Beispielen Hypothesen aufzustellen.	„Kim: Also dann lautet die Hypothese: Die Energie ist proportional zur Kraft.“
2.3.6 angeleitetes Experiment durchführen	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit einfache Experimente nach Anleitung durchzuführen und diese auszuwerten.	„Sina: Ok, dann müssen wir jetzt laut Zettel das Kabel da in die Birne stecken. Hm die leuchtet nicht. Ist wohl kaputt.“
2.3.7 eigenes Experiment planen	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit einfache Experimente zu planen, durchzuführen und die Ergebnisse zu dokumentieren.	„Nele: Also wir können ja einfach mal hier ein Kabel und da ein Kabeln ran stöpseln, und dann gucken wir, wie hoch die Spannung ist.“
2.3.8 Daten auswerten	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit gewonnene Daten auszuwerten, ggf. auch durch einfache Mathematisierungen.	„Kai: Ok, also das hier ist eine Gerade, also ist der Zusammenhang von Energie und Kraft wohl linear?!“
2.3.9 Verallgemeinerung der Ergebnisse	Die Situation könnte lernrelevant sein bezüglich der Fähigkeit die Gültigkeit empirischer Ergebnisse und deren Verallgemeinerung zu beurteilen.	„Lutz: Dann können wir jetzt sagen, dass Licht immer so gebrochen wird.“

Eine intendiert kritische Situation wurde dann als tatsächlich kritisch akzeptiert, wenn fünf der sechs physikdidaktisch Tätigen diese Situation mindestens einem der Codes zugeordnet haben (Inhaltsvalidität). In einem zweiten Schritt wurde dieses Vorgehen auf die Hauptcodes (Code 2.2 „Bezüglich Vorstellungen“ und Code 2.3 „Bezüglich Prozessfähigkeit der Erkenntnisgewinnung“) übertragen, da die Feinkörnigkeit der Subcodes zu keiner überzeugenden Übereinstimmung oder sinnvollen Unterscheidung führte.

Eine kritische Situation im Drehbuch wurde letztendlich für inhaltlich valide erklärt, wenn mindestens in einem der Codes 2.2 oder 2.3 mindestens fünf der sechs physikdidaktisch Tätigen in ihrer Kodierung übereinstimmten. Die inhaltliche Validierung des gesamten Items (bestehend aus einer kritischen Situation im Drehbuch/Video und einer offenen oder geschlossenen Frage) war erst nach der Entwicklungsstudie II abgeschlossen, in der die offenen und geschlossenen Fragen entwickelt wurden.

5.1.4 Ergebnisse und Diskussion der inhaltlichen Validierung der Drehbücher

Die sechs Rater_innen für die Inhaltsvalidierung haben gemäß des oben vorgestellten Kodierschemas die fünf Drehbücher in Bezug auf die professionelle Unterrichtswahrnehmung bewertet. Dazu wurden alle kritischen Situationen in den Drehbüchern daraufhin untersucht, ob sie der professionellen Unterrichtswahrnehmung einer Lehrkraft bedürfen und welcher Lernaspekt (Schülervorstellungen oder Erkenntnisgewinnung) in der kritischen Situation zum Tragen kommt. Die folgenden Tabellen 5.7 bis 5.11 zeigen die Ergebnisse dieser Kodierungen für die Codes 2.2 (betrifft Schülervorstellungen) und 2.3 (betrifft Erkenntnisgewinnung).

Tabelle 5.7: Ergebnisse der Kodierung für die Beurteilung der Inhaltsvalidität – Drehbuch 1.

Situation	1	2	3	4	5	6	7	8	9	10
Code 2.2	6	2	5	4	5	1	5	4	1	4
Code 2.3	3	6	4	5	5	6	3	4	6	5

Wie in Tabelle 5.7 zu sehen ist, sind sich die physikdidaktisch Tätigen nur bezüglich der Situation 8 des ersten Drehbuchs nicht einig über die Signifikanz der kritischen Situation, da für beide Codes nur vier Kodierende Evidenz für die professionelle Unterrichtswahrnehmung in der Beschreibung der Situation gefunden haben. Situationen 1, 3 und 7 werden als Situationen mit eher lernrelevanten Schülervorstellungen kodiert, während die Situationen 2, 4, 6, 9 und 10 offenbar mehr Anlass zur Wahrnehmung der Erkenntnisgewinnungsprozesse geben. In Situation 5 müssen laut der in der Physikdidaktik Tätigen sowohl Aspekte der Schülervorstellungen als auch der Erkenntnisgewinnung professionell wahrgenommen werden. Die Einschätzungen der Rater_innen stimmen bis auf Situation 5 mit den In-

tentionen bezüglich der kritischen Situationen in Drehbuch 1 (siehe Unterabschnitt 5.1.2) überein.

Tabelle 5.8: Ergebnisse der Kodierung für die Beurteilung der Inhaltsvalidität – Drehbuch 2.

Situation	1	2	3	4	5	6	7
Code 2.2	6	0	4	4	2	0	4
Code 2.3	6	6	5	5	6	5	4

Die Kodierungen zum zweiten Drehbuch führten zur Exklusion der Situation 7, da sich nur vier der sechs Rater_innen über die Einschätzung der kritischen Situation einig sind. Situation 1 betrifft laut der physikdidaktisch Tätigen Schülervorstellungen. Allerdings finden sich in den Situationen 3 und 4 neben der Erkenntnisgewinnung auch lernrelevante Aspekte bezüglich Schülervorstellungen. Die Situationen 2, 5 und 6 werden dagegen durch die Rater_innen eher der Erkenntnisgewinnung zugeordnet. Intendiert war, dass die Situationen 1 und 3 die professionelle Unterrichtswahrnehmung bezüglich Schülervorstellungen und alle anderen kritischen Situationen die Erkenntnisgewinnung der Lernenden adressieren.

Tabelle 5.9: Ergebnisse der Kodierung für die Beurteilung der Inhaltsvalidität – Drehbuch 3.

Situation	1	2	3	4	5	6	7	8
Code 2.2	4	6	5	5	1	0	4	3
Code 2.3	6	4	6	5	6	6	5	5

Situation	9	10	11	12	13	14	15
Code 2.2	5	4	3	6	2	2	1
Code 2.3	5	5	6	3	6	6	6

Alle 15 intendiert kritischen Situationen des dritten Drehbuchs wurden durch die Kodierung der physikdidaktisch Tätigen als kritisch bestätigt. Dabei zählen mehrheitlich die Situationen 2 und 12 ausschließlich als relevant bezüglich Schülervorstellungen. Die Situationen 1, 5, 6, 7, 8, 10, 11, 13, 14 und 15 betreffen eher die Erkenntnisgewinnung. Beide Aspekte werden durch die Situationen 4 und 9 gleichzeitig adressiert. Für die Situation 3 sind sich zwar alle sechs Kodierende einig, dass professionelle Unterrichtswahrnehmung zu Erkenntnisgewinnung der Lernenden angesprochen wird. Allerdings sind sich auch fünf der sechs Rater_innen einig über wahrzunehmende Schülervorstellungen, was außerdem die Intention war. Die Einschätzungen der Rater_innen zu Situation 11 stimmt nicht mit der Intention überein.

Tabelle 5.10: Ergebnisse der Kodierung für die Beurteilung der Inhaltsvalidität – Drehbuch 4.

Situation	1	2	3	4	5	6	7
Code 2.2	6	6	6	6	6	5	6
Code 2.3	0	1	1	2	3	3	1

Situation	8	9	10	11	12	13	14
Code 2.2	5	5	5	5	6	5	5
Code 2.3	1	0	1	2	1	2	2

Die Ergebnisse der Kodierung zum Drehbuch 4 zeigen, dass alle geplant kritischen Situationen durch die physikdidaktisch Tätigen bestätigt wurden. Darüber hinaus sind alle kritischen Situationen durch die Mehrheit dem Code 2.2 zugeordnet worden. Alle kritischen Situationen scheinen wie durch das Drehbuch intendiert die professionelle Unterrichtswahrnehmung zu Schülervorstellungen anzusprechen.

Das Drehbuch 5 umfasst sechs kritische Situationen, die alle auch als kritisch von den Rater_innen hervorgehoben wurden. Von diesen kritischen Situationen betreffen Situation 2 und 3 die professionelle Unterrichtswahrnehmung bezüglich Schülervorstellungen.

Tabelle 5.11: Ergebnisse der Kodierung für die Beurteilung der Inhaltsvalidität – Drehbuch 5.

Situation	1	2	3	4	5	6
Code 2.2	4	6	6	4	5	4
Code 2.3	6	0	4	5	5	5

Situation 1, 4 und 6 wurden von der Mehrheit mit dem Code 2.3 versehen und betreffen somit offenbar eher den Aspekt der Erkenntnisgewinnung. Beide Aspekte wurden durch die Rater_innen in der Situation 5 aufgefunden. Intendiert war, dass die Situationen 2, 3, 4, 5 und 6 die professionelle Unterrichtswahrnehmung zu Schülervorstellungen adressieren.

Wie in Tabelle 5.12 dargestellt ist, wurden insgesamt 52 intendiert kritische Situationen in die Drehbücher aufgenommen. Von diesen 52 möglichen kritischen Situationen wurden durch die in der Physikdidaktik Tätigen 50 kritische Situationen durch die oben beschriebene Kodierung bestätigt. Die Ergebnisse weisen auf eine hohe Rate (50 von 52) von kritischen Situationen aus dem Physikunterricht hin, die von physikdidaktisch Tätigen als für eine Physiklehrkraft wahrnehmungswürdig identifiziert wurden. Wie in Abschnitt 4.2 zur Validität erläutert, weist das darauf hin, dass der Test inhaltlich valide ist, weil die Situationen laut Urteil der Experten tatsächlich der professionellen Unterrichtswahrnehmung entsprechen.

Tabelle 5.12: Ergebnisse zur inhaltlichen Validität.

Drehbuch	Anzahl der vorher festgelegten signifikanten Situationen (Schritt C)	Anzahl der inhaltlich validen signifikanten Situationen (Schritt D)
# 1	10	9
# 2	7	6
# 3	15	15
# 4	14	14
# 5	6	6
Total	52	50

mung bedürfen.. Dieses Ergebnis ist ein Indiz dafür, dass Interpretationen von Daten des PUPiL-Instruments einen berechtigt Schluss auf die professionelle Unterrichtswahrnehmung erlauben. Es bleibt allerdings zu klären, ob der Test auch curricular und ökologisch valide ist. Dies geschieht im nächsten Abschnitt. Eine erfolgreiche Konstruktvalidierung würde die Ergebnisse dieser Entwicklungsstudie zusätzlich stützen.

Zwischendiskussion Die Ergebnisse zur Inhaltsvalidierung werden bereits an dieser Stelle diskutiert, weil die Auswahl der inhaltlich validierten kritischen Situationen Grundlage für die curriculare und ökologische Validierung durch die Lehrkräfte war.

Im ersten Drehbuch konnte durch das Rating keine Übereinstimmung der physikdidaktisch Tätigen bezüglich der Situation 8 erreicht werden. Situation 8 beschreibt eine gelingende Unterrichtssituation, in der eine angemessene beziehungsweise ausbaufähige Schülervorstellung genannt wird. Die abgebildete Unterrichtsstunde wird aber durch die inadäquate Vorstellung über die schmelzenden Teeteile dominiert. Wie in Abschnitt 3.1 zum Experten-Novizen-Vergleich berichtet, konzentrieren sich erfahrene Lehrkräfte (und vermutlich Lehrkräfte im Allgemeinen) eher auf Atypisches und somit beispielsweise eher auf unangebrachte Schülervorstellungen als auf ausbaufähige. Daher ist es nicht verwunderlich und sogar erwartungskonform, dass eine kritische Situation mit gelingendem Lernen nicht zwingend als wahrnehmungswürdig kodiert wird, obwohl sie der Lehrkraft Rückmeldung über Lernprozesse und Lernstände geben kann. Die Situation bleibt weiter im Drehbuch enthalten, dient aber nicht als Item im endgültigen PUPiL-Instrument. In der Situation 5 äußert ein Schüler die inadäquate Schülervorstellung und formuliert dabei eine Hypothese. Dieser Umstand kann dazu geführt haben, dass die Rater_innen die kritische Situation nicht eindeutig einem der zwei Aspekte zugeordnet haben.

Die Situation 7 im Drehbuch 2 ist die letzte beschriebene Aktion der Unterrichtsstunde und zeigt eine inadäquate Verallgemeinerung der Beobachtungen in der Schülergruppe. Es wird dabei aber offenbar eine Schülervorstellung geäußert, die von den Rater_innen als ebenso wichtig wahrgenommen wurde wie der Prozess der Erkenntnisgewinnung. Diese kritische Situation wird daher aus dem finalen PUPiL-Instrument ausgeschlossen. Auch die Situationen 3 und 4 werden nicht eindeutig durch die Rater_innen eingeordnet. In Situation 3 ist die intendierte Schülervorstellung in eine Analogie aus dem Alltag einge-

bettet, was dazu geführt haben kann, dass fünf der sechs physikdidaktisch Tätigen auch die Erkenntnisgewinnung als lernrelevant wahrgenommen haben. Situation 4 beschreibt einen Schüler, der ein Gedankenexperiment durchführt. Dabei nennt er eine Vorstellung über Magneten. Dies hat bei den Rater_innen möglicherweise dazu beigetragen, auch diesen Aspekt in der kritischen Situation als lernrelevant zu kodieren.

In Drehbuch 3 nehmen die Rater_innen in Situation 11 eher das Erklären des Phänomens als die Schülervorstellung, dass Blitze etwas mit Kälte zu tun haben könnten, wahr. Es wird deutlich, dass die beiden Aspekte Schülervorstellungen und Erkenntnisgewinn nicht disjunkt sind, viele Teilprozesse der Erkenntnisgewinnung basieren auf Wissen oder im Fall von Lernenden auf Vorstellungen. Es ist daher nicht verwunderlich, dass bei dieser kritischen Situation der nicht intendierte Aspekt als kritisch markiert wurde. Für den Zweck dieser Studie ist es nicht relevant, ob Schülervorstellungen und Erkenntnisgewinnung disjunkt sind, sondern nur, dass kritische Situationen mit mindestens einem dieser Aspekte die professionellen Unterrichtswahrnehmung von Lehrkräften bedürfen. Ähnliches gilt für die Situationen 4 und 9. Auch diese wurden von den Rater_innen nicht eindeutig zugeordnet. Daher bleibt die Zuordnung wie im Drehbuch intendiert. In der Beschreibung eines Phänomens in Situation 3 wird die Schülervorstellung, dass elektrostatische Anziehung wie Kleben funktioniert, deutlich. Die Rater_innen sind sich einig, dass in dieser kritischen Situation die Beschreibung der Lernenden von Lehrkräften wahrgenommen werden müsste. Fünf der Rater_innen sind sich aber auch über die Lernrelevanz der geäußerten Schülervorstellung einig. Daher bleibt die intendierte Kategorisierung als relevant in Bezug auf Schülervorstellungen bestehen.

Das Drehbuch 4 beschreibt eine Unterrichtsstunde, in der die Lernenden rein kognitiv arbeiten. Es war daher zu erwarten, dass die physikdidaktisch Tätigen den Aspekt der Erkenntnisgewinnung nicht als wahrnehmungswürdig erachten.

Bezüglich Drehbuch 5 ergibt sich ein ähnliches Bild wie für Drehbuch 4. In der Situation 5 sind laut den Rater_innen beide Aspekte gleichermaßen relevant. Die Situationen 4 und 6 weisen für die Rater_innen außerdem eher Aspekte der Erkenntnisgewinnung als – wie intendiert – Schülervorstellungen auf. Nichtsdestotrotz sind alle Situationen als kritisch markiert worden.

Die inhaltliche Validität wurde für alle fünf entwickelten Drehbücher unter Streichung zweier intendiert kritischer Situationen erreicht mit der Einschränkung, dass nicht immer eindeutig ist, welcher der Aspekte Schülervorstellungen oder Erkenntnisgewinnung eher der professionellen Unterrichtswahrnehmung bedarf. Alle Drehbücher zusammen betrachtet führen zu maximal 50 kritischen Situationen. Aufgrund der begrenzten Testzeit, des Arbeitsgedächtnisses und der Aufmerksamkeit der Testpersonen wurde beschlossen, die Anzahl der Drehbücher auf drei zu begrenzen. Damit gehen 30 bis 40 kritische Situationen in die curriculare und ökologische Validierung über. Nach Erfahrungen aus der Präpilotierung mit dem Probevideo nehmen angehende Physiklehrkräfte ungefähr die Hälfte der Situationen wahr, die sie wahrnehmen sollten. Es wurden die drei Drehbücher 1, 3 und 4 für die Verarbeitung im finalen PUPiL-Instrument gewählt, da diese eine hohe Anzahl von kritischen Situationen enthalten. Jede Testperson bearbeitet nur zwei der drei Videos.

Falls alle inhaltlich validierten kritischen Situationen auch der curricularen und ökologischen Validierung Stand halten, können jene Testpersonen, die zukünftig Video 1 und 3 bearbeiten, insgesamt 24 kritische Situationen wahrnehmen. Für die Kombination aus Video 1 und 4 beinhalten die Drehbücher 23 kritische Situationen. Die Kombination von Video 3 und Video 4 würde zu einer Gesamtzahl von 29 kritischen Situationen führen, die von zukünftigen Testpersonen wahrgenommen werden sollten. Die Drehbücher 1, 3 und 4 umfassen insgesamt 38 kritische Situationen, welche noch bezüglich der curricularen und ökologischen Passung bewertet werden mussten.

5.1.5 Schritt 5 – Fragebogen für die curriculare und ökologische Validität

Für die Sicherung der curricularen und ökologischen Validität wurden fünf praktizierende Physiklehrkräfte befragt. Ziel ist es, durch die (kleine) Stichprobe an Lehrkräften Einschätzungen bezüglich der curricularen und ökologischen Validität der Drehbücher aus der Perspektive zukünftiger Testpersonen zu erhalten. Die Fähigkeit, Unterricht professionell wahrzunehmen, kann von Testpersonen nur gezeigt werden, wenn das Geschehen so authentisch wie möglich ist, weil das PUPiL-Instrument und die Erhebungssituation von den Testpersonen dann als glaubwürdig akzeptiert werden kann. Neben den (bis dahin nur inhaltlich validierten und zunächst ausgewählten drei) Drehbüchern, erhielten die Lehrkräfte je Drehbuch einen zweigeteilten Fragebogen (siehe Anhang B). Jede der folgenden Aussagen auf dem Fragebogen wurde durch Ankreuzen (Stimme zu; Stimme nicht zu) bewertet und bezog sich auf das ganze Drehbuch. Unter jeder Aussage wurden die Lehrkräfte gebeten, ihre Antwort in einem leeren Textfeld zu begründen.

Folgende Aussagen wurden bezüglich der *curricularen Validität* bewertet:

1. Das in der Unterrichtssequenz behandelte Thema ist in den curricularen Vorgaben der Hamburger Bildungsstandards enthalten.
2. Die der Unterrichtssequenz zugrundeliegenden Lernziele werden durch die curricularen Vorgaben Ihrer Schule abgedeckt.

Diese Aussagen sollten schriftlich durch Zustimmung oder Ablehnen bewertet werden. Anschließend wurden die Lehrkräfte um eine schriftliche Begründung für ihre Zustimmung oder Ablehnung gebeten. Durch diese Bewertung wurde garantiert, dass nur kritische Situationen in das Instrument eingehen, die curricular und ökologisch validen Unterricht zeigen. Wenn das Instrument die Fähigkeit, Unterricht professionell wahrzunehmen, valide messen soll, muss dies anhand von Unterricht geschehen, den die (angehenden) Lehrkräfte so oder so ähnlich erfahren können. Es wurden daher von vornherein Themen gewählt, die in den Bildungsstandards und Lehrplänen einiger Hamburger Schulen enthalten sind. Da die nationalen Bildungsstandards aber einen Interpretationsspielraum bieten, konnte durch diese Fragen herausgefunden werden, ob die Interpretation der Drehbuchautorin mit der Interpretation der praktizierenden Lehrkräfte übereinstimmt. Somit wurde die Themenauswahl durch das Urteil der Lehrkräfte validiert.

Bezüglich der *ökologischen Validität* bewerteten die Lehrkräfte folgende sechs Aussagen:

1. In dieser Unterrichtssequenz ist die **Interaktion zwischen den Lernenden und der Lehrkraft** authentisch für den Physikunterricht.
2. In dieser Unterrichtssequenz ist die **Interaktion zwischen den Lernenden** authentisch für den Physikunterricht.
3. In dieser Unterrichtssequenz ist die **Sozialform** authentisch für den Physikunterricht.
4. In dieser Unterrichtssequenz sind die **Arbeitsweisen der Lernenden** authentisch für den Physikunterricht.
5. In dieser Unterrichtssequenz sind die **Unterrichtsmethoden** authentisch für den Physikunterricht.
6. Diese Unterrichtssequenz könnte im Physikunterricht einer allgemeinbildenden Schule so tatsächlich stattfinden.

Auch bezüglich Interaktionen und anderen Oberflächenstrukturen von Unterricht, sollen die Drehbücher möglichst curricular und ökologisch valide sein, um die Lehrkräfte durch das PUPiL-Instrument in eine annähernd reale Lehr-Lern-Situation zu bringen. In dieser können sie ihre Fähigkeit, Unterricht professionell wahrzunehmen, nur zeigen, wenn das Geschehen so authentisch wie möglich ist und damit das PUPiL-Instrument und die Erhebungssituation von den Testpersonen als glaubwürdig akzeptiert werden. Darum beurteilten die Lehrkräfte die Drehbücher im Hinblick auf die obigen sechs Aussagen.

5.1.6 Ergebnisse der curricularen und ökologischen Validierung der Drehbücher

Fünf Lehrkräfte haben mithilfe des oben vorgestellten Fragebogens die curriculare und ökologische Validität der drei im vorherigen Schritt ausgewählten Drehbücher bewertet. Sie untersuchten dazu entlang des Fragebogens, ob die Drehbücher authentischen Unterricht beschreiben. Es ergibt sich für alle Drehbücher ein ähnliches Bild.

Drehbuch 1 Für das erste Drehbuch wurde nur von vier der fünf Lehrkräften eine Bewertung abgegeben, weil die bewertende Person 1 sich für die Beurteilung des Themengebiets an Stadtteilschulen nicht befähigt fühlte. Wie Tabelle 5.13 zeigt, stimmen zwei der vier Lehrkräfte der Aussage zu, dass die Brown'sche Molekularbewegung von den Hamburger Bildungsplänen abgedeckt ist. Sie begründen dies mit der Einordnung in die Themen Wärmelehre oder Energie, welche Teil des mittleren Schulabschlusses seien. Eine Lehrkraft hat zu dieser Aussage keine Angabe gemacht, eine weitere Lehrkraft hat die Aussage abgelehnt mit der Begründung, dass es nicht notwendiger Bestandteil der Hamburger Bildungspläne sei.

Bezüglich der zweiten Aussage haben alle vier Lehrkräfte zugestimmt, dass die Brown'sche Molekularbewegung an ihrer Schule durch einen Lehrplan vorgesehen sei. Die entsprechenden Begründungen lauten:

- „Nicht die Lernziele, aber das Thema im Curriculum.“
- „Eher aufgeteilt in Klasse 7 und 10.“
- „Gehört zum Thema 'Energie'.“

Tabelle 5.13: Ergebnis der Lehrkräftebefragung zur curricularen und ökologischen Validität – Drehbuch 1.

	Curriculare Validierung		Ökologische Validierung					
	Frage 1: Hamburger Bildungspläne	Frage 2: Lehrplan der Schule	Frage 1: Interaktion Lehrkraft und Lernende	Frage 2: Interaktion zwischen den Lernenden	Frage 3: Sozialform	Frage 4: Arbeitsweisen der Lernenden	Frage 5: Unterrichts- methoden	Frage 6: Real mögliche Unterrichts- sequenz
Anzahl der Zustimmungen	2	4	4	4	4	4	4	4

Bezüglich der Fragen zur ökologischen Passung sind sich alle vier Lehrkräfte einig: Das Drehbuch 1 ist in Hinblick auf die Interaktionen, Sozialform, Arbeitsweisen und Unterrichtsmethoden authentisch (siehe Tabelle 5.13). Die Interaktionen der Lernenden erhalten unterschiedliche Kommentare: Von „Kann ich mir so vorstellen“ über „SuS³² würden vielleicht Bestätigung bei Lehrkraft einfordern“ bis hin zu „Schülerantwort teilweise zu konstruiert“.

Die Interaktionen der Lernenden untereinander werden als etwas zu gesittet beschrieben, die Verwendung des Begriffs „schmelzen“ des Tees deutet für eine Lehrkraft auf eine Schule mit hohem Migrationsanteil hin.

Die Sozialform in Drehbuch 1 wurde als üblich bezeichnet.

Die Arbeitsweisen der Lernenden wurden als typisch bewertet. So begründet eine Lehrkraft, dass Lernende häufig oberflächlich handelten und vorschnell Schlüsse zögen, weil sie sich oder andere überschätzen.

Die Bemerkungen zu den Unterrichtsmethoden variieren. Eine Lehrkraft empfindet sie als typisch, eine andere würde den Lernenden mehr Bedenkzeit einräumen. Wieder eine andere Lehrkraft stuft die Aufgabe als problemorientiert ein, die daher viel Zeit in Anspruch nimmt und deshalb wahrscheinlich eher selten eingesetzt wird. Dennoch schätzt die Lehrkraft den Einsatz der Methoden als denkbar ein.

Zur sechsten Frage wurde angemerkt, dass Lehrkräfte wahrscheinlich schon beim ersten Aufkommen des Begriffs „schmelzen“ eingegriffen hätten. Eine andere Lehrkraft hält die Stunde im Drehbuch 1 für realisierbar, sie würde die naturwissenschaftlichen Arbeitsweisen lehren.

³²Anmerkung der Autorin: Dies ist eine gängige Abkürzung für „Schülerinnen und Schüler“.

Drehbuch 3 Für das dritte Drehbuch haben alle fünf Lehrkräfte ihre Stimme abgegeben. Tabelle 5.14 zeigt, dass drei der fünf Lehrkräfte das Thema der Elektrostatik den Hamburger Bildungsstandards zuschreiben (eine Ablehnung, eine Enthaltung). Die Begründungen weisen darauf hin, dass das Thema nicht zu den Mindestanforderungen zu zählen scheint und für eine neunte Klasse zu spät angesetzt ist. Alle Lehrkräfte finden die Thematik dennoch im Lehrplan ihrer Schule wieder. Es scheint, dass die Elektrostatik an mindestens zwei Schulen grundlegend ist und in einer anderen schon in Klasse 7 unterrichtet wird.

Die ökologische Passung des Drehbuchs 3 wird für fünf der sechs Fragen von allen Lehrkräften bestätigt.

Tabelle 5.14: Ergebnis der Lehrkräftebefragung zur curricularen und ökologischen Validität – Drehbuch 3.

	Curriculare Validierung		Ökologische Validierung					
	Frage 1: Hamburger Bildungspläne	Frage 2: Lehrplan der Schule	Frage 1: Interaktion Lehrkraft und Lernende	Frage 2: Interaktion zwischen den Lernenden	Frage 3: Sozialform	Frage 4: Arbeitsweisen der Lernenden	Frage 5: Unterrichts- methoden	Frage 6: Real mögliche Unterrichts- sequenz
Anzahl der Zustimmungen	3	5	5	5	5	4	5	5

Allerdings zeigen die Kommentare zu den einzelnen Fragen, dass das Drehbuch eher schlechten, nicht wünschenswerten Unterricht zeigt. Es sollte laut den befragten Lehrkräften weniger Unterrichtszeit auf das Thema verwendet werden. Die Antworten der Lernenden wirken für eine der bewertenden Lehrkräfte zu konstruiert. Die Arbeitsweisen und die Sozialform werden aber als typisch für den Physikunterricht eingestuft.

Drehbuch 4 Für das vierte Drehbuch stimmen alle Lehrkräfte den beiden Aussagen zur curricularen Passung ohne Einschränkung zu (vergleiche Tabelle 5.15).

Tabelle 5.15: Ergebnis der Lehrkräftebefragung zur curricularen und ökologischen Validität – Drehbuch 4.

	Curriculare Validierung		Ökologische Validierung					
	Frage 1: Hamburger Bildungspläne	Frage 2: Lehrplan der Schule	Frage 1: Interaktion Lehrkraft und Lernende	Frage 2: Interaktion zwischen den Lernenden	Frage 3: Sozialform	Frage 4: Arbeitsweisen der Lernenden	Frage 5: Unterrichts- methoden	Frage 6: Real mögliche Unterrichts- sequenz
Anzahl der Zustimmungen	5	5	5	5	5	4	5	5

Wieder haben fünf der sechs Aussagen zur ökologischen Passung von Drehbuch 4 alle Lehrkräfte zugestimmt. Wie im Drehbuch 3 gibt es eine Enthaltung bezüglich der Arbeitsweisen der Lernenden. Die Interaktionen zwischen Lehrkraft und Lernenden werden

zum Teil als typisch (3 Bemerkungen) und zum Teil (2 Bemerkungen) als eher untypisch bezeichnet. Eine Lehrkraft schlägt vor, dass die Lehrkraft im Drehbuch die geäußerten Schülervorstellungen an der Tafel festhalten sollte, um besseren Unterricht zu geben.

Die Interaktionen der Lernenden, Sozialformen, Arbeitsweisen und Unterrichtsmethoden werden als (sehr) typisch beschrieben.

Eine Lehrkraft fasst auf Frage 6 hin zusammen, dass der Unterricht nicht gut sei, aber vorkommen könne.

Alle drei Drehbücher betreffend, wurde von den befragten Lehrkräften am ehesten fehlendes oder unpassendes Verhalten der Lehrkräfte bemängelt.

5.2 Abschließende Diskussion und Zusammenfassung

Die Ergebnisse der curricularen und ökologischen Validierung können aufgrund der geringen Stichprobengröße ($N = 5$) nicht verallgemeinernd betrachtet werden. Dennoch geben die Antworten der Lehrkräfte Hinweise darauf, wie die Drehbücher und später die Videos auf zukünftige Testpersonen wirken könnten.

5.2.1 Curriculare Validität

Kritische Anmerkungen zu Frage 1 (nach dem Hamburger Bildungsplan) könnten daher rühren, dass die Hamburger Bildungsstandards für Physikunterricht (Beschluss der Kultusministerkonferenz, 2004) wörtlich und nicht interpretierend gelesen wurden. Da sich zu dieser Frage nicht jede befragte Lehrkraft positioniert hat, hätte die Frage offener und auffordernder formuliert werden müssen.

Um herauszufinden, ob der beschriebene Unterricht im Lehralltag vertreten sein kann, ist die Frage 2 (nach dem Lehrplan der Schule) von höherer Bedeutung als Frage 1. Auch wenn eine konkrete Thematik in den Bildungsstandards nicht explizit gefordert wird, können sich Schulen über schulinterne Curricula dazu entscheiden, diese Thematik mit in ihren Physikunterricht aufzunehmen. Bei dieser Befragung wurde durch alle Lehrkräfte bestätigt, dass die beschriebenen Themen Teil des schulinternen Curriculums sind. Daraus wird geschlossen, dass derartige Unterricht prinzipiell vorkommen kann und somit für das finale PUPiL-Instrument einsetzbar ist.

5.2.2 Ökologische Validität

Die Kritik der befragten Lehrkräfte bezüglich einer mangelnden Rückmeldung der Lehrkraft ist berechtigt. In einem guten Unterricht wäre ein moderierendes Verhalten samt lernförderlicher Rückmeldungen zu erwarten. Dennoch muss in diesen Drehbüchern darauf verzichtet werden, damit der Fokus der professionellen Unterrichtswahrnehmung auf den Lernenden und nicht auf der Lehrkraft liegt.

Auch das „gesittete“ und „konstruierte“ Verhalten der Lernenden wurde zu Recht kritisiert. In den Drehbüchern wurden die „Nebenschauplätze“ durch Verhalten der anderen anwesenden Lernenden bewusst nicht beschrieben, um beim Dreh der Videos authentischeres Verhalten zu erhalten. Wäre jede Handlung der Lernenden bereits festgelegt, müsste die schauspielerische Leistung der Lernenden sehr gut sein. Ohne Anweisung fühlen sie sich vermutlich wie Lernende im Unterricht und verhalten sich dementsprechend (weniger gesittet und konstruiert).

Bezüglich Frage 5 wurde bemängelt, dass die Lehrkraft teilweise zu viel Zeit für einfache Schritte im Unterricht aufwendet. Für die Umsetzung im PUPiL-Instrument ist es aber wichtig, viel Zeit auf die einzelnen Aspekte des Unterrichts zu verwenden, um mehr Indizien für die professionelle Unterrichtswahrnehmung zu liefern.

Es wurde außerdem gefordert, dass die Lehrkraft (früher) auf die Schülervorstellungen eingehen sollte, was aber im Drehbuch nicht geschieht. Mit der Einschätzung haben die Lehrkräfte auf Grundlage ihrer Erfahrungen Recht. Allerdings ist auch an dieser Stelle ein Eingreifen der Lehrkraft im Drehbuch nicht möglich, weil die erste Dimension der professionellen Unterrichtswahrnehmung nicht mehr gemessen werden kann, wenn im Video bereits explizit darauf aufmerksam gemacht wurde.

Wie oben bereits angekündigt, wurde begründet auch Unterricht beschrieben, der den Merkmalen guten Unterrichts nicht entspricht. Die Anmerkungen der bewertenden Lehrkräfte bestätigen, dass dies gelungen ist, auch wenn sich die Lehrkräfte eher guten Unterricht wünschen würden.

Abschließend waren die verschiedenen Meinungen zur der Kategorisierung typischer oder untypischer Interaktionen zu erwarten, da fünf verschiedene Lehrkräfte beurteilen, die unterschiedliche individuelle Erfahrungen mit Physikunterricht und Lernenden haben. Für die Gestaltung der Drehbücher als Grundlage der Konstruktion des Messinstruments sind Kompromisse hinsichtlich der ökologischen Validität daher akzeptabel.

Der Beurteilung durch die Lehrkräfte folgend, können alle drei Drehbücher als curricular und ökologisch valide erklärt werden. Die Einschränkungen durch mehr oder weniger typische Interaktionen zwischen den Lehrkräften und Lernenden sowie das Handeln der Lehrkräfte fällt im Gegensatz zu den zu bewertenden anderen Aussagen weniger ins Gewicht. Die curriculare und ökologische Validität wurde demnach für alle drei Drehbücher sichergestellt. Während der Verfilmung der validierten Drehbücher mithilfe des Teams des Medienzentrums der Fakultät für Erziehungswissenschaft der Universität Hamburg wurde durch die Regie der Autorin sichergestellt, dass alle 38 kritischen Situationen in der intendierten Art und Weise wahrnehmbar sind. In der Entwicklungsstudie II (siehe nächstes Kapitel) wurden Distraktoren für den geschlossenen Teil des Fragebogens für jedes Video entwickelt. Die Kombination aus kritischer Situation und geschlossener Frage dient im PUPiL-Instrument der Hauptstudie als Item. Auf Grundlage der Erfahrungen aus dem Pilot eines ersten Videos wurde für die Entwicklungsstudie II und Hauptstudie die Aufgabe für die Testpersonen festgelegt. Außerdem wurde für die Verfilmung die Kameraperspektive aus Sicht der Lehrkraft (wie im Pilot) durch eine statische Fokussierung der Paare oder Gruppen ersetzt, damit die Testpersonen zum einen nicht in eine

advokatorische Bewertung der Lehrkräfte geleitet und zum anderen durch die unruhige Kameraführung nicht irritiert werden.

Für die in dieser Arbeit operationalisierte Fähigkeit der physikspezifischen professionellen Unterrichtswahrnehmung von Schülervorstellungen und Erkenntnisgewinnung in den Themenbereichen Brown'sche Molekularbewegung, Elektrizität und Energieumwandlung zeigt die Entwicklungsstudie I, dass 38 der intendiert kritischen Situationen zur Inhaltsvalidität beitragen, weil sie Teil der interessierenden Gesamtheit möglicher Items im Sinne der Inhaltsvalidität sind (vgl. Abschnitt 4.2 zur Validität), um die adressierte Fähigkeit zu messen.

Kritisch zu beachten ist allerdings, dass nicht alle bekannten Schülervorstellungen zu den Themen in den Drehbüchern behandelt werden. Es werden auch nicht alle Teilprozesse der Erkenntnisgewinnung von den Lernenden durchlaufen. Diese Tatsachen schränken die Inhaltsvalidität zwar ein, der Test wäre sonst aber nicht mehr einsetzbar: Wenn alle relevanten Inhalte in ihrer erschöpfenden Breite und Tiefe abgehandelt werden würden, ergäbe sich eine Testzeit der Erhebung, die vermutlich weit über einem vertretbaren Rahmen läge.

6 Entwicklungsstudie II

In dieser Teilstudie wurde die professionelle Unterrichtswahrnehmung explorativ und qualitativ in Hinblick auf die Passung und Schwierigkeit des Tests untersucht. Daneben steht die Präsentation der Distraktorenentwicklung im Fokus dieses Kapitels. Die Distraktoren stellen gemeinsam mit den kritischen Situationen der Videos aus der Entwicklungsstudie I die Items des PUPiL-Instruments in der Hauptstudie dar. Für die Entwicklung eines validen Tests ist es wichtig, dass neben den kritischen Situationen aus den Videos auch die geschlossenen Antwortmöglichkeiten repräsentativ im Sinne der Inhaltsvalidität für alle möglichen Antwortmöglichkeiten zu den Situationen sind. Damit die Items nur Inhalte repräsentieren, die sich auf für die (zukünftigen) Testpersonen Relevantes beziehen, werden die Distraktoren der Fragen auf Basis von Performanz bezüglich der professionellen Unterrichtswahrnehmung von Lehramtsstudierenden und Lehrkräften entwickelt. Repräsentative Items sind eine Voraussetzung für eine erfolgreiche Inhaltsvalidierung (vgl. Hartig et al., 2012, S. 152). Diese Forderung ist zwar praktisch nur schwer umsetzbar, dennoch werden in diesem Kapitel Schritte, die versuchen eine solche Inhaltsvalidität zu erreichen, beschrieben.

Die Distraktoren für einen geschlossenen Fragebogen wurde in zwei Zyklen der Entwicklungsstudie II entwickelt. Die zweite Durchführung (Abschnitt 6.2) baut auf den Ergebnissen des ersten Durchgangs von Entwicklungsstudie II auf, um Verbesserungen an der Erhebung vorzunehmen. Dieses Kapitel ist demnach ebenfalls zweigeteilt. Es wird zuerst die erste Durchführung in Abschnitt 6.1 präsentiert, bevor in Abschnitt 6.2 auf die zweite Durchführung eingegangen wird. In abschließenden Diskussionen werden die Ergebnisse interpretiert und beurteilt.

6.1 Entwicklungsstudie II.1

Zu Beginn dieser Teilstudie lagen drei validierte und gefilmte Videos vor, die gestellte Unterrichtssituationen zeigen. Ziel der Entwicklungsstudie II war es,

- a) erste Erkenntnisse über die professionelle Unterrichtswahrnehmung, die mit den vorliegenden Videos erfasst werden kann, zu erlangen, also die Passung des Instruments zur intendierten Stichprobe der Hauptstudie und somit die Schwierigkeit des Instruments einschätzen zu können, und
- b) aus den Antwortmustern der Testpersonen geeignete Distraktoren zu jeder kritischen Situation zu gewinnen.

Die Entscheidung, den Fragebogen für die Hauptstudie (teilweise) geschlossen zu gestalten, basierte auf der Fokussierung einer physikspezifischen professionellen Unterrichtswahrnehmung. Wenn das Instrument nur mit offenen Antwortformaten ausgestattet wäre, hätte ein Fokus auf diese fachspezifische Fähigkeit nur schwer umgesetzt werden können. Testpersonen hätten dann die Möglichkeit, alle visuell wahrnehmbaren Ereignisse in den Videos zu adressieren. Die Information zur physikspezifischen professionellen Unterrichtswahrnehmung würde in der Fülle an wahrnehmbaren Aspekten (vergleiche Abschnitt 3.2) untergehen. Nichtsdestotrotz gibt das PUPiL-Instrument durch den halboffenen Fragebogen in der Hauptstudie die Möglichkeit, Aussagen über eine eher generische professionelle Unterrichtswahrnehmung abzuleiten (vgl. hierzu Lefstein & Snell, 2011).

6.1.1 Methode

Für die Entwicklungsstudie II.1 wurde eine Stichprobe mit $N = 12$ Testpersonen zusammengestellt. Die Testpersonen haben sich nach Werbung in didaktischen Lehrveranstaltungen oder nach direkter Ansprache freiwillig für die Teilnahme an der Erhebung gemeldet. Die Stichprobe umfasste jeweils zwei Physiklehramtsstudierende des 2., 6. und 10. Semesters, sowie zwei Physiklehrkräfte im Vorbereitungsdienst und zwei Physiklehrkräfte sowie jeweils eine Person im Geographielehramtsstudium des 4. und 6. Semesters (vgl. Tabelle 6.1). Geographielehramtsstudierende wurden unter anderem in die Stichprobe mitaufgenommen, um auch über die physikspezifische professionelle Unterrichtswahrnehmung angehender Geographielehrkräfte Erkenntnisse zu gewinnen und überzeugende Distraktoren für beide Gruppen entwickeln zu können (vgl. H1). In der Erhebungssituati-

Tabelle 6.1: Stichprobe der Entwicklungsstudie II.1.

Fach	2. Semester	4. Semester	6. Semester	10. Semester	Lehrkräfte in Vorbereitung	Lehrkräfte
Physik	2	0	2	2	2	2
Geographie	0	1	1	0	0	0

on wurden die Testpersonen darüber aufgeklärt, dass das Ziel der Erhebung ein generelles Verständnis von Wahrnehmung sei und nicht die Beurteilung der Fähigkeit jeder einzelnen Testpersonen verfolgt würde (für den Leitfaden siehe Anhang C). Anschließend wurde ihnen die Methode des lauten Denkens erklärt, welche sie anhand einer Aufgabe üben sollten (Verbesserung technischer Geräte: Erfinden Sie fünf Verbesserungen für Ihr Handy.³³). Nach Absolvieren der Übung bekamen die Testpersonen eine Rückmeldung über ihr Verhalten und Hinweise, wie sie ihr lautes Denken verbessern könnten.

Das laute Denken (oder Erinnern) wurde als introspektive Erhebungsmethode (vgl. Konrad, 2010) angewandt, um den Testpersonen die Möglichkeit zu geben, sich in die Lehrkräfte in den Videos hineinzusetzen. Dies erfolgte, um im Rahmen der Erhebungssituation

³³Diese Aufgabe zur Übung des lauten Denkens wurde in Feser (eingereicht) entwickelt.

Verbaldaten zu erhalten, von denen möglichst valide auf das Konzept der professionellen Unterrichtswahrnehmung geschlossen werden kann. Auf diese Weise kann abgeleitet werden, ob tatsächlich die Fähigkeit der professionellen Unterrichtswahrnehmung zu den Ergebnissen führt, oder ob eine für irrelevant erachtete Fähigkeit Anteil an den Ergebnissen hat. Wenn geschlussfolgert werden kann, dass vor allem die professionelle Unterrichtswahrnehmung zur Lösung der Aufgaben beiträgt, ist das ein Indiz für Konstruktvalidität (vgl. erste Frage bezüglich Konstruktvalidität in Abschnitt 4.2 zur Validität).

Theoretisch lässt sich die Praxis des lauten Denkens in die drei Formen *Introspektion*, *unmittelbare Retrospektion* und *verzögerte Retrospektion* einteilen (vgl. Konrad, 2010, S. 476). Diese drei Formen unterscheiden sich in der zeitlichen Nähe des Artikulierens zum kognitiven Prozess (vgl. hier und in den folgenden Sätzen Konrad, 2010). Die *Introspektion* findet während des Denkens statt, da die Gedanken des Kurzzeitgedächtnisses bereits verbalisierbar vorliegen. Testpersonen äußern solche Gedanken auch, wenn sie nicht dazu aufgefordert wurden. Diese Daten können Einblick in ihre professionelle Unterrichtswahrnehmung geben. Dahingegen müssen nicht enkodierte Gedanken zunächst in Sprache überführt werden. Dies ist dann der Fall, wenn die Testpersonen einer Aufforderung nachgehen und nur bestimmte Teile ihres Kurzzeitgedächtnisses wiedergeben sollen (*unmittelbare Retrospektion*). Werden Gedanken im Nachhinein erklärt, so spricht man von *verzögerter Retrospektion*. Diese Form kommt in der Erhebung der Entwicklungsstudie II nur vor, wenn sich die Testpersonen entscheiden Gedanken zu mehreren Ereignissen auf einmal und erst nachträglich zu äußern. In der Erhebung der Entwicklungsstudie II kam hauptsächlich die zweite Form vor, da die Testpersonen folgende Aufgabe erhielten, bevor sie zwei der drei Videos angesehen haben:

1. Bitte stoppen Sie das jeweilige Video immer dann, wenn Sie etwas in Bezug auf ...
 - ... Vorstellungen über fachliche Begriffe und Konzepte der Schüler_innen,
 - ... Erkenntnisgewinnungsprozesse der Schüler_innen,
 - ... Vorstellungen über Erkenntnisgewinnung der Schüler_innen wahrnehmen³⁴.

(Das Stoppen des Videos muss unmittelbar und in sehr kurzer Zeit nach dem Prozess des Erkennens geschehen, damit dem unmittelbaren Handlungs- und Ereignisdruck realen Unterrichts Rechnung getragen werden kann.)
2. Begründen Sie laut denkend, warum ihnen diese Situation aufgefallen ist.
3. Stellen Sie sich zu jeder Situation vor, Sie stünden neben den agierenden Lernenden, äußern Sie laut denkend, wie Sie handeln würden.

Durch das laute Denken werden die kognitiven Prozesse im besten Fall nicht geändert, aber durch die nun notwendige Kodierung der Gedanken verlangsamt (vgl. Konrad, 2010, S. 479). Die Schwäche dieser Methode ist, dass möglicherweise nicht alle Gedanken erfasst werden können, die für die Studie von Bedeutung sind, da ein Intervenieren der Erhe-

³⁴Es wurde zu diesem Zeitpunkt noch unterschieden, ob die Testpersonen Vorstellungen über Konzepte und Begriffe oder Erkenntnisgewinnung wahrnahmen. Über die Analyse der Daten wurde aber deutlich, dass letzterer Aspekt immer als Teil der Erkenntnisgewinnung gehandelt wurde.

bungsleitung die kognitiven Prozesse und somit die Belastbarkeit der Daten beeinflussen würde. Die Testpersonen würden in eine eher reflektierende Haltung übergehen. Wie in Abschnitt 2.4 zur Abgrenzung von Reflexion und Diagnose bereits erläutert, sind Reflexion und professionelles Wahrnehmen verschiedene Prozesse. Für eine valide Messung der professionellen Unterrichtswahrnehmung sollte es daher vermieden werden, reflexive Prozesse zuzulassen.

Jedes Video wurde durch dieses Vorgehen von acht Testpersonen angesehen. Die Videos beginnen mit vorgelesenen und abgedruckten Kontextinformationen über das Vorwissen der Lernenden und Ziele der Lehrkraft für die Unterrichtsstunde, während im Hintergrund die Lernenden im Klassenraum Platz nehmen. Zusätzlich erhalten die Testpersonen den Aufgabentext, eine Definition zu Schülervorstellungen, die Liste der Aspekte der Erkenntnisgewinnung nach den Bildungsstandards und die Materialien der Lernenden abgedruckt auf einem Papier (siehe Interviewleitfaden Anhang C).

Die Daten wurden entlang der verschiedenen Erkenntnisinteressen dieser Entwicklungsstudie II auf zweierlei Weise analysiert. Die Darstellung der Analysen samt Ergebnissen und Interpretation der Ergebnisse erfolgt daher in zwei getrennten Unterabschnitten.

6.1.2 Teilanalyse A: Prüfung der Passung und Schwierigkeit des Tests

Nach einer Transkription³⁵ und Segmentierung, wurden die Verbaldaten mit einem deduktiven Kategoriensystem qualitativ analysiert (Intercoderreliabilität nach Brennan & Prediger (1981) $\kappa_n = 0.86$). Das Material wurde skalierend strukturiert (vgl. skalierende Strukturierung bei Mayring, 2015, S. 101-109). Das verwendete deduktive Kodierschema ist in Tabelle 6.2 dargestellt. In weiteren Hinweisen zu diesem Kodierschema ist abgedruckt, was in welcher der kritischen Situationen als lernrelevant wahrgenommen werden sollte. Es waren außerdem nur diejenigen Situationen zu kodieren, die in der Entwicklungsstudie I als kritisch validiert wurden. Demnach wurden die Aussagen der Testpersonen – angelehnt an die Operationalisierung der professionellen Unterrichtswahrnehmung (vergleiche Abschnitt 4.4 zur Operationalisierung) – zuerst daraufhin analysiert, ob sie die intendiert kritischen Situationen betreffen. Wenn das der Fall war, wurde untersucht, ob die Aussagen der Testpersonen das kritische Moment der Situation treffen (beispielsweise die Schülervorstellung, dass Tee im warmen Wasser schmilzt). Dabei wurden verschiedene Ausprägungen der professionellen Unterrichtswahrnehmung unterschieden (vgl. Abbildung 6.1): Eine Aussage kann ...

1. das Erkennen einer kritischen Situation ausdrücken (Perception).

³⁵Die Transkripte wurden in Anlehnung an den Vorschlag von Mayring (2015) angefertigt. Es wurden die Gedankenstriche für Pausen bei Mayring durch Punkte ersetzt. Ausdrücke wurden zwar bereinigt, dialektische Färbungen wurden allerdings, anders als bei Mayring vorgeschlagen, transkribiert.

2. den Kern der kritischen Situation richtig benennen und/oder erläutern (Interpretation).
3. das Fassen eines Handlungsplans nach expliziter oder impliziter Interpretation der Situation beschreiben und/oder begründen (Decision Making).

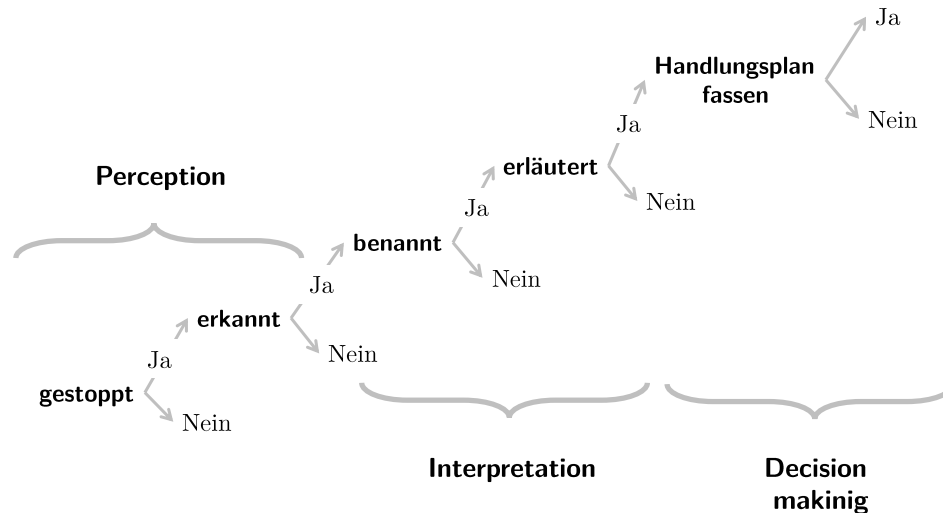


Abbildung 6.1: Ausprägungsstufen der professionellen Unterrichtswahrnehmung bezüglich kritischer Situationen der Videos während des lauten Denkens.

Die Antworten der Testpersonen geben Aufschluss über die erreichte Ausprägungsstufe der physikspezifischen professionellen Unterrichtswahrnehmung.

6.1.2.1 Ergebnisse – Passung und Schwierigkeit des Tests

Nach einer deduktiven qualitativen Inhaltsanalyse kann berichtet werden, von wie vielen Testpersonen die jeweiligen Situationen in den Videos (a) erkannt, (b) benannt, (c) erläutert oder (d) in einen Handlungsplan überführt wurden (vgl. Abbildung 6.1 und Abbildung 6.2). An dieser Stelle werden die Ergebnisse exemplarisch an fünf repräsentativen Situationen berichtet.

- Für Situation 1 aus Video 1 wurde von drei Testpersonen ein Handlungsplan gefasst. Diese Ausprägung der professionellen Unterrichtswahrnehmung enthält immer (entweder explizit oder implizit) die vorherigen Ausprägungen. Eine der Testpersonen führt ihre Wahrnehmung folgendermaßen aus (in kursiv ist der Handlungsplan abgedruckt):

ANN1967: „Wir haben hier die (.) äh das Konzept, dass Tee schmilzt.
Da würde ich als Lehrerin sa-, auf jeden Fall (darauf) eingehen, dass Tee nicht schmilzt, sondern (.) dass man die Extrakte von Tee aus dem Blättern herauszieht.“

Neben diesen drei Testpersonen haben keine weiteren Testpersonen die Situation thematisiert.

- Situation 5 aus Video 1 wurde von fünf Testpersonen nur erkannt, von einer Testperson adäquat benannt. Vier weitere Testpersonen haben die Ausprägung „Handlungsplan fassen“ erreicht. Eine davon äußert beispielsweise laut denkend (in kursiv ist der Handlungsplan abgedruckt):

BRI1772: „So hier unterbreche ich auch einmal kurz. Ähm (..) das ist mir vorhin schon aufgefallen (...) Der Aslan spricht von den Teeteilchen, die irgendwie schmelzen. Offensichtlich besteht da (..) ähm ein ein Konzept, eine Schülervorstellung, dass (..) sozusagen der Tee (...) ähm ja schmelzen muss, damit er sich im Wasser verteilt. Und äh (...) da muss man natürlich irgendwie noch mal drauf eingehen, denn ähm (...), das ist natürlich äh, das ist natürlich nicht so. (lacht) (...) Gut, das würde ich aber (...) klar, was würde man jetzt in der Situation gleich machen (fragend)? Ja, weiß nicht, würde man da fragen, würde man die Parallele ziehen, äh: Verhält sich Tee genauso wie Eis? Wie ist das von der Konsistenz her? Was passiert mit Eis, wenn es warm wird? Dann schmilzt es, es s. Wird Tee, den man so lose hat ähm und ihn dann warm macht, ähm schmilzt der dann auch? (Atmet ein) Also da müsst man noch mal irgendwie (...) das noch mal irgendwie (..) durchgehen.“

Situation 5 ist außerdem die Situation, zu der die Testpersonen am häufigsten ihre Wahrnehmung geäußert haben.

- Situation 4 aus Video 3 (Situation 13 in Abbildung 6.2) ist ein Beispiel dafür, dass trotz Erläuterung der erkannten und benannten Situation keine Testperson einen Handlungsplan entwickelt hat. Eine Testperson erkennt die Situation, zwei weitere Testpersonen benennen die Charakteristik der Situation, während eine Person in der Lage ist, diese zu erläutern (kursiv):

BEA1767: „Genau und er bringt jetzt gleich das Wort Ladung ins Spiel. Ähm für ihn ist das mit dem Kleben auch ganz eindeutig falsch. Also ich geh- auch davon aus, dass sie -ne falsche Vorstellung hat. Aber man weiß es halt, find- ich, noch nicht genau. Und (..) ähm (..) bei ihm is- jetz- halt die Frage, er sagt, er hat es im Fernsehen gesehen. Und bringt auch das Wort Ladung rein. Is- jetz- die Frage, ob er schon irgend -ne Vorstellung dahinter hat oder ob er sich einfach erfolgreich Begriffe aus der Fernsehsendung gemerkt hat.“

- Die Reaktion der Testpersonen auf die erste Situation in Video 4 (Situation 25 in Abbildung 6.2) zeigt, dass nur eine Testperson die Situation als wahrnehmungswürdig erkannt hat, aber keine Testperson in der Lage war, das kritische Moment der Situation zu benennen oder zu erläutern. An der Formulierung der Testperson ist zu erkennen, wie schwer begreifbar die Situation ist:

MAR1653: „Also da hatte der gerade eine Schülervorstellung, ich habs jetzt nur grad nicht ganz im Kopf, was für eine, aber irgendwie so eine Schülervorstellung über Energie, wann man die (..) ähm (...), ich habs jetzt grad nich- nicht so richtig aufgepasst, aber dass sie eben ähm (..) ja so ein Beispiel eben. Und der macht jetzt ein Beispiel.“

- Die Situation 10 des vierten Videos (Situation 34 in Abbildung 6.2) wurde von keiner einzigen Testperson als wahrnehmungswürdig erkannt.

Die Ergebnisse zu kritischen Situationen, die ähnliche Muster in Abbildung 6.2 zeigen, können analog abgelesen werden.

Insgesamt haben zu 25 der 38 Situationen weniger als die Hälfte der Testpersonen ihre

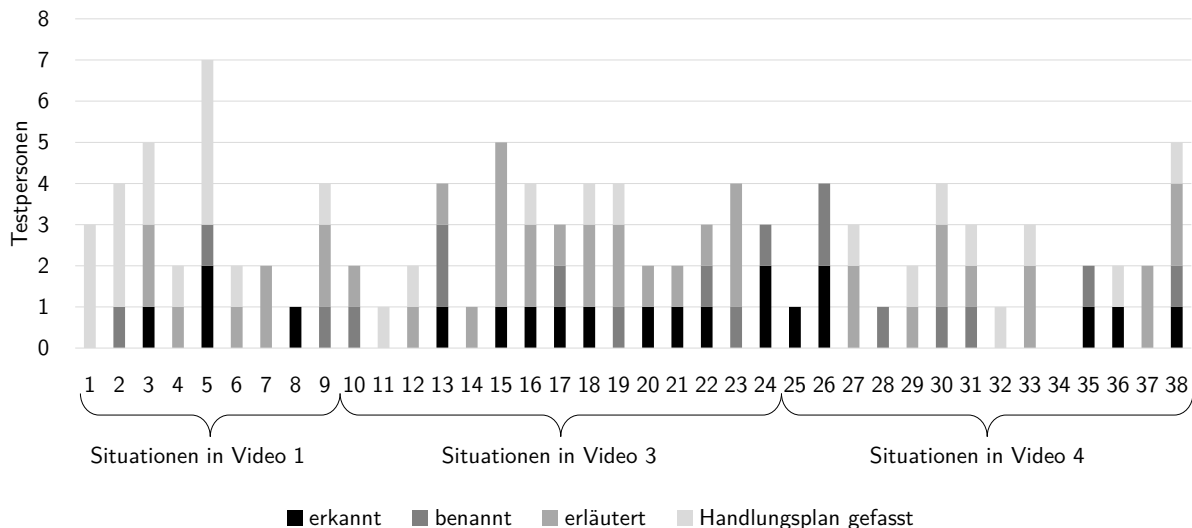


Abbildung 6.2: Ergebnisse Prüfung der Passung und Schwierigkeit des Tests.

Wahrnehmung mitgeteilt. Das bedeutet auch, dass zu 13 Situationen sich (mehr als) die Hälfte der Testpersonen geäußert hat. In 11 Situationen haben nur zwei Testpersonen den kritischen Kern der Situation mindestens erkannt. Zu sechs Situationen hat jeweils nur eine Testperson ihre Wahrnehmung durch Bearbeiten der Aufgaben ausgedrückt.

6.1.2.2 Interpretation und Diskussion – Passung und Schwierigkeit des Tests

In diesem Unterkapitel werden die Interpretationen der Ergebnisse präsentiert. Den einzelnen Interpretationen folgen Diskussionen, die aufzeigen, welche Befunde aus dieser Teilstudie die folgenden Schritte beeinflusst haben.

Video 1 zeigt Unterricht zum Thema Brown'schen Molekularbewegung anhand des Lösens von Tee. Die Unterrichtssituationen zeigen zum einen Schülervorstellungen, die für das Fachlernen hinderlich sein könnten, zum anderen Erkenntnisgewinnung mit gelingenden und nicht gelingenden Momenten. Die Ergebnisse der Auswertung zeigen, dass die Situationen 4, 6, 7 und 8 schwerer wahrzunehmen zu sein scheinen als alle anderen Situationen, da sie nur von wenigen Testpersonen überhaupt erkannt wurden.

Situationen 4, 6 und 7 wurden jeweils nur von zwei Testpersonen thematisiert. Das ist ein Indiz dafür, dass die Bearbeitung des Tests an diesen Stellen schwerfällt. Situation 4 handelt vom Aufstellen nicht überprüfbarer Hypothesen durch Lernende, während Situation 6 zeigt, dass die Lernenden ihr Experiment nicht reflektiert durchführen. Situation 7 handelt vom erneuten Benennen der Schülervorstellung, dass Teeteilchen schmelzen. Möglicherweise nehmen die Testpersonen diese Aussage wahr, aber äußern sich trotz der Aufforderung laut zu denken nicht darüber, weil sie in einer vorherigen Situation bereits laut dazu gedacht haben.

Tabelle 6.2: Kodierschema zur Reaktion auf die Videos in der Entwicklungsstudie II.

Titel	Code-Definition	Ankerbeispiel
1. Erkannt	Die Testpersonen benennen, was sie wahrgenommen haben. Sie haben etwas anderes wahrgenommen als die intendiert kritische Situation.	„Ich seh, dass Hypothese im (.) Punkt zwei (...) nicht ganz richtig geschrieben ist (lacht). (...) naja, is ja nicht so schlimm.“
2. Benannt	Die Testpersonen benennen, was sie wahrgenommen haben und treffen damit den Kern der kritischen Situation.	„[...] Da stellt er auch wieder Bezug auf Alltagsbeispielen, diese Phänomene da.“ „Genau, das is- -n bisschen, dass was ich gerade meinte. Ähm er sagt jetzt-, ja an der Tafel steht das ja so. Also (.) er hat vielleicht noch nicht ganz den Prozess dahinter verstanden, warum sie sich jetzt auch -ne Begründung dafür überlegen sollen. Warum sie eben diese Hypothese aufstellen sollen, sondern er macht das halt erstmal, weil's da so steht. (..) Aber ich denke, eben dadurch, dass er es jetzt so macht, wird er auch irgendwie diesen Prozess verinnerlichen können. [...]“
3. Erläutert	Die Testpersonen erläutern ihre Wahrnehmung. Sie können ihre Wahrnehmung explizit benennen oder nicht. Es wird davon ausgegangen, dass einen Erläutern immer ein Benennen vorausgeht.	Nicht-Handeln: „Joa, die Schüler kommen hier gut ins Gespräch. Da würd- ich als Lehrer tatsächlich auch das weiterlaufen lassen.“
4. Handlungsplan gefasst	Die Testpersonen erklären, was sie als Leinrkraft unternehmen würden. Die Benennung und Erläuterung ihrer Wahrnehmung kann explizit oder implizit geschehen. Es wird davon ausgegangen, dass dem Fassen eines Handlungsplans eine Interpretation vorausgeht. Es kann auch geplant werden, <i>nicht</i> zu handeln.	Handeln: „[...] nochmal darauf hinweisen (.) ehm durch, durch (.) Fragen auf diesen Begriff des Schmelzens eingehen, da könnte man eh ich würde fragen ehm „Schmilzt der Tee wirklich?“, um das nochmal zu hinterfragen eh (3), dass es eh beim Schmelzen um ein Aggregatzustandswechsel geht von einem festen zu, eh von einem ja festen zum flüssigen Aggregatzustand eh (.) und, dass das nicht zu der Situation passt, die er da beschreiben möchte.“

Es wäre für das Erkenntnisinteresse dieser Teilstudie wünschenswert gewesen, mit ihnen in ein Gespräch über die einzelnen Situationen zu kommen, allerdings würde ein klassisches Interview die Nähe zur eigentlichen Unterrichtshandlung der Testpersonen schmälern. Somit wäre auch die Konstruktvalidität verringert. Aus diesem Grund wurde die Erhebung im Stil des lauten Denkens geführt. Die achte Situation in Video 1³⁶ zeigt eine angemessene Beschreibung und angemessene Erklärung der Beobachtungen eines Lernenden. Wie bereits erläutert, ist es erwartungskonform, wenn Testpersonen gelingende Lernsituationen nicht wahrnehmen, weil diese nicht zwingend nach einer (sichtbaren) Handlung verlangen. Situationen 2, 3, 5 und 9 scheinen insgesamt für mehr als zwei der Testpersonen lösbar, das heißt, sie konnten die Situation mindestens erkennen und im besten Fall einen Handlungsplan fassen, um in der Situation lernförderlich zu agieren. Es lässt sich in den Daten beobachten, dass die Schülervorstellung des „Schmelzens von Tee“ für die Testpersonen im Fokus steht.

Das Video 3 behandelt ein Thema aus der Elektrostatik. Neun der 15 kritischen Situationen wurden von weniger als der Hälfte der Testpersonen wahrgenommen. Die kritische Situation, die von den Testpersonen am häufigsten bearbeitet wurde, ist Situation 6 (Situation 15 in Abbildung 6.2). Diese wird von vier Testpersonen erläutert und einmal benannt. In der Situation ergänzt ein Schüler die Beschreibung einer Beobachtung seines Mitschülers. Die Wahrnehmung dieses Ereignisses könnte leicht fallen, weil kurz zuvor durch die Beschreibung des Mitschülers bereits ein Indikator für die Lernrelevanz des Ereignisses dargeboten wurde. Damit könnte auch erklärt werden, warum Situation 5, die unmittelbar vor Situation 6 zu beobachten ist, nur von einer Person erläutert wurde. Es bedarf möglicherweise mehr als nur eines Indikators für lernrelevante Prozesse im Unterricht, um die professionelle Unterrichtswahrnehmung gelingender Augenblicke ins Bewusstsein der Testpersonen zu bringen. Gleichzeitig konnte für Video 1 beobachtet werden, dass eine spätere Wiederholung eines Indikators für eine kritische Situation weniger häufig wahrgenommen wurde. Auffällig, aber nicht verwunderlich ist, dass keine Person einen Handlungsplan äußert. Im Fall einer gelingenden Situation trifft eine Lehrkraft vermutlich die Entscheidung, nicht zu handeln und implizit das Gesehene in die Beurteilung des Lernstandes der Lernenden miteinzubeziehen. Während des lauten Denkens führt diese eher unbewusste Entscheidung dazu, dass keine Handlungspläne genannt werden.

Weitere Auffälligkeiten können nach der Analyse der Antworten zu diesem Video nicht festgestellt werden. Anders als bei Video 1 werden gelingende und nicht gelingende Unterrichtssituationen gleich häufig (nicht) adressiert.

Das Video 4 zeigt Physikunterricht zum Thema Energieumwandlung und betrifft Schülervorstellungen über den sogenannten Energieverbrauch. Abbildung 6.2 zeigt, dass nur drei der 14 Situationen von (mehr als) der Hälfte der acht Testpersonen mindestens erkannt wurden. Diese Situationen handeln von einem Beitrag eines Schülers, der ein angemessenes Beispiel zur Energieumwandlung liefert (Situation 2), einer Aussage einer Schülerin, die erkennen lässt, dass das Konzept der Energieumwandlung verstanden wird (Situation 6) und vom Schüler aus Situation 2, der die Reibung als mögliche Erklärung für das zu erklä-

³⁶Achtung: Situation 8 im Video entspricht bei der Beschreibung der ursprünglichen Drehbücher Situation 9, da Situation 8 durch die inhaltliche Validierung nicht bestätigt wurde.

rende Phänomen benennt. Dieses sind alles Situationen, die eher gelingende Konzeptbildung bei den Lernenden zeigen. In diesem Video – wieder im Gegensatz zu Video 1 – sind die am häufigsten wahrgenommenen Situationen nicht lernhinderlich, sondern -förderlich. Und andersherum sind diejenigen Situationen, die nur von einer Testperson mindestens erkannt wurden (Situation 1, 4 und 8) eher lernhinderlich. Es lässt sich daher vermuten, dass die professionelle Unterrichtswahrnehmung neben der Fachspezifität auch eine themenspezifische Ausprägung haben könnte. In einigen Themen sind Lernschwierigkeiten leichter wahrzunehmen, bezüglich anderer Themen eher die lernförderlichen Situationen. Es ist außerdem auffällig, dass Situation 10 (Situation 34 in Abbildung 6.2) überhaupt nicht bewusst wahrgenommen wurde. In der Situation wird durch eine Schülerin das erste Mal ein Konzept der nicht nutzbaren Energie eingeführt, welches den Lernenden hilft zu verstehen, dass Energie nicht verbraucht, sondern umgewandelt wird. Wie in Bezug zu Video 2 bereits erwähnt, kommt es vor, dass Testpersonen erst nach wiederholter Indikation einen Situationstypen wahrnehmen.

Einen Handlungsplan äußert je eine Testperson für acht der 14 Situationen. Das betrifft bei zwei Situationen gelingende Lernprozesse (Situation 5 und 9), bei zwei Situationen (6 und 7) einen Übergang von der Verbrauchs- zur Umwandlungsvorstellungen und bei den übrigen vier Situationen weniger gelingende Lernprozesse (Situation 3, 8, 12 und 14).

Zusammengefasst fällt auf, dass den Testpersonen das Äußern ihrer Wahrnehmung schwer fällt. Eine mögliche Ursache hierfür ist, dass die Testpersonen die Handhabung des Instruments nicht gewohnt sind. Außerdem ist trotz Aufforderung auch ein „Nicht-Handeln“ zu äußern, in den Daten zu beobachten, dass nach gelingenden Situationen (in Abbildung 6.2 Situationen 3, 8, 10, 13, 14, 15, 16, 17, 19, 21, 23, 25, 29, 33, 34) zu zwei Drittel der Situationen keine Handlungspläne geäußert werden. Im Vergleich dazu gab für ein Drittel der Situationen, in denen nicht angemessene Schülervorstellungen oder Erkenntnisgewinnung gezeigt werden, keine Testperson einen Handlungsplan an. Im Mittel wurde zu 35% aller Situationen kein Handlungsplan angegeben. Obwohl die Fähigkeit Unterricht professionell wahrzunehmen theoretisch als ein dreidimensionales Konstrukt definiert wurde, lassen diese Ergebnisse erkennen, dass die Teilfähigkeit, einen Handlungsplan zu äußern, auf die hier praktizierte Art und Weise weniger valide erfasst werden kann. Möglicherweise führen geringe Abänderungen in der Erhebungssituation bereits zu belastbaren Ergebnissen. Außerdem liefern die Daten aus der Erhebung mit Anteilen des lauten Denkens nicht genügend Grundlage, um Distraktoren für einen geschlossenen Fragebogen zu entwickeln und abzusichern. Die Daten reichten aber aus, um auf deren Basis ein erstes induktives Kategoriensystem zu entwickeln und zu testen. Mit diesem konnten die Daten adäquat beschrieben und Distraktoren entwickelt werden. Nach bisheriger Auswertung sind nur 13 von 38 kritischen Situationen durch mehr als die Hälfte der Testpersonen bearbeitet worden. In einem zweiten Durchgang wird ein Training der Testpersonen mit einem formatnahen aber inhaltsfernen Video (professionelle Wahrnehmung im Straßenverkehr) vorgeschaltet, bevor das Instrument für die Hauptstudie vorbereitet werden kann. Die Konstruktvalidierung des PUPiL-Instruments wird dann in einer Hauptstudie abgeschlossen.

6.1.3 Teilanalyse B: Entwicklung von Distraktoren

In einer zweiten qualitativen Inhaltsanalyse (Mayring, 2015) wurden Antworten der Testpersonen nach einem induktiven Kodierschema hinsichtlich der Aspekte der Situationen untersucht, die die Testpersonen wahrgenommen und geäußert haben. Das Kodierschema ist ein Teilergebnis dieser Entwicklungsstudie II und wird mit den nötigen Angaben zur Güte im folgenden Unterkapitel vorgestellt. Dafür wurde das Material in einer zusammenfassenden Inhaltsanalyse nach Mayring (2015, S. 67-83) bearbeitet.

Aus der Theorie zu gutem Unterricht (vgl. Lipowsky, 2006; Merzyn, 2015; Meyer, 2003; Seidel et al., 2006) kann abgeleitet werden, dass Unterricht nicht nur aus fachdidaktischen Aspekten wie Schülervorstellungen und Erkenntnisgewinnung besteht, sondern ebenso generische Aspekte wie Klassenraumführung oder andere didaktische Aspekte lernrelevant werden. Auch wenn die Testpersonen explizit nur um ihre Wahrnehmung bezüglich Schülervorstellungen und Erkenntnisgewinnung gebeten wurden, konnte beobachtet werden, dass andere Aspekte wahrgenommen und geäußert wurden (Introspektion). Die offenen Antworten einiger Testpersonen deuten durch die Fokussierung auf andere Unterrichtsaspekte daraufhin, dass die professionelle Unterrichtswahrnehmung dieser Testpersonen sich alleine dadurch auszeichnet.

Anhand von Verbaldaten je einer Testperson aus dem Bachelor (Lehramt Physik und Geographie), Master (Lehramt Physik), Vorbereitungsdienst und Berufsleben (BRI1854, CHR1784, BEA1767, RAM1900, GAB1761) wurde induktiv ein Kodierschema entwickelt und anschließend auf alle Daten der Entwicklungsstudie II angewandt. Für die Kodierung wurden die Passagen des lauten Denkens den bestehenden Kategorien zugeordnet. Wenn dies nicht möglich war, wurde eine neue Kategorie hinzugefügt oder die Codedefinition bestehender Kategorien angepasst.

6.1.3.1 Ergebnisse – Entwicklung von Distraktoren

Entgegen der Erwartungen, die an die Entwicklungsstudie II gestellt wurden, stehen wie oben ausgeführt nur wenige Aussagen der Testpersonen für die Entwicklung von Distraktoren zur Verfügung. Die Verbaldaten wurde dennoch dafür genutzt, ein induktives Kategoriensystem abzuleiten, um aus den aktuellen Daten und Daten der zweiten Entwicklungsstudie II Distraktoren für den geschlossenen Teil des Fragebogens zu entwickeln. Durch qualitative Analyse erster Daten des lauten Denkens wurden vier Kategorien entwickelt (vergleiche Tabellen 6.3 und 6.4):

1. Vorstellungen zu fachlichen Begriffen und Konzepten (kurz: Schülervorstellungen)
2. Erkenntnisgewinnung
3. Classroom Management
4. anderer didaktischer Aspekt

Die Kategorie „Vorstellungen zu fachlichen Begriffen und Konzepten“ beinhaltet – ähnlich wie in der Entwicklungsstudie I – Aussagen der Testpersonen zu wahrgenommenen

Vorstellungen der Lernenden. Es werden in dieser (und jeder anderen) Kategorie auch diejenigen Aussagen berücksichtigt, die kontrafaktisch sind, also Ereignisse beschreiben, die nicht im Video vorkommen. Außerdem beinhaltet die Kategorie nicht nur die intendiert kritischen Situationen und auch nicht nur den Kern dieser, sondern alle Aussagen der Testpersonen zu Vorstellungen der Lernenden. Dadurch wird eine reichhaltigere Grundlage für die Entwicklung der Distraktoren gelegt. Auch wenn die Aussagen inhaltlich nicht zu den kritischen Situationen passen, geben sie Aufschluss darüber, welche Aspekte durch Testpersonen von sich aus adressiert werden und welche Strukturen diese Antworten liefern. Die Interpretation reicht als Dimension vom Beschreiben der Situation hin zu evidenzbasierter Beurteilung. Auf welcher Interpretationsebene – beschreibend oder interpretierend – die Distraktoren formuliert werden müssen, kann aus eben diesen Antworten abgeleitet werden.

In die Kategorie „Erkenntnisgewinnung“ wurden Aussagen der Testpersonen einsortiert, wenn diese Prozesse der Erkenntnisgewinnung von Lernenden betreffen. Die Beschreibung der Kategorie wurde theoriegeleitet durch die Bildungsstandards (Beschluss der Kultusministerkonferenz, 2004) ergänzt, auch wenn nicht zu allen Unterpunkten in den Bildungsstandards Aussagen im Datenmaterial gefunden wurden. Auch in dieser Kategorie wurden kontrafaktische Aussagen beachtet. Bei der Kodierung fiel auf, dass einige Testpersonen von Erkenntnisgewinnung sprachen, aber „Verstehen“ meinten:

ANN1967: „Hier haben wir auch wieder ein- Erkenntnisgewinnung v-von der Schülerin. Wobei sie immer noch denkt, dass Energie was mit Objekten zu tun hat, das man Energie nur auf ein Objekt umw- oder auftr- oder anwenden kann, was ja nicht immer der Fall ist. Beispiel: Potentielle Energie.“

Daher wurde im Kodierschema festgehalten, dass solche Aussagen nicht in diese Kategorie sortiert werden sollen, sondern in Kategorie 1.

Unter der Kategorie „Classroom Management“ werden Aspekte des Classroom Managements (nach Seidel, 2009) sowie Aussagen zu extrinsischer Motivation, fachunabhängiger Ereignisse und Lehrkräftezentrierung zusammengefasst, auch wenn dies nicht den gängigen Definitionen von Classroom Management entspricht. Eine detailliertere Unterscheidung dieser Kategorie ist für das Ziel dieser Entwicklungsstudie II nicht notwendig, da eine Differenzierung nach dem Grad an Fachspezifität ausreicht, um die aufgestellten Hypothesen zu testen.

Als letzte Kategorie werden daher Aussagen zu fachspezifischen Ereignissen, die aber nicht Vorstellungen zu Begriffen und Konzepten oder Erkenntnisgewinnung betreffen, über „andere didaktische Aspekte“ zusammengefasst.

Es kann aus diesen Daten auch abgeleitet werden, dass die Antworten, wenn sie nicht den Kern der kritischen Situation treffen, auf zwei Weisen inadäquat sind:

- Die Antworten adressieren einen Unterrichtsaspekt wie Classroom Management oder andere didaktische Aspekte, nach dem in der Aufgabe nicht gefragt ist.
- Die Antworten adressieren einen Unterrichtsaspekt, wie Schülervorstellungen oder Erkenntnisgewinnung, sind aber kontrafaktisch.

Tabelle 6.3: Kodierschema zur Distraktorenentwicklung in der Entwicklungsstudie II – Teil 1.

Kategorien	Code-Definition	Ankerbeispiel
1. Vorstellungen zu fachlichen Begriffen und Konzepten	Die Testpersonen beschreiben, dass Lernende richtige, falsche, anschlussfähige oder überhaupt Vorstellungen zu dem betrachteten Begriffen und Konzepten äußern.	„Die zweite Schülerin glaubt, dass sich Tee in kaltem Wasser lösen kann. Allerdings würde es länger dauern.“ „Der Junge hat richtig erkannt, dass das Phänomen etwas mit elektrischen Ladungen zu tun haben muss.“
	Die Testpersonen beschreiben, dass Lernende Prozesse der Erkenntnisgewinnung vollziehen. Dies sind zumeist das Aufstellen von Hypothesen oder das Durchführen von Experimenten. Verstehen $\neq$ Erkenntnisgewinn. Die Lernenden ...	
	– beschreiben Phänomene und führen sie auf bekannte physikalische Zusammenhänge zurück,	
	– wählen Daten und Informationen aus verschiedenen Quellen zur Bearbeitung von Aufgaben und Problemen aus, prüfen sie auf Relevanz und ordnen sie,	„Schülerin stellt Hypothese auf (dauert länger), bringt dann Vergleich (Eiswürfel).“
2. Erkenntnisgewinnung	– verwenden Analogien und Modellvorstellungen zur Wissensgenerierung,	
	– wenden einfache Formen der Mathematisierung an, nehmen einfache Idealisierungen vor,	„Der Schüler hat sehr gut beobachtet und nicht direkt gedeutet.“
	– stellen an einfachen Beispielen Hypothesen auf,	
	– führen einfache Experimente nach Anleitung durch und werten sie aus,	
	– planen einfache Experimente, führen sie durch und dokumentieren die Ergebnisse,	
	– werten gewonnene Daten aus, ggf. auch durch einfache Mathematisierungen,	
	– beurteilen die Gültigkeit empirischer Ergebnisse und deren Verallgemeinerung	

Tabelle 6.4: Kodierschema zur Distraktorenentwicklung in der Entwicklungsstudie II – Teil 2.

Kategorie	Code-Definition	Ankerbeispiel
3. Classroom Management	- Die Testpersonen bemerken, dass Aspekte einer guten Klassenführung (nicht) berücksichtigt sind. - Die Ereignisse sind fachunspezifisch, eher lehrerzentriert oder handeln von extrinsischer Motivation.	„Die Lehrerin hat den Arbeitsauftrag an die Tafel geschrieben, was sich als sinnvoll herausstellt, da der Schüler bereits jetzt nochmal an der Tafel nach schaut was genau zu tun ist.“ „Die Jungs hinten links scheinen auch nicht zu arbeiten.“
4. anderer didaktischer Aspekt	- Kommunikationskompetenz - Sprachliche Aspekte (Fachsprache, DAZ, ...) - Gender + Genderbild - Fachspezifisch - Eher Schüler-zentriert - Intrinsische Motivation	„Die Schüler halten sich an die Vorgaben an der Tafel und diskutieren ihre Vorstellungen.“ „Deutschdefizit beim rechten Schüler.“ „Der Schüler diktiert seiner Partnerin die Vermutung vor.“

Tabelle 6.5: Ergebnisse der InterCoder-Reliabilität für die Erhebung mit Kappa nach Testpersonen.

Interview	ANN1967	BEA1767	BIR1742	BRI1772	BRI1894	CHR1705
κ_n	0.888	1	0.903	0.862	0.736	1

Interview	CHR1784	ERD1831	GAB1671	MAR1653	RAM1900	SAB1835
κ_n	0.853	0.866	0.874	0.908	0.841	0.716

Tabelle 6.6: Ergebnisse der InterCoder-Reliabilität für die Kategorien mit Kappa.

Kategorie	Schüler- vorstellungen	Erkenntnis- gewinnung	Classroom Management	anderer did. Aspekt	Sonstige	Gesamt
κ_n	0.928	0.918	0.955	0.968	0.979	0.869

In dem Kodierschema, welches zur Entwicklung der Distraktoren entstanden ist, werden außerdem folgende allgemeine Hinweise gegeben:

- Die Kodiereinheiten (zu kodierende Textstellen) sind die Beschreibung und Interpretation der wahrgenommenen Situationen. Allerdings kann der Handlungsplan, den einige Testpersonen fassen, helfen die Interpretation zu verstehen.
- Die gestellten Unterrichtsvideos sollen nur als Orientierung gelten, es soll keinesfalls kodiert werden, was im Video zu sehen ist.
- In die Kategorie „Sonstige“ ist jede transkribierte Passage des lauten Denkens einzuordnen, die nicht zu einer der oberen vier Kategorien passt.

Das gesamte Datenmaterial wurde von zwei Personen unabhängig voneinander kodiert. Anschließend wurde eine Überprüfung der InterCoderübereinstimmung nach Brennan & Prediger (1981, S. 693) mittels $\kappa_n = \frac{\sum P_{ii} - \frac{1}{n}}{1 - \frac{1}{n}}$ vorgenommen, wobei P_{ii} die relative Anzahl der Übereinstimmungen je Kategorie und n die Anzahl der Kategorien ist. Die InterCoderreliabilität – aufgeschlüsselt nach den einzelnen Testpersonen und den Kategorien sowie dem gesamten Material – ist in Tabelle 6.5 und Tabelle 6.6 aufgelistet. Für einen Kappa-Wert zwischen 0.4 und 0.6 spricht man von einer *moderaten Übereinstimmung*, zwischen 0.6 und 0.8 von einer *hohen Übereinstimmung* und zwischen 0.8 und 1 von einer *sehr hohen Übereinstimmung* (vgl. Landis & Koch, 1977, S. 165). Nach dieser Einteilung ergibt sich, dass das Kodierschema für die Erhebung mit zwei Testpersonen (BRI1894 und SAB1835) zu einer hohen Übereinstimmung führt. Für die restlichen zehn Testpersonen ergibt sich eine sehr hohen Übereinstimmung zwischen den zwei Kodierenden.

Die Kodierenden kommen für alle Kategorien und für das gesamte Kodiersystem zu einer sehr hohen Übereinstimmung.

Die Kodierung der Verbaldaten zeigt auf Grundlage der wenigen Daten, dass die verschiedenen Testpersonen auch verschiedene Fokusse in ihrer Wahrnehmung zu legen scheinen (vergleiche Abbildung 6.3).

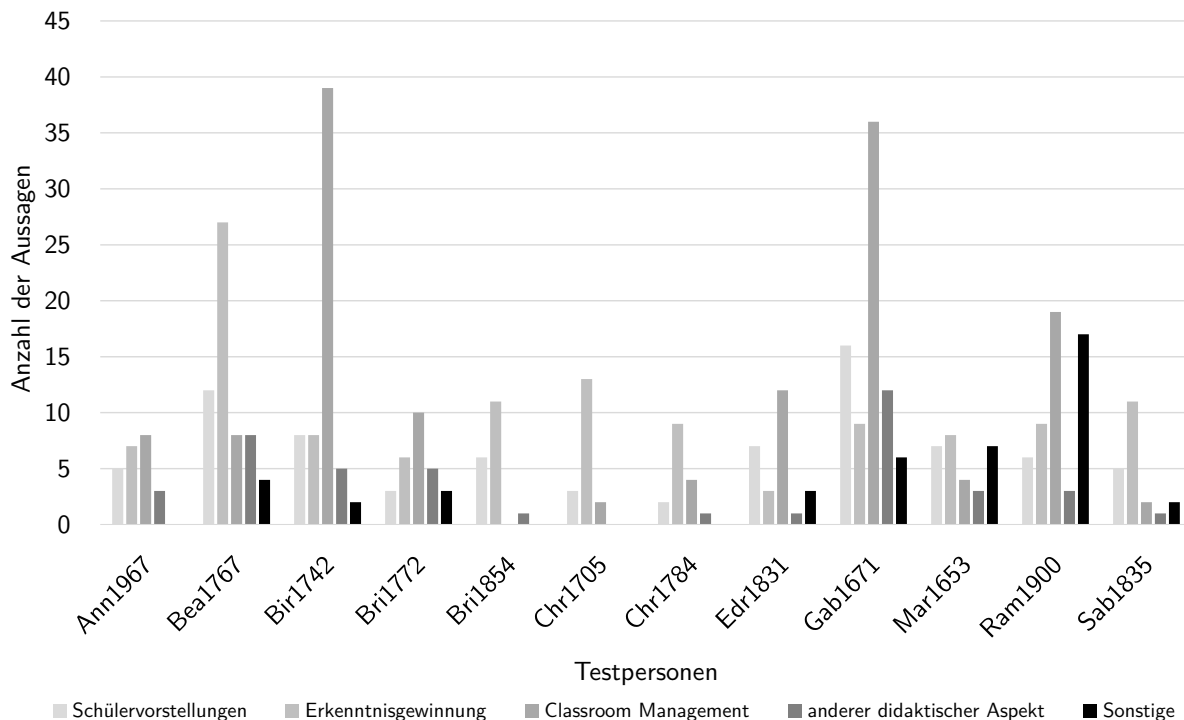


Abbildung 6.3: Verteilung der Aussagen aller Testpersonen der Entwicklungsstudie II.1 auf die Kategorien.

ANN1967: Es wurden insgesamt 23 Situationen in den Videos 1 und 4 wahrgenommen. Keine der Aussagen konnte in die Kategorie „Sonstige“ eingeordnet werden. Die Verteilung der Kategorien in den Aussagen scheint einen Schwerpunkt bezüglich der Kategorie 2 (Erkenntnisgewinn) und 3 (Classroom Management) zu haben. ANN1867 ist im sechsten Bachelorsemester im Lehramtsstudium Physik.

BEA1767: In den Videos 1 und 3 wurden insgesamt 59 Situationen wahrgenommen. Mit 27 Äußerungen zur Kategorie 2 liegt ein klarer Schwerpunkt auf der Wahrnehmung von Erkenntnisgewinnungsprozessen, gefolgt von 12 Aussagen zu Schülervorstellungen. Damit beantworten 66,1% der Aussagen die Aufgabe, die eigene Wahrnehmung bezüglich Schülervorstellungen und Erkenntnisgewinnung zu äußern. BEA1767 ist im vierten Mastersemester im Lehramtsstudium Physik.

BIR1742: Diese Testperson äußerte ihre Wahrnehmung zu 62 Situationen in Video 3 und 4. Dabei liegt anscheinend ein eindeutiger Schwerpunkt auf der Wahrnehmung von Situationen, in denen Aspekte der Klassenführung lernrelevant zu sein scheinen, da circa 63% der Aussagen in die Kategorie 3 eingeordnet

werden konnten. Die Kategorien 1 und 2 werden jeweils nur mit acht Aussagen versehen. BIR1742 ist im vierten Bachelorsemester im Lehramtsstudium Geographie.

BRI1772: Es wurden insgesamt 27 Situationen der Videos 1 und 4 als lernrelevant durch die Aussagen der Testperson markiert. Davon betreffen mehr als ein Drittel Aspekte des Classroom Managements. Die Kategorien 1 und 2 umfassen drei und sechs Äußerungen und damit zusammen ebenfalls ein Drittel der Aussagen. BRI1772 ist im vierten Mastersemester im Lehramtsstudium Physik.

BRI1854: In den Videos 3 und 4 wurden insgesamt 18 Situationen als lernrelevant wahrgenommen. Von diesen 18 Äußerungen betreffen sechs Aussagen Aspekte von Schülervorstellungen und 11 Aussagen Aspekte von Erkenntnisgewinnung. Somit äußert die Testperson fast nur Wahrnehmungen entsprechend der gegebenen Aufgabe, aber insgesamt wenig im Vergleich zu den anderen Testpersonen. BRI1854 ist im zweiten Bachelorsemester im Lehramtsstudium Physik (Berufsschule).

CHR1705: Diese Testperson äußerte Wahrnehmung zu 18 Situationen der Videos 3 und 4. Es handelt sich dabei in 13 der Fälle um Erkenntnisgewinnungsprozesse. Auch CHR1705 äußert damit verhältnismäßig selten ihre Wahrnehmung zu den Videos 3 und 4. CHR1705 studiert im sechsten Bachelorsemester im Lehramtsstudium Physik.

CHR1784: Es wurden insgesamt 16 Situationen in Video 1 und 3 wahrgenommen. Davon betreffen neun Aussagen die Erkenntnisgewinnung. CHR1784 studiert im sechsten Bachelorsemester Geographie.

ERD1831: In den Videos 1 und 4 wurden insgesamt 26 Situationen wahrgenommen. Dabei betreffen 12 Aussagen Aspekte des Classroom Managements, sieben Äußerungen Aspekte der Schülervorstellungen und drei Aussagen Erkenntnisgewinnung der Lernenden. ERD1831 ist Lehrkraft für Physik mit fünf Jahren Berufserfahrung.

GAB1761: Diese Testperson hat mit 79 Äußerungen die meisten Aussagen zu Video 1 und 3 bezüglich ihrer Wahrnehmung mitgeteilt. Am häufigsten betreffen diese Aussagen Aspekte des Classroom Managements (36 Aussagen). Aber 16 Aussagen adressieren außerdem Schülervorstellungen und neun Aussagen Erkenntnisgewinnung der Lernenden. GAB1761 ist eine Lehrkraft für Physik mit mehr als fünf Jahren Berufserfahrung und zusätzlich Erfahrung in der Betreuung von (angehenden) Lehrkräften.

MAR1653: Es wurden insgesamt 29 Situationen in den Videos 3 und 4 wahrgenommen. Die Verteilung ist dabei breit: Sieben Aussagen zu Schülervorstellungen, acht Aussagen zu Erkenntnisgewinnung, vier Äußerungen zu Aspekten des Classroom Managements und drei Aussagen zu anderen didaktischen Aspekten. Außerdem wurden sieben Aussagen der Kategorie „Sonstige“ zugeordnet. Dies waren

vor allem Aussagen zu Fragen bezüglich des Umgangs mit dem Instrument, wie beispielsweise: „Kann ich noch ein bisschen zurück? Weil ich war // I: Mhm. // nee. ok“ oder Vorlesen der Aspekte der Bildungsstandards. MAR1653 ist eine Lehrkraft im Vorbereitungsdienst für den Physikunterricht.

RAM1900: In Video 1 und 3 wurden 54 Aussagen zur Wahrnehmung getätigt. Der Großteil dieser Aussagen betreffen Aspekte des Classroom Managements (19 Äußerungen) und die Aussagen der Kategorie „Sonstiges“ (17 Äußerungen). Letzteres betraf häufig, dass die Testperson nicht verstanden hat, worum es im Video ging: „Ich bin mir jetzt nicht sicher, ob ich jetzt den Zusammenhang verstanden habe, das lag aber wahrscheinlich akustisch dran.“ RAM1900 ist eine Lehrkraft im Vorbereitungsdienst für den Physikunterricht.

SAB1835: Es wurden insgesamt 21 Situationen in Video 1 und 4 wahrgenommen, von denen 11 Äußerungen Situationen mit Bezug zu Erkenntnisgewinnungsprozessen der Lernenden betreffen. Fünf Aussagen handeln von Schülervorstellungen. SAB1835 studiert im zweiten Semester im Lehramtsstudium Physik.

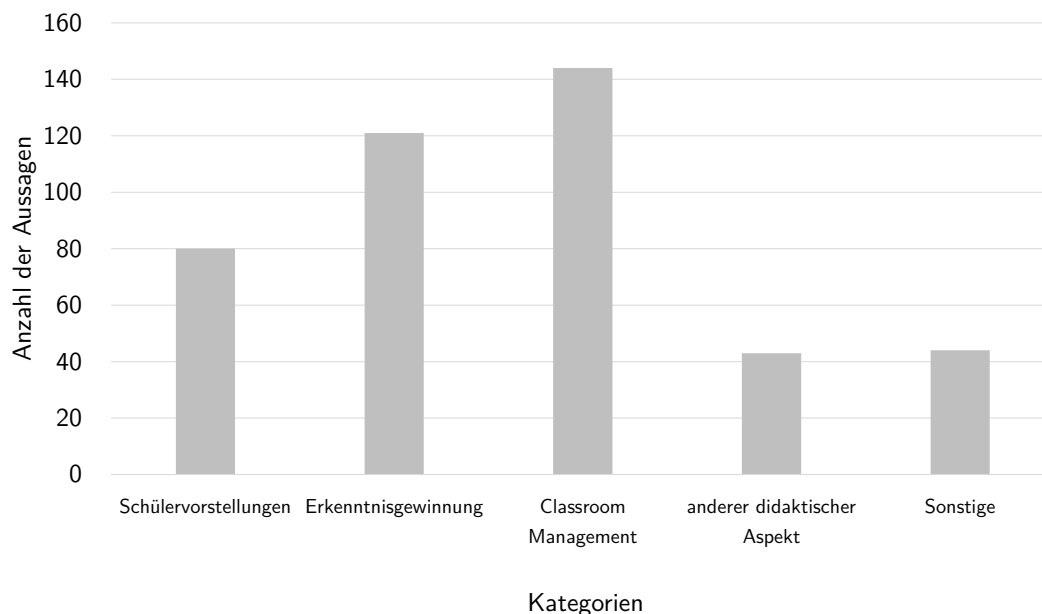


Abbildung 6.4: Verteilung der Aussagen der Testpersonen in der Entwicklungsstudie II.1 auf die Kategorien.

Die Verteilung über das ganze Datenmaterial auf die Kategorien kann folgendermaßen zusammengefasst werden (siehe auch Abbildung 6.4): Kategorie 1 (Schülervorstellungen) wurde insgesamt mit 80 Aussagen adressiert, Kategorie 2 (Erkenntnisgewinnung) mit 121, Kategorie 3 (Classroom Management) mit 144, Kategorie 4 (anderer didaktischer Aspekt) mit 43 und Kategorie 5 (Sonstige) mit 44 (vergleiche Abbildung 6.4). Insgesamt haben alle Testpersonen zusammen 432 Aussagen getätigt. Das bedeutet, dass Aspekte des Classroom Managements mit einem Drittel aller Aussagen am häufigsten mindestens

erkannt wurden und andere didaktische Aspekte mit circa zehn Prozent am seltensten (abgesehen von der Kategorie „Sonstige“).

6.1.3.2 Interpretation und Diskussion – Entwicklung von Distraktoren

In diesem Unterkapitel werden die Interpretationen der Ergebnisse präsentiert. Den einzelnen Interpretationen folgen Diskussionen, die aufzeigen, welche Befunde aus dieser Teilstudie die folgenden Schritte beeinflussten.

Das entwickelte Kodiersystem diente zur Entwicklung der Distraktoren für einen geschlossenen Fragebogen im PUPiL-Instrument der Hauptstudie. Der geschlossene Fragebogen betrifft nur diejenigen Situationen, die in der Entwicklungsstudie I für valide erklärt wurden.

Die Kategorie „anderer didaktischer Aspekte“ beinhaltet viele Teilaspekte der Fachdidaktik wie Kommunikation, Heterogenität oder Gender im Physikunterricht und wirkt damit wenig spezifisch. Die Intercoderübereinstimmung mit 0.968 ist allerdings zufriedenstellend. Außerdem liefern die Aussagen der Testpersonen zu dieser Kategorie vielfältige Distraktoren für den geschlossenen Fragebogen.

Es kann kritisch betrachtet werden, dass auch die Aussagen kodiert wurden, die die Testpersonen zu nicht-kritischen Situationen tätigten. Diese konkreten Aussagen gelangen nicht in das PUPiL-Instrument, können aber hilfreiche Informationen enthalten, wie Distraktoren zu den kritischen Situationen inhaltlich und strukturell aufgebaut sein sollten, um für die anvisierte Stichprobe überzeugend zu sein und somit die Validität des PUPiL-Instruments zu erhöhen.

In der Betrachtung der Antwortmuster der einzelnen Testpersonen lässt sich beobachten, dass bei der Hälfte der Testpersonen ein Fokus auf der Wahrnehmung von Aspekten des Classroom Managements zu liegen scheint. Dieser Umstand ist zwar hilfreich bei der Entwicklung von attraktiven Distraktoren, weist aber zugleich darauf hin, dass die Testpersonen sich in dem Test nicht rein auf ihre physikspezifische Wahrnehmung konzentrieren (Top-Down-Prozess). Situationen, in denen Aspekte des Classroom Managements lernrelevant werden, scheinen die Aufmerksamkeit der Testpersonen auf sich zu lenken (Bottom-Up-Prozess).

Bei der kleinen Stichprobe kann noch keine Aussage darüber getroffen werden, ob bestimmte Gruppen von Lehrenden eher Classroom Management-Aspekte oder Schülervorstellungen beziehungsweise Erkenntnisgewinnung wahrnehmen. Allerdings ist auffällig, dass alle praktizierenden Lehrkräfte mehr Aussagen zu Aspekten des Classroom Managements tätigten als zu einem anderen Unterrichtsaspekt und die Studierenden am Anfang des Studiums wenig, aber aufgabengerecht wahrnehmen. Des Weiteren legt scheinbar keine der Testpersonen einen eindeutigen Fokus auf Schülervorstellungen, obwohl es in den Videos insgesamt mehr kritische Situationen gibt, in denen Schülervorstellungen relevant werden als Situation zur Erkenntnisgewinnung.

Bei drei Testpersonen ist auffällig, dass mehr als fünf Aussagen der Kategorie „Sonstige“ zugeordnet wurden. Einige der Äußerungen weisen darauf hin, dass die Testpersonen

im Umgang mit dem Instrument nicht sicher waren oder (inhaltliche oder akustische) Verständnisprobleme mit den Videos hatten. In Verbindung mit der analysierten Schwierigkeit (vergleiche Unterunterabschnitt 6.1.2.1) führen diese Aussagen zu der Vermutung, dass der Test insgesamt zu schwer ist, was mit dem ungeübten und nicht alltäglichen Umgang mit den Videos im Instrument zusammenhängen kann.

Um abzusichern, dass die eher geringe Passung des Instruments zur Stichprobe an dem unbekannten Umgang mit den Videos liegt, wurde die in diesem Kapitel beschriebene Untersuchung mit ähnlichen Testpersonen erneut durchgeführt. Dabei wurde eine Übung vorangestellt, die den Umgang mit den Videos schulen soll.

6.2 Entwicklungsstudie II.2

6.2.1 Methode

An die Ergebnisse der Entwicklungsstudie II.1 anschließend liegen ein Interviewleitfaden und drei validierte Videos vor. Der Interviewleitfaden wurde insofern überarbeitet, als dass das laute Denken während der Videos in einer Übung angeleitet wird. Neben der Frage, inwiefern diese Abänderung zu einer besseren Passung des Tests zur Stichprobe führt, bleibt das Ziel aus Entwicklungsstudie II.1 auch für die Entwicklungsstudie II.2 bestehen: Beurteilung der Schwierigkeit des Tests und Entwicklung von Distraktoren. Resultate der Entwicklungsstudie II.2 sind dementsprechend Distraktoren für einen geschlossenen Fragebogen für das PUPiL-Instrument in der Hauptstudie.

Für die Durchführung konnte eine Stichprobe mit $N = 10$ Testpersonen zusammengestellt werden, welche der Stichprobe aus Entwicklungsstudie II.1 in Teilen entspricht. Die Stichprobe umfasst beispielsweise wie bereits in der Entwicklungsstudie II.1 jeweils zwei Physiklehramtsstudierende des 3. Semester (Bachelor) und 7. Semesters (Master) und zwei Physiklehrkräfte. Abweichend wurde eine Person nach Beendigung des Studiums und drei Physiklehrkräfte im Vorbereitungsdienst in die Stichprobe integriert. Allerdings umfasste die Stichprobe dieses Mal keine Geographielehramtsstudierenden (vgl. Tabelle 6.7).

Tabelle 6.7: Stichprobe der Entwicklungsstudie II.2.

Fach	3. Semester	7. Semester	Nach dem Studium	Lehrkräfte in Vorbereitung	Lehrkräfte
Physik	2	2	1	3	2

Das Vorgehen der Untersuchung bleibt nahezu identisch mit der Entwicklungsstudie II.1. Es wird aber für das laute Denken statt der Übung „Erfinden Sie fünf Verbesserungen für Ihr Handy“ eine Aufgabe eingebunden, in der die Testpersonen den Umgang mit den Videos und das laute Denken üben können. Um eine möglichst gute Wirkung auf die Performanz bei den eigentlichen Aufgaben zu erzielen, ist es wichtig, dass die Übung von

der Konstruktion her identisch ist. Da aber kein bloßer Trainingseffekt gemessen werden soll, behandelt das Übungsvideo einen anderen Inhaltsbereich. Die Testpersonen sahen ein kurzes, bearbeitetes Video, welches Straßenverkehr an einer vielbefahrenen Kreuzung in Hamburg zeigt. In dem Video verletzen Verkehrsteilnehmende einige Verkehrsregeln und bringen Radfahrende auf diese Weise in Gefahr. Die Testpersonen erhielten die Aufgabe, das Video immer dann anzuhalten, wenn sie etwas in Bezug auf die Verkehrssicherheit der Radfahrenden wahrnahmen. Anschließend sollten sie begründen, warum sie das Video angehalten haben und benennen, was sie tun würden, wenn sie diese Situation als Mitglied der Verkehrspolizei gesehen hätten. Für die ersten zwei Minuten wurde das Vorgehen von der Erhebungsleitung vorgeführt. Danach hat sich die Erhebungsleitung im Hintergrund aufgehalten und die Testpersonen üben lassen. Abschließend erhielten die Testpersonen auf Grundlage der Beobachtungen durch die Erhebungsleitung ein Feedback zu ihrem Umgang mit dem Video. Wenn alle Fragen geklärt waren, wurde die Erhebung mit den Aufgaben zu den validierten Videos wie in der Entwicklungsstudie II.1 fortgesetzt.

Die Verbaldaten der Erhebung wurden transkribiert und mithilfe der beiden Kodierschemata aus Entwicklungsstudie II.1 analysiert (vergleiche Unterabschnitt 6.1.2 und Unterabschnitt 6.1.3). Um herauszufinden, ob das Übungsvideo den Testpersonen dabei geholfen hat, ihre professionelle Unterrichtswahrnehmung anhand der Unterrichtsvideos zu zeigen, wurden die Ergebnisse der Entwicklungsstudie II.1 mit den der Entwicklungsstudie II.2 verglichen. Dazu wurden absolute Differenzen der Anzahl von Aussagen zu den verschiedenen Kategorien betrachtet. Die Befunde dieser qualitativen Analysen werden anschließend mit nichtparametrischen quantitativen Analysen ergänzt.

6.2.2 Teilanalyse A – Prüfung der Passung und Schwierigkeit des Tests

6.2.2.1 Ergebnisse – Passung und Schwierigkeit des Tests

Analog zu Unterabschnitt 6.1.2.1 kann nach einer deduktiven qualitativen Inhaltsanalyse berichtet werden, von wie vielen Testpersonen die jeweiligen Situationen in den Videos (a) erkannt, (b) benannt, (c) erläutert oder (d) in einen Handlungsplan überführt wurden (vgl. Abbildung 6.1 und Abbildung 6.5). Ob das Übungsvideo seinen Zweck erfüllt hat, kann aus den Unterschieden im Antwortverhalten der Testpersonen in den verschiedenen Stichproben abgeleitet werden. Video 1 und 4 wurden von sieben Testpersonen und Video 3 von sechs Testpersonen bearbeitet. Der Tabelle 6.8 zum Vergleich der Entwicklungsstudie II.1 und Entwicklungsstudie II.2 ist zu entnehmen, dass in der Entwicklungsstudie II.2 13 Aussagen mehr bezüglich dem *Erkennen* einer Situation geäußert wurden als in Entwicklungsstudie II.1. Es wurden im zweiten Durchlauf der Entwicklungsstudie II nur zwei Aussagen getätigt, die richtig erkannte Situationen *benennen*. Die Anteile an *erläuternden* Aussagen und solchen zum *Fassen von Handlungsplänen* steigen im Vergleich zum ersten Durchlauf der Studie an.

Tabelle 6.8: Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II bezüglich Ausprägung der professionellen Unterrichtswahrnehmung. Elemente mit „1“ bilden Ergebnisse der Entwicklungsstudie II.1 und Elemente mit „2“ Ergebnisse der Entwicklungsstudie II.2 ab.

Situation	Video 1	Video 3	Video 4	Insgesamt
erkannt 1	4	10	6	20
erkannt 2	11	10	12	33
benannt 1	3	8	7	18
benannt 2	1	1	0	2
erläutert 1	8	21	12	41
erläutert 2	21	34	10	75
Handlungsplan gefasst 1	15	5	8	28
Handlungsplan gefasst 2	14	22	31	69

Aus der Analyse der Differenzen in Tabellen 6.9 bis 6.13 resultiert, dass nach beiden Teilstudien nur noch circa 36,8% der Situationen aus den drei Videos von weniger als der Hälfte der Testpersonen aus beiden Teilstudien zusammen genommen mindestens erkannt wurden (Video 1 und 4 wurden insgesamt von 15 Testpersonen gesehen, Video 3 von 14 Testpersonen). Dies betraf im ersten Durchlauf der Entwicklungsstudie II ohne Übung noch circa 65,8% der Situationen.

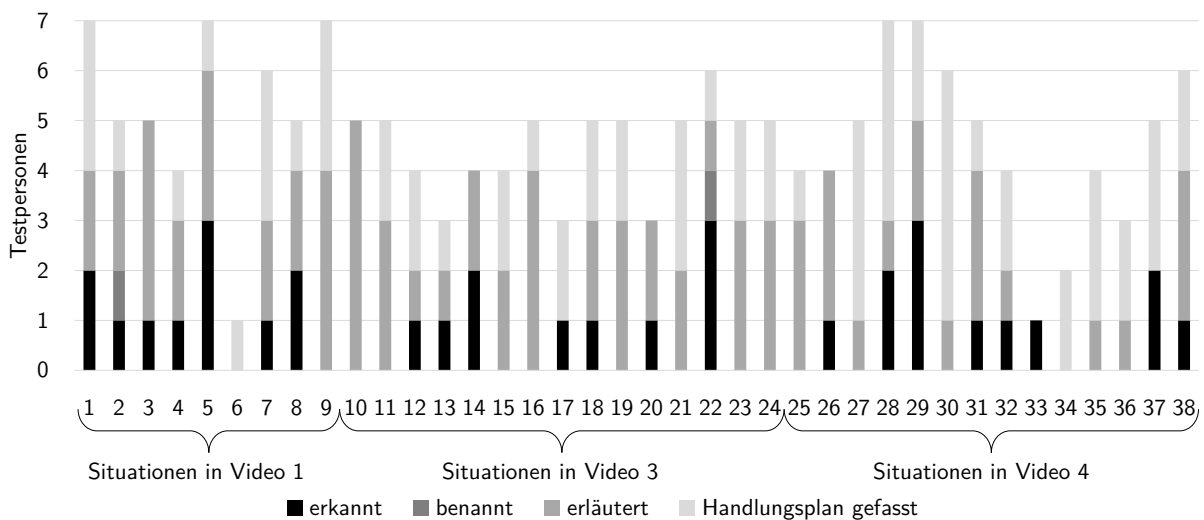


Abbildung 6.5: Ergebnisse Prüfung der Passung und Schwierigkeit des Tests Entwicklungsstudie II.2.

Tabelle 6.9: Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 1. Die Zelleneinträge repräsentieren die Anzahl an Äußerungen durch die Testpersonen. Elemente mit „1“ bilden Ergebnisse der Entwicklungsstudie II.1 und Elemente mit „2“ Ergebnisse der Entwicklungsstudie II.2 ab. Bei der Änderung handelt es sich um eine absolute Differenz.

Situation	1	2	3	4	5	6	7	8	9
erkannt 1	0	0	1	0	2	0	0	1	0
erkannt 2	2	1	1	1	3	0	1	2	0
Änderung erkannt	2	1	0	1	1	0	1	1	0
benannt 1	0	1	0	0	1	0	0	0	1
benannt 2	0	1	0	0	0	0	0	0	0
Änderung benannt	0	0	0	0	-1	0	0	0	-1
erläutert 1	0	0	2	1	0	1	2	0	2
erläutert 2	2	2	4	2	3	0	2	2	4
Änderung erläutert	2	2	2	1	3	-1	0	2	2
Handlungsplan gefasst 1	3	3	2	1	4	1	0	0	1
Handlungsplan gefasst 2	3	1	0	1	1	1	3	1	3
Änderung Handlungsplan	0	-2	-2	0	-3	0	3	1	2

Tabelle 6.10: Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 2. Die Zelleneinträge repräsentieren die Anzahl an Äußerungen durch die Testpersonen. Elemente mit „1“ bilden Ergebnisse der Entwicklungsstudie II.1 und Elemente mit „2“ Ergebnisse der Entwicklungsstudie II.2 ab. Bei der Änderung handelt es sich um eine absolute Differenz.

Situation	10	11	12	13	14	15	16	17
erkannt 1	0	0	0	1	0	1	1	1
erkannt 2	0	0	1	1	2	0	0	1
Änderung erkannt	0	0	1	0	2	-1	-1	0
benannt 1	1	0	0	2	0	0	0	1
benannt 2	0	0	0	0	0	0	0	0
Änderung benannt	-1	0	0	-2	0	0	0	-1
erläutert 1	1	0	1	1	1	4	2	1
erläutert 2	5	3	1	1	3	3	4	0
Änderung erläutert	4	3	0	0	1	-2	2	-1
Handlungsplan gefasst 1	0	1	1	0	0	0	1	0
Handlungsplan gefasst 2	0	2	2	1	0	2	1	2
Änderung Handlungsplan	0	1	1	1	0	2	0	2

Tabelle 6.11: Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 3. Die Zelleneinträge repräsentieren die Anzahl an Äußerungen durch die Testpersonen. Elemente mit „1“ bilden Ergebnisse der Entwicklungsstudie II.1 und Elemente mit „2“ Ergebnisse der Entwicklungsstudie II.2 ab. Bei der Änderung handelt es sich um eine absolute Differenz.

Situation	18	19	20	21	22	23	24
erkannt 1	1	0	1	1	1	0	2
erkannt 2	1	0	1	0	3	0	0
Änderung erkannt	0	0	0	-1	2	0	-2
benannt 1	0	1	0	0	1	1	1
benannt 2	0	0	0	0	1	0	0
Änderung benannt	0	-1	0	0	0	-1	-1
erläutert 1	2	2	1	1	1	3	0
erläutert 2	2	3	2	2	1	3	3
Änderung erläutert	0	1	1	1	0	0	3
Handlungsplan gefasst 1	1	1	0	0	0	0	0
Handlungsplan gefasst 2	2	2	0	3	1	2	2
Änderung Handlungsplan	1	1	0	3	1	2	2

Tabelle 6.12: Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 4. Die Zelleneinträge repräsentieren die Anzahl an Äußerungen durch die Testpersonen. Elemente mit „1“ bilden Ergebnisse der Entwicklungsstudie II.1 und Elemente mit „2“ Ergebnisse der Entwicklungsstudie II.2 ab. Bei der Änderung handelt es sich um eine absolute Differenz.

Situation	25	26	27	28	29	30	31
erkannt 1	1	2	0	0	0	0	0
erkannt 2	0	1	0	2	3	0	1
Änderung erkannt	-1	-1	0	2	3	0	1
benannt 1	0	2	0	1	0	1	1
benannt 2	0	0	0	0	0	0	0
Änderung benannt	0	-2	0	-1	0	-1	-1
erläutert 1	0	0	2	0	1	2	1
erläutert 2	3	3	1	1	2	1	3
Änderung erläutert	3	3	-1	1	1	-1	2
Handlungsplan gefasst 1	0	0	1	0	1	1	1
Handlungsplan gefasst 2	1	0	4	4	2	5	1
Änderung Handlungsplan	1	0	3	4	1	4	0

Tabelle 6.13: Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 5. Die Zelleneinträge repräsentieren die Anzahl an Äußerungen durch die Testpersonen. Elemente mit „1“ bilden Ergebnisse der Entwicklungsstudie II.1 und Elemente mit „2“ Ergebnisse der Entwicklungsstudie II.2 ab. Bei der Änderung handelt es sich um eine absolute Differenz.

Situation	32	33	34	35	36	37	38
erkannt 1	0	0	0	1	1	0	1
erkannt 2	1	1	0	0	0	2	1
Änderung erkannt	1	1	0	-1	-1	2	0
benannt 1	0	0	0	1	0	0	1
benannt 2	0	0	0	0	0	0	0
Änderung benannt	0	0	0	-1	0	0	-1
erläutert 1	0	2	0	0	0	2	2
erläutert 2	1	0	0	1	1	0	3
Änderung erläutert	1	-2	0	1	1	-2	1
Handlungsplan gefasst 1	1	1	0	0	1	0	1
Handlungsplan gefasst 2	2	0	2	3	2	3	2
Änderung Handlungsplan	1	-1	2	3	1	3	1

Außerdem ist der Anstieg an mindestens erkannten Situationen durch mehr als die Hälfte der Testpersonen in Entwicklungsstudie II.2 groß: In Video 1 wurden acht der neun Situationen mindestens erkannt, in Video 3 wurden alle 15 Situationen mindestens erkannt und in Video 4 wurden 11 der 14 Situationen von mehr als der Hälfte der Testpersonen mindestens erkannt (vgl. Abbildung 6.5). Den Tabellen ist außerdem zu entnehmen, dass insgesamt 30 Situationen ein bis sechs mal mehr mindestens erkannt wurden als in Entwicklungsstudie II.1, vier Situationen wurden ein bis zwei mal seltener mindestens erkannt und zu vier Situationen hat sich die Häufigkeit der Nennung in den Verbaldaten nicht verändert.

An dieser Stelle wird die Entwicklung der Ergebnisse exemplarisch an einigen Situationen genauer betrachtet. Alle weiteren Ergebnisse sind in Abbildung 6.5 und in den Tabellen 6.9 bis 6.13 auf die gleiche Weise abzulesen:

- Es gibt vier Situationen, zu denen die Anzahl der Adressierungen durch Testpersonen der Entwicklungsstudie II.1 zu II.2 *gleich bleibt* (Situationen 3 und 5 von Video 1, Situation 8 von Video 3 und Situation 2 von Video 4). Obwohl die Anzahl der Nennungen gleich geblieben ist, haben sich die Verteilungen geändert. Deshalb folgt nun eine genauere Sichtung der Zusammenstellung entlang der Kategorien.

Zu Situation 3 aus Video 1 gab es insgesamt fünf Nennungen. Im ersten Durchlauf verteilten sich diese auf das Erkennen (1), Erläutern (2) und Fassen eines Handlungsplans (2). Im Vergleich dazu wurde kein Handlungsplan der Testpersonen aus dem zweiten Durchgang gefasst, dafür wurde die Situation aber vier Mal erläutert. Bei Situation 5 ist die Umverteilung der Nennungen ebenfalls zulasten des Fassens eines Handlungsplans gegangen: Statt vier Testpersonen, äußert nur noch eine Testperson einen Handlungsplan. Dafür wird die Situation häufiger erkannt und erläutert.

Bei Situation 8 aus Video 3 (in der Tabelle 6.10 Situation 17) hingegen verteilen sich die Nennungen hin zum Fassen eines Handlungsplans um.

Zu Situation 2 in Video 4 (in Tabelle 6.12 Situation 26) hat weder im ersten noch im zweiten Durchgang eine Testperson einen Handlungsplan entwickelt, dafür steigt im zweiten Durchgang aber die Anzahl der Erläuterungen von null auf drei.

- Es gibt vier Situationen, zu denen sich die Anzahl der Adressierungen durch die Testpersonen vom ersten zum zweiten Durchgang *verringert* hat (Situationen 6 aus Video 1, Situationen 4, 6 aus Video 3 und Situation 9 aus Video 4).

Situation 6 wurde insgesamt nur drei mal mindestens erkannt. Nur eine Testperson äußert sich in der Entwicklungsstudie II.2, indem sie einen Handlungsplan fasst.

In beiden Situationen aus Video 3 wurden im zweiten Durchgang mehr Handlungspläne gefasst, dafür wurden aber weniger Aussagen in anderen Kategorien getätigt. Die Situation 6 aus Video 1 und Situationen 9 sowie 10 aus Video 4 (Situation 33 und 34 in Tabelle 6.13) wurden am seltensten mindestens erkannt (insgesamt nur zwei bis vier mal). Die Situation 9 aus Video 4 wurde im zweiten Durchgang nur noch ein Mal erkannt, weitere Aussagen konnten im Gegensatz zum ersten Durchgang nicht verzeichnet werden. Zu Situation 10 aus Video 4 gab es in der Entwicklungsstudie

II.1 keine Äußerungen der Testpersonen. Im zweiten Durchgang wurden hingegen zwei Handlungspläne gefasst.

- Bezüglich 30 Situationen wurden in der Entwicklungsstudie II.2 *mehr* Aussagen geäußert, die sich vom Erkennen einer kritischen Situation hin zum Fassen eines Handlungsplans erstrecken.

Die größte Zunahme von einer Nennung auf sieben Nennungen kann für Situation 4 aus Video 4 (Situation 28 in Tabelle 6.12) verzeichnet werden. Diese positive Veränderung betrifft dabei die Kategorien erkennen (+2), erläutern (+1) und Handlungsplan fassen (+4) zulasten der Kategorie des Benennens (-1).

Ob und inwiefern die hier präsentierten Ergebnisse zu einer Besserung der Passung des PUPiL-Instruments zur intendierten Stichprobe beigetragen haben, wird im nächsten Unterkapitel diskutiert.

6.2.2.2 Interpretation und Diskussion – Passung und Schwierigkeit des Tests

In diesem Unterkapitel werden die Interpretationen der Ergebnisse präsentiert. Den einzelnen Interpretationen folgen Diskussionen, die aufzeigen, welche Befunde aus dieser Teilstudie die folgenden Schritte beeinflussten.

In Entwicklungsstudie II.1 hat sich gezeigt, dass der Umgang mit den Videos für die Testpersonen nicht zu den erwarteten Ergebnissen geführt hat. Darum wurde diese Studie mit (anderen) Testpersonen ($N = 10$) wiederholt. Allerdings bekamen die Testpersonen bei diesem Durchgang eine Möglichkeit, den Umgang mit den Videos im Voraus zu üben.

Im Vergleich zur Entwicklungsstudie II.1 wurden nun nur noch 15 der 38 Situationen von weniger als der Hälfte der Testpersonen in der Entwicklungsstudie II.2 mindestens erkannt. Das bedeutet, dass in Entwicklungsstudie II.2 mehr Situationen von mehr Testpersonen in den Fokus genommen wurden. Es lässt sich beobachten, dass die Situationen, die von weniger als der Hälfte der Testpersonen mindestens erkannt wurden, verschiedene Aspekte beinhalten: Bei vier Situationen kann ein angemessener Erkenntnisgewinnungsprozess der gezeigten Lernenden beobachtet werden (Situationen 4, 5, 6 und 8 aus Video 3). Bei zwei weiteren Situationen sind die Aktivitäten der Lernenden in Hinblick auf eine Erkenntnisgewinnung allerdings fragwürdig (Situationen 4 und 6 aus Video 1). Die Situationen 3 und 11 aus Video 3, Situationen 1, 8, 11 und 12 aus Video 4 wurden von weniger als der Hälfte der Testpersonen mindestens erkannt und betreffen inadäquate Schülervorstellungen. Nur drei Situationen, die adäquate Schülervorstellungen thematisieren, werden von weniger als der Hälfte der Testpersonen wahrgenommen (Situationen 5, 9 und 10 aus Video 4). Das deutet darauf hin, dass die Übung mit dem Video zum Verkehr oder die neu zusammengesetzte Stichprobe zu einer besseren Wahrnehmung von Situationen führt, die in Entwicklungsstudie II.1 als vermeintlich schwer interpretiert wurden. Das Übungsvideo gab allerdings nur die Möglichkeit, die Wahrnehmung misslingender Augenblicke im Verkehr zu üben. Es kann also geschlossen werden, dass das Übungsvideo die Interaktion mit den Videos im Instrument im Allgemeinen unterstützt hat und es somit zu einer

erleichterten Handhabung kam. Das Übungsvideo bleibt daher Bestandteil des PUPiL-Instruments für die Hauptstudie.

Schwerer wahrzunehmende Situationen sind auf alle Aspekte (adäquate oder inadäquate Schülervorstellungen und Erkenntnisgewinnung) verteilt. Wenn in der folgenden Hauptstudie eine Situation weniger oft wahrgenommen wird als andere, kann daraus geschlossen werden, dass diese tatsächlich schwerer wahrzunehmen ist und es nicht am Umgang mit dem PUPiL-Instrument liegt.

Ein weiteres Indiz dafür, dass das Übungsvideo seinen Zweck erfüllt, ist die erhöhte Anzahl an „erkannten“ Situationen. Die Anzahl der Aussagen, die ein „Erkennen“ von kritischen Situationen andeuten, ist um 13 Aussagen gestiegen. Das bedeutet nicht, dass vorher mehr Situationen unerkannt blieben, sondern dass nun mehr Testpersonen wenigstens erkannt haben, dass eine Situation etwas Wahrnehmungsbedürftiges an sich hat. Es kann geschlussfolgert werden, dass das Übungsvideo einen Trainingseffekt bewirkt, da mehr mit den Videos des Instruments interagiert wurde. Trotzdem gab es Testpersonen, die sich über den Umgang mit den Unterrichtsvideos nicht ganz im Klaren waren. Beispielsweise fragte eine Testperson während des lauten Denkens: „[...] Ist das okay, wenn ich das so mache, wie jetzt?“

Eine weitere Testperson rechtfertigte nach dem Sehen des ersten Videos ihre ausführlichen Redebeiträge folgendermaßen:

„Ja, ich glaub schon. //I: Okay.// Nochmal alles sehr sehr krass zusammengefasst wiederholt, aber... //I: Ja.// Ich dachte, falls ich mal was vergessen hab, ist es jetzt klar.“

Diese Aussagen weisen darauf hin, dass sich Testpersonen in der Erhebungssituation unter anderem wie in einer Prüfungssituation fühlen könnten. Ihre Aufmerksamkeit für die Videos wird damit geteilt. Es werden kognitive Ressourcen dafür aufgebracht, über das eigene Sprechen und dessen Inhalt nachzudenken.

Auch wenn es immer noch Testpersonen gibt, die vor allem das Handeln der im Video gezeigten Lehrkraft kritisieren, scheint die Übung mit dem Verkehrsvideo zumindest bei einer Testperson den Fokus der geäußerten Wahrnehmung verlegt zu haben: „Ich fin- (.) ach so. Den Lehrer beurteile ich nicht, nä.“

Insgesamt scheint der Umgang mit den Videos für einige Testpersonen vertrauter geworden zu sein.

Es lässt sich des Weiteren beobachten, dass nur noch zwei Aussagen der Testpersonen rein beschreibend sind, das entspricht einer Abnahme von 16 Aussagen vom ersten zum zweiten Durchgang der Entwicklungsstudie II. Dieser Rückgang ist positiv zu bewerten, weil gleichzeitig mehr Aussagen vorliegen, die erkannte Situationen sowohl beschreiben als auch erläutern (34 zusätzliche Aussagen) und zu diesen Handlungspläne fassen (41 zusätzliche Aussagen), obwohl im zweiten Durchgang der Entwicklungsstudie II nur zehn statt 12 Testpersonen teilgenommen haben. Möglicherweise ist die Unterweisung, dass die Testpersonen bei dem Verkehrsvideo die Gefahr für die Radfahrenden erläutern sollen, für diese Veränderung verantwortlich. Allerdings zeigt sich wie in der Entwicklungsstudie II.1, dass

das Fassen eines Handlungsplans nach wie vor, vor allem für die Lehramtsstudierenden, durch die eingesetzte Methode nicht zwingend an den Wahrnehmungsprozess anschließt. Darum wird für die Hauptstudie das Konstrukt der professionellen Unterrichtswahrnehmung auf Grundlage der empirischen Erfahrungen nur als Perception und Interpretation erfasst. Die Testpersonen werden gebeten, in einem freien Textfeld einen Handlungsplan zu fassen, nachdem sie ihre Interpretation dokumentiert haben. Es werden für das Fassen eines Handlungsplans keine Distraktoren entwickelt, um in Zukunft möglicherweise mehr empirische Erkenntnisse über die Art und Angemessenheit der Handlungspläne gewinnen zu können. Für die Auswertung der professionellen Unterrichtswahrnehmung werden aber nur die Antworten zu den Multiple-Choice-Fragen berücksichtigt. Bezüglich der Inhaltsvalidität bedeutet dies, dass die geschlossenen Fragen als Items nur die ersten beiden Dimensionen der professionellen Unterrichtswahrnehmung (Perception, Interpretation) erfassen können. Somit sind die einzelnen Items nicht repräsentativ für die Gesamtheit an Items, die nötig wären, um die professionelle Unterrichtswahrnehmung im ursprünglichen Sinn (Perception, Interpretation und Decision Making) zu messen. Eine Einschränkung des Konstrukts der professionellen Unterrichtswahrnehmung scheint aber auf Grundlage der bisherigen empirischen Befunde notwendig. Somit bilden die geschlossenen Fragen einen Teil der nötigen Items, um die zweidimensionale professionelle Unterrichtswahrnehmung zu messen (vgl. Abschnitt 4.2 zur Validität).

Das prinzipielle Vorgehen der Studie wurde bis auf den Einsatz eines Übungsvideos nicht geändert. Es konnte nach wie vor nur diejenige Wahrnehmung gemessen werden, die ins Bewusstsein der Testpersonen gelangt ist, von ihnen willentlich und von sich aus kommuniziert wurde. Damit bleibt der Teil der situationsspezifischen Fähigkeit, der nicht ins Bewusstsein der Testpersonen dringt, noch im Verborgenen. Das PUPiL-Instrument ist dennoch in der Lage, die professionelle Unterrichtswahrnehmung valider zu messen, als die oben vorgestellten Instrumente (vgl. Abschnitt 3.2 zu den bestehenden Instrumenten zur Erfassung professioneller Unterrichtswahrnehmung). Diese Instrumente deuten aber bereits einen Wandel von Paper-Pencil-Tests hin zu performanznäheren Tests an. Nur dank dieser Entwicklung war es möglich, in dieser Studie noch näher an eine valide Erhebung der (physikspezifischen) professionellen Unterrichtswahrnehmung zu gelangen.

Nach dieser Analyse kann die Passung von Instrument und Stichprobe als angemessen gewertet werden. Außerdem lag nun mehr Datenmaterial zu jeder kritischen Situation vor, auf deren Grundlage Distraktoren formuliert werden konnten. Für Situationen wie Situation 9 aus Video 4 werden Distraktoren aus dem Datenmaterial zu anderen Situationen abgeleitet. Diese weisen dann eine ähnliche Struktur auf, wurden aber für die jeweilige Situation angepasst. Die Entwicklung der Distraktoren wird im nächsten Unterkapitel vorgestellt.

6.2.3 Teilanalyse B – Entwicklung von Distraktoren

6.2.3.1 Ergebnisse – Entwicklung von Distraktoren

Die Entwicklung der Distraktoren für den geschlossenen Fragebogen des PUPiL-Instruments in der Hauptstudie basiert auf dem Datenmaterial von 22 Testpersonen mit Anteilen des lauten Denkens und dem Kodiersystem, das auf der Grundlage von 12 dieser Erhebungen entwickelt wurde (vergleiche Unterunterabschnitt 6.1.3.1). Nach der Kodierung des gesamten Datenmaterials mit dem oben beschriebenen Kodiersystem, wurden die Daten weiter entlang der zusammenfassenden Inhaltsanalyse nach Mayring (2015, S. 67-83) reduziert.

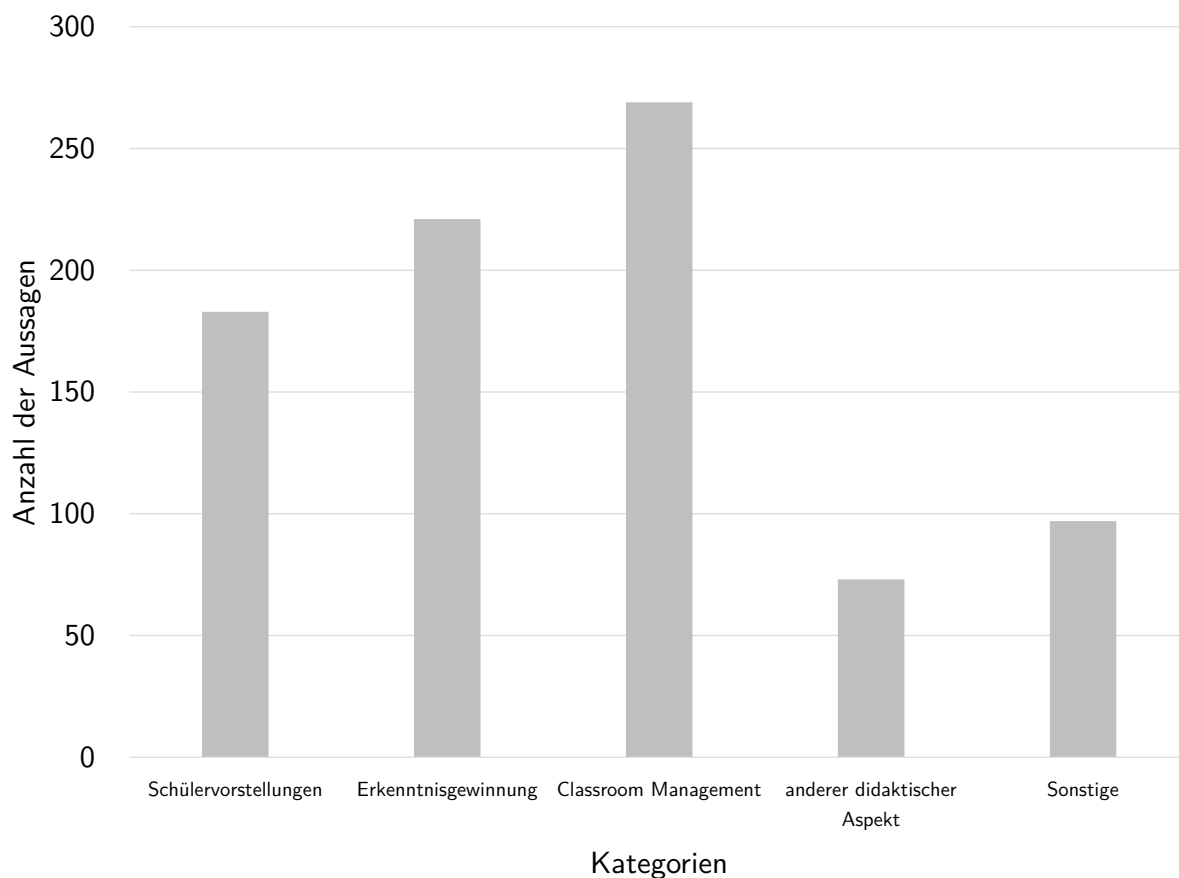


Abbildung 6.6: Verteilung der Aussagen auf die Kategorien in Entwicklungsstudie II.1 und II.2.

Es wurden insgesamt 746 Aussagen³⁷ der Testpersonen als relevant erachtet³⁸ und in mindestens eine der vier Kategorien einsortiert. Wie in Abbildung 6.6 abgebildet, ergab die

³⁷Die Anzahl der Aussagen ist in dieser Entwicklungsstudie II mit der Anzahl der Kodiereinheiten für das entsprechende induktiv entwickelte Kodierschema gleichzusetzen. Die Anzahl der Aussagen dieser Analyse ist nicht zu vergleichen mit der Anzahl der Aussagen der Analyse zur Passung und Schwierigkeit des Instruments, da für die Distraktorenentwicklung alle (auch zu nicht kritischen Situationen gehörende) Aussagen gezählt wurden.

³⁸Das sind alle Aussagen außer diejenigen, die in der Kategorie „Sonstige“ einsortiert sind.

Kodierung, dass von den Äußerungen durch die 22 Testpersonen 183 Schülervorstellungen, 221 Erkenntnisgewinnung, 269 Classroom Management und 73 andere didaktische Aspekte betrafen.

Die Kodierung der Verbaldaten des zweiten Durchgangs zeigt wieder, dass die verschiedenen Testpersonen auch verschiedene Fokusse in ihrer Wahrnehmung zu legen scheinen (siehe Abbildung 6.7):

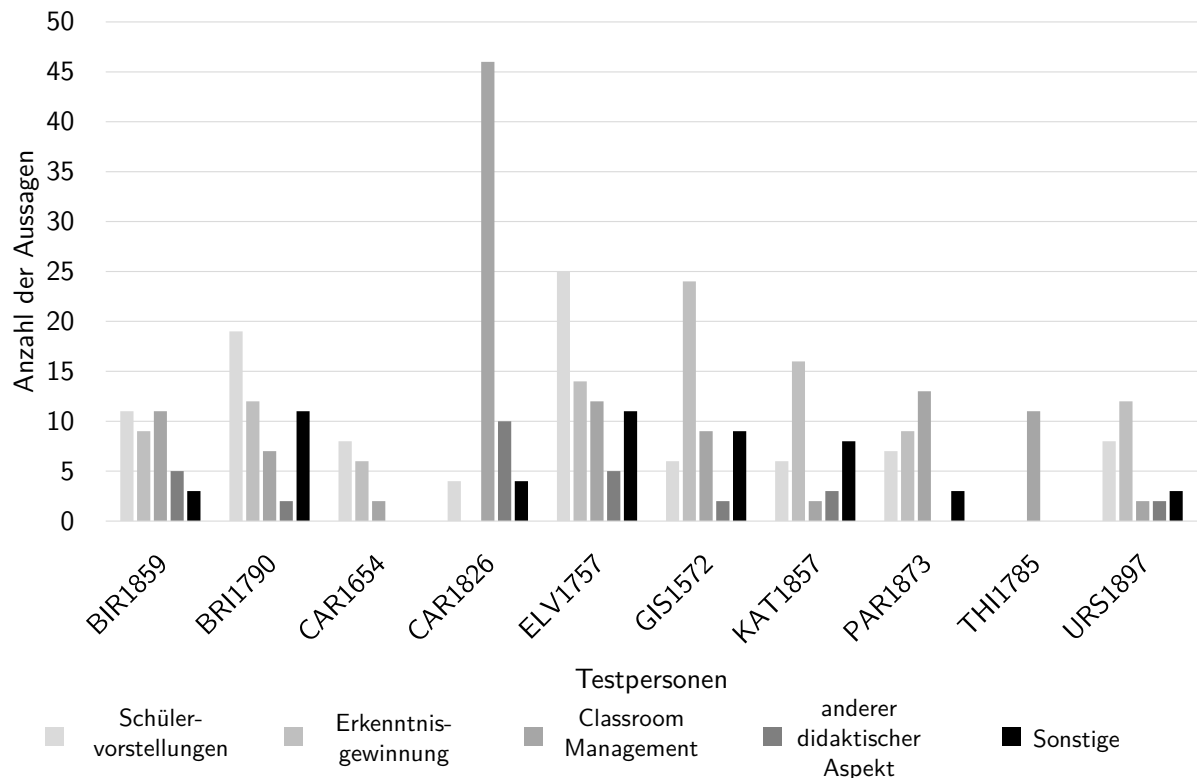


Abbildung 6.7: Verteilung der Aussagen aller Testpersonen aus der Entwicklungsstudie II.2 auf die Kategorien.

BIR1859: Es wurden insgesamt 39 Situationen in den Videos 1 und 4 wahrgenommen. Die Aussagen verteilen sich dabei relativ gleichmäßig auf die Kategorien Schülervorstellungen, Erkenntnisgewinnung und Classroom Management. BIR1859 ist im dritten Bachelorsemester im Lehramtsstudium Physik.

BRI1790: In den Videos 1 und 4 wurden insgesamt 51 Situationen wahrgenommen. Mit 19 Äußerungen zur Schülervorstellungen und 12 Aussagen zu Erkenntnisgewinnung liegt ein Schwerpunkt auf der Wahrnehmung dieser Prozesse, gefolgt von 11 Aussagen zu Sonstigem. Damit bearbeiten 54,4% der Aussagen die Aufgabe, die eigene Wahrnehmung bezüglich Schülervorstellungen und Erkenntnisgewinnung zu äußern. BRI1790 ist im siebten Bachelorsemester im Lehramtsstudium Physik.

CAR1654: Diese Testperson äußerte ihre Wahrnehmung zu 16 Situationen in Video 1 und 4. Dabei liegt anscheinend ein Schwerpunkt auf der Wahrnehmung von Si-

tuationen, in denen Aspekte der Schülervorstellungen und Erkenntnisgewinn lernrelevant zu sein scheinen, da circa 87,5% der Aussagen in die Kategorien 1 und 2 eingeordnet werden konnten. Die Kategorie 3 wird nur mit zwei Aussagen versehen. CAR1654 hat das Lehramtsstudium Physik gerade beendet.

CAR1826: Es wurden insgesamt 64 Situationen in den Videos 1 und 4 wahrgenommen. Nur vier Äußerungen bearbeiten die Aufgabe angemessen. Von den 64 Aussagen betreffen 71,9% Situationen, in denen Klassenraumsituationen lernrelevant scheinen. Es werden außerdem 10 Aussagen auf andere didaktische Aspekte bezogen. Keine andere Testperson äußert sich so häufig zu dieser Kategorie. CAR1826 ist Lehrkraft im Vorbereitungsdienst für Physik.

ELV1757: In den Videos 3 und 4 wurden insgesamt 67 Situationen wahrgenommen. Mit 58,2% der Äußerungen zu Schülervorstellungen und Erkenntnisgewinnung liegt ein Schwerpunkt auf der geforderten Wahrnehmung der Prozesse, gefolgt von 12 Aussagen zu Classroom Management. ELV1757 ist Lehrkraft im Vorbereitungsdienst für Physik.

GIS1572: Diese Testperson äußerte ihre Wahrnehmung zu 50 Situationen in Video 1 und 3. Dabei liegt anscheinend ein eindeutiger Schwerpunkt auf der Wahrnehmung von Situationen, in denen Aspekte der Erkenntnisgewinnung lernrelevant zu sein scheinen, da 48% der Aussagen in die Kategorie 2 eingeordnet werden konnten. Die übrigen Kategorien wurden jeweils mit weniger als zehn Aussagen versehen. GIS1572 ist eine Physiklehrkraft mit mehr als fünf Jahre Berufserfahrung.

KAT1857: Es wurden insgesamt 35 Situationen in den Videos 1 und 3 wahrgenommen. Mit 16 Aussagen wurden auch von dieser Testperson Situationen, in denen Erkenntnisgewinnung lernrelevant wird am häufigsten wahrgenommen. KAT1857 ist Lehrkraft im Vorbereitungsdienst für Physik.

PAR1873: In den Videos 3 und 4 wurden insgesamt 32 Situationen wahrgenommen. Mit 13 Äußerungen zu Classroom Management und 16 Aussagen zu Erkenntnisgewinnung und Schülervorstellungen zusammen ist die Wahrnehmung dieser Testperson relativ gleichmäßig verteilt. Damit bearbeiten 50% der Aussagen die Aufgabe, die eigene Wahrnehmung bezüglich Schülervorstellungen und Erkenntnisgewinnung zu äußern. PAR1873 ist im ersten Mastersemester im Lehramtsstudium Physik (Berufsschule).

THI1785: Diese Testperson äußerte ihre Wahrnehmung zu 11 Situationen in Video 3 und 4. Dabei liegt anscheinend ein eindeutiger Schwerpunkt auf der Wahrnehmung von Situationen, in denen Aspekte der Erkenntnisgewinnung lernrelevant zu sein scheinen, da alle Aussagen in die Kategorie 2 eingeordnet werden konnten. Insgesamt hat diese Testperson mit nur 11 Aussagen am wenigsten mit den Videos interagiert. THI1785 ist im dritten Bachelorsemester im Lehramtsstudium Physik.

URS1897: Es wurden insgesamt 27 Situationen in den Videos 1 und 3 wahrgenommen. Die Aussagen betreffen dabei zum Großteil Schülervorstellungen und Erkenntnisgewinnung. Damit bearbeiten 74,1% der Aussagen die Aufgabe der Erhebung. URS1897 hat als Physiklehrkraft mehr als 15 Jahre Berufserfahrung.

Die Verteilung aller Aussagen des zweiten Durchgangs ist in Abbildung 6.8 im Vergleich zu denjenigen der Entwicklungsstudie II.1 abgebildet.

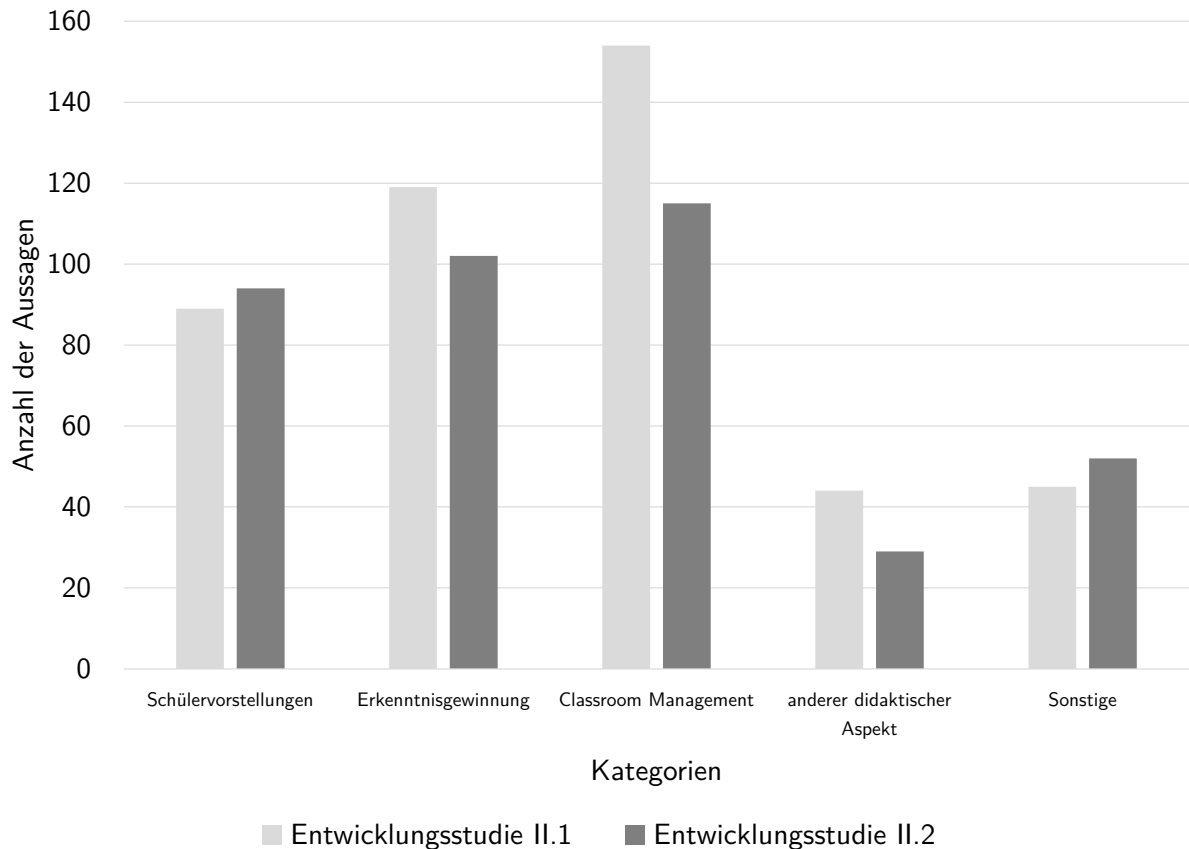


Abbildung 6.8: Vergleich der beiden Teilstudien bezüglich Verteilung der Aussagen auf die Kategorien.

Man kann in Abbildung 6.8 im Vergleich zu Abbildung 6.4 erkennen, dass es weniger Aussagen zu Erkenntnisgewinnung, Classroom Management und andere didaktische Aspekte, aber zu Situationen zu Schülervorstellungen und nicht relevanten Situationen („Sonstige“) mehr Äußerungen gibt. Ein nichtparametrischer Test auf Signifikanz der Übereinstimmung der Verteilungen von Entwicklungsstudie II.1 und II.2 ergibt allerdings, dass keiner der Unterschiede signifikant wird (vergleiche Tabelle 6.14). Für die Reduktion der Aussagen wurden zunächst alle Aussagen zu einer Kategorie je Situation zusammengefasst, sodass für jede Situation mindestens eine Aussage zu Schülervorstellungen, Erkenntnisgewinnungsprozesse, Classroom Management und anderen didaktischen Aspekten vorlag. Wenn zu einer Situation und Kategorie nur eine oder gar keine Aussage vorlag, war diese Reduktion nicht möglich. Darum wurden die Aussagen zu jeder Kategorie quer über alle (auch nicht kritische) Situationen zusammengefasst und damit etwaige Lücken gefüllt.

Tabelle 6.14: U-Test für den Vergleich der Verteilung auf die verschiedenen Kategorien.

	Begriffe und Konzepte	Erkenntnis- gewinnung	Classroom Management	anderer didaktischer Aspekt	Sonstige
Mann-Whitney-U	50.0	53.0	57.0	44.5	49.5
Wilcoxon-W	128.0	108.0	112.0	99.5	127.5
Z	-.660	-.462	-.198	-1.025	-.696
Asymptotische Signifikanz (2-seitig)	.509	.644	.843	.305	.486

Für Situation 10 aus Video 4 ist es dennoch nicht gelungen, attraktive Distraktoren zu entwickeln. In der Hauptstudie werden die offenen Aussagen der Testpersonen entsprechend des Kodierschemas beurteilt und in die Analysen miteinbezogen. Die reduzierten Aussagen der Testpersonen zu jeder Situation wurden in ein Multiple-Choice-Design überführt. Da die Verbaldaten der Entwicklungsstudie II bereits zeigen, dass die Testpersonen trotz fokussierter Aufgabenstellung auch andere Aspekte im Unterricht wahrnehmen und äußern (wie zum Beispiel Aspekte des Classroom Managements), ist es denkbar über ein Ordered Multiple-Choice-Design (Briggs, Alonzo, Schwab & Wilson, 2006) weitere Informationen über die allgemeine professionelle Unterrichtswahrnehmung³⁹ zu gewinnen. Dieses Aufgabenformat unterscheidet sich vom klassischen Multiple-Choice-Format bezüglich der diagnostischen Aussagen, die man durch die Auswertung erhalten kann (ebd., S. 33). Indem die Distraktoren entlang einer Construct-Map (ebd., S. 37-39) entwickelt werden, kann aus dem Antwortverhalten der Testpersonen (bei Briggs et al. (2006) sind das Lernende) abgeleitet werden, auf welchem Verständnisniveau sich diese bezüglich des getesteten Konstrukts befinden (ebd., S. 34). Je nach Antwort der Testperson wird demnach nicht nur unterschieden, ob die richtige oder eine falsche Antwort gewählt wurde (dichotomes Verständnis von Fähigkeiten), sondern vielmehr welches Verständnis die jeweilige Testperson hat (ebd., S. 38). Die Construct-Map, die als Grundlage für die Ordered Multiple-Choice-Items des PUPiL-Instruments dient, orientiert sich an der Operationalisierung des Konstrukts der professionellen Unterrichtswahrnehmung, wie in Abschnitt 4.4 zur Operationalisierung beschrieben und dem induktiv entwickelten Kodiersystem. Da es in dieser Studie um eine physikspezifische professionelle Unterrichtswahrnehmung geht, wird das höchste Level der Construct-Map folgendermaßen beschrieben:

Level 4: Die Testpersonen erkennen Situationen, die aus einer physikdidaktischen Sicht für das fachliche Lernen der Lernenden relevant sind. Diese Situationen betreffen Schülervorstellungen und/oder Erkenntnisgewinnungsprozesse der Lernenden.

³⁹Wenn keine domänenspezifische Unterrichtswahrnehmung gemeint ist, wird die professionelle Unterrichtswahrnehmung in dieser Arbeit als „allgemein“ bezeichnet.

Geht man von einer fachspezifischen Fähigkeit aus, ist diese für jene Testpersonen weniger stark ausgeprägt, die einen eher generischen Fokus auf Lernen in Form von Aspekten des Classroom Managements legen. Selbst wenn die Antwort der Testperson einen anderen **didaktischen** Aspekt beträfe, ist diese Antwort mit weniger Punkten zu bewerten als die entsprechend Level 4, weil das Urteil aus der Inhaltvalidierung (vgl. Unterabschnitt 5.1.3 zur Validierung der Drehbücher) und dementsprechend die Aufgabenstellung an die Testpersonen explizit einen Fokus auf Schülervorstellungen und Erkenntnisgewinnung fordern. Das bedeutet für die mittleren Level der Construct-Map Folgendes:

Level 3: Die Testpersonen erkennen Situationen, die aus einer physikdidaktischen Sicht für das fachliche Lernen der Lernenden relevant sind. Diese Situationen betreffen kommunikative Kompetenzen, sprachliche Aspekte, Gender und weitere schülerzentrierte Aspekte des Lernens, aber nicht Schülervorstellungen und Erkenntnisgewinnung.

Level 2: Die Testpersonen erkennen Situationen, die auch aus allgemeindidaktischer Sicht für das fachliche Lernen der Lernenden relevant sind. Diese Situationen fallen unter Aspekte des Classroom Managements, die weniger fachliche Spezifität aufweisen und somit viel eher Lernen im Allgemeinen beeinflussen.

Wie in Unterabschnitt 6.1.3.1 bereits beschrieben, sollen die hier konstruierten Distraktoren auf zwei Ebenen falsch sein: Zum einen durch die unter Level 2 und 3 fallenden Aspekte, welche nicht aufgabengerecht sind und zum anderen durch das Wahrnehmen kontrafaktischer Aspekte, die aber aufgabengerecht sind. Deshalb ergibt sich für das letzte Level folgende Beschreibung:

Level 1: Die Testpersonen erkennen Situationen, die aus einer rein physikdidaktischen Sicht für das fachliche Lernen der Lernenden relevant sind, adressieren aber Geschehnisse, die nicht zu beobachten sind. Dieser Geschehnisse betreffen entweder Schülervorstellungen oder Erkenntnisgewinnungsprozesse der Lernenden.

Entlang dieser Level wurden für alle Situationen Distraktoren aufgelistet. Um eine mögliche Ratewahrscheinlichkeit für die zukünftigen Testpersonen möglichst gering zu halten und den Test dennoch überschaubar zu gestalten, wurden statt vier Antwortmöglichkeiten (entlang der Level) fünf Antwortmöglichkeiten entwickelt. Die Ratewahrscheinlichkeit liegt damit nur noch bei einem Fünftel. Da der Anteil an Beiträgen zu Aspekten des Classroom Managements am höchsten ist, wurde diese fünfte Antwortmöglichkeit zu der Kategorie Classroom Management konstruiert. Die Distraktoren und der Attraktor wurden überwiegend auf einem beschreibenden Level formuliert, weil ein Attraktor mit einer angemessenen Erläuterung zu attraktiv gegenüber den Distraktoren gewesen wäre. Bei einigen Distraktoren mussten dennoch erklärende Anteile integriert werden, weil sie ohne diese zu oberflächlich und damit nicht attraktiv wären.

Zu jeder Situation liegen demnach vier Distraktoren und ein Attraktor in der oben beschriebenen Form vor. Diese dienten als Grundlage für eine Diskussion mit vier Personen, die in der Physikdidaktik tätig sind, um zu einem Konsens bezüglich der sprachlichen Passung, Struktur und des Inhalts der Distraktoren zu kommen.

Tabelle 6.15: Antwortmöglichkeiten der Level 1 bis 4 für die Situation 2 aus Video 4.

Level	Antwortmöglichkeit
4	Der Schüler nennt ein Beispiel, in dem sich ein fachlich angemessenes Verständnis von Energieumwandlung anbahnt.
3	Der Schüler formuliert sprachlich einfache Sätze.
2	Ein Schüler ergänzt die Aussage eines anderen Schülers unaufgefordert.
2	Alle Schüler_innen folgen dem Unterrichtsgespräch.
1	Die Schüler_innen erklären physikalische Zusammenhänge, indem sie bereits gewonnene Erkenntnis damit in Verbindung bringen.

Ergebnis dieser Diskussion sind fünf Antwortmöglichkeiten zu 37 kritischen Situationen (siehe Anhang D). In Tabelle 6.15 ist exemplarisch dargestellt, wie die Antwortmöglichkeiten für die zweite Situation in Video 4 gestaltet sind. Es wurde anschließend anhand der Szenen in den Videos bestimmt, zu welchen Zeiten im Video die Frage (siehe Unterabschnitt 6.1.1) mit den jeweiligen geschlossenen Antworten erscheinen soll, wenn eine Testperson das Video anhält (siehe Ende dieses Kapitels). Einem projektexternen Bachelorstudenten im Informatikstudium wurde der Auftrag erteilt, ein webbasiertes Instrument aus den Videos, dem Fragebogen und weiteren oben beschriebenen Bestandteilen zu programmieren. Für die Dauer der Erhebung in der Hauptstudie war das PUPiL-Instrument passwortgeschützt online erreichbar.

6.2.3.2 Interpretation und Diskussion – Entwicklung von Distraktoren

In diesem Unterkapitel werden die Interpretationen der Ergebnisse präsentiert. Den einzelnen Interpretationen folgen Diskussionen, die aufzeigen, welche Befunde aus dieser Teilstudie die folgenden Schritte beeinflussten.

Neben der Entwicklung von Distraktoren für das PUPiL-Instrument in der Hauptstudie hat die Analyse der Antworten aller Testpersonen erneut etwas Aufschluss über die Fähigkeit, gestellten Unterricht physikspezifisch wahrzunehmen, gegeben.

Es lässt sich beobachten, dass insgesamt mehr Situationen wahrgenommen wurden, in denen Classroom Management lernrelevant werden könnte, als etwa Schülervorstellungen und Erkenntnisgewinnung. Das ist insofern plausibel, als dass Aspekte des Classroom Managements häufig in den gezeigten Situationen vorkommen, da Lernende miteinander und mit der Lehrkraft interagieren. Es war außerdem zu beobachten, dass die Verbaldaten mit den meisten Aussagen bezogen auf alle Aspekte von praktizierenden Lehrkräften kamen, welche wie oben bereits analysiert einen Fokus auf Classroom Management legen. In der fachdidaktischen Forschung und somit in der fachdidaktischen universitären Ausbildung wird darauf aber ein geringerer Fokus gelegt. Die Tatsache, dass häufig eine eher Classroom Management bezogene professionelle Unterrichtswahrnehmung gezeigt wird, hat Folgen für die Frage, ob sich die Antworten der Testpersonen auf das Konstrukt (hier die physikspezifische professionelle Unterrichtswahrnehmung) beziehen oder auch uner-

wartete Prozesse zur Lösung führen (vgl. Abschnitt 4.2 zur Validität). Entweder ist das PUPiL-Instrument nicht in der Lage, *nur* die physikspezifische professionelle Unterrichtswahrnehmung valide zu messen oder die professionelle Unterrichtswahrnehmung lässt sich nicht weiter in eine fachspezifische und eine generische Fähigkeit unterteilen. Ein weiterer Grund für die Befunde der Entwicklungsstudie II könnte sein, dass das Studium wenig Gelegenheiten bietet, eine fachspezifische professionelle Unterrichtswahrnehmung auszubilden und im Berufsleben die Wahrnehmung von Classroom Management Aspekten notwendiger ist als eine fachspezifische Unterrichtswahrnehmung, um Unterricht überhaupt zu ermöglichen. Ein Fokus auf eine fachspezifische professionelle Unterrichtswahrnehmung könnte möglicherweise auch erst nach einer generischen professionellen Unterrichtswahrnehmung gefördert werden. Die Befunde der Hauptstudie können helfen, diese Vermutungen zu stützen oder zu verwerfen.

Dass insgesamt mehr Aussagen zu Erkenntnisgewinnung als zu Schülervorstellungen fallen, ist verwunderlich. Außer in der Kombination von Video 1 und 3 sehen die Testpersonen mehr kritische Situationen, in denen Schülervorstellungen lernrelevant werden als Situationen bezüglich Erkenntnisgewinnung. Allerdings sind Erkenntnisgewinnungsprozesse häufig leicht zu beobachtende Handlungen, während Schülervorstellungen subjektiv konstruierte Konzepte sind. Diese sind schwerer auszumachen, weil die subjektive Konstruktion der Lernenden von Lehrkräften zunächst aus den Aussagen der Lernenden rekonstruiert werden muss. Man kann außerdem davon ausgehen, dass einige der in den Videos gezeigten Schülervorstellungen auch bei den Testpersonen noch aktiv (neben dem physikalisch angemessenen Konzept) bestehen.

Die Betrachtung der Antwortmuster der einzelnen Testpersonen zeigt, dass die Ergebnisse aus der Entwicklungsstudie II.1 reproduziert werden konnten. Das sieht man daran, dass vor allem die praktizierenden Lehrkräfte klare Fokusse (auf Erkenntnisgewinnung oder Classroom Management) legten, der Großteil der anderen Testpersonen ihre Aussagen aber relativ gleich verteilt zu den ersten drei Kategorien tätigten. Für die Hauptstudie kann erwartet werden, dass sich ebenfalls ähnliche Muster ergeben könnten, wobei keine praktizierenden Lehrkräfte in die Stichprobe der Hauptstudie mitaufgenommen werden. Wenn sich die Muster reproduzieren, kann die Frage gestellt werden, ob das PUPiL-Instrument die physikspezifische Unterrichtswahrnehmung misst oder ob es vielmehr misst, wie gut die Testpersonen die Aufgabenstellung verstanden haben und behalten können.

Bei vier Testpersonen ist auffällig, dass mehr als fünf Aussagen der Kategorie „Sonstige“ zugeordnet wurden. Anders als im ersten Durchgang dieser Entwicklungsstudie II betreffen die Äußerungen nicht Unsicherheit zum Umgang mit dem PUPiL-Instrument, sondern vielmehr Ereignisse im Unterricht, die nicht in die vier inhaltlichen Kategorien fallen und weniger relevant für die Bearbeitung der Aufgabe sind. Das können beispielsweise Aussagen über Intentionen der Lehrkraft sein. Es kann daher geschlussfolgert werden, dass die Testpersonen sicherer im Umgang mit den Videos waren. Dieser Effekt kann aber nicht eindeutig der Übung anhand eines ähnlichen Videos zugesprochen werden. Es könnte ebenso gut an den einzelnen Testpersonen der Stichprobe in der Entwicklungsstudie II.2 liegen.

Der Einsatz von Ordered Multiple-Choice-Items geschieht in dieser Studie weniger geprüft, als es derzeit in der bildungswissenschaftlichen Forschung üblich ist. Es gibt einige Studien, die sich mit der Frage beschäftigen, ob dieses Frageformat valider ist als ein dichotomes und ob es offene Frageformate ersetzen kann (vgl. zum Beispiel Briggs et al., 2006; Hadenfeldt, Repenning & Neumann, 2014; Lin, Chu & Meng, 2010). In dieser Studie wird Ordered Multiple-Choice aber nicht angewendet, weil eine höhere Validität erhofft wird, sondern weil der Anspruch besteht, aus den gesammelten Daten eine möglichst gute Aufklärung über die professionelle Unterrichtswahrnehmung der Testpersonen durch den Test zu erhalten. Statt nur die Information zu generieren, ob eine Testperson das fachdidaktisch, normativ Angemessene wahrgenommen hat oder nicht, wird außerdem in Erfahrung gebracht, auf welchen Unterrichtsaspekten ansonsten der Fokus der professionellen Unterrichtswahrnehmung liegt.

In der Umsetzung der entwickelten Distraktoren in ein webbasiertes Instrument ergab sich die Schwierigkeit, dass festgelegt werden musste, in welchen Intervallen eine kritische Situation erkannt und das Video von einer Testperson angehalten werden muss, damit die Situation durch die Testperson als wahrgenommen gilt. In dem PUPiL-Instrument hat diese Entscheidung zur Folge, ob die Testperson geschlossene Antworten vorgelegt bekommt oder in einem Textfeld ihre Wahrnehmung selber frei äußern kann. Halten sie das Video während einer kritischen Situation an, wird ihnen eine geschlossene Frage vorgelegt, in allen anderen Fällen werden sie mit einem Satz um die Mitteilung ihrer Wahrnehmung in einem leeren Textfeld gebeten. Ein Intervall für die kritische Situation wurden zunächst so gesetzt, dass es beginnt, wenn der Kern der Situation bereits gesehen werden konnte und endet, wenn alle notwendigen Ereignisse der Situation gesehen werden konnten, spätestens aber bevor eine neue kritische Situation beginnt oder es einen Szenenwechsel gibt. Den Testpersonen soll nicht die Möglichkeit gegeben werden, über die Situation reflektierend nachzudenken oder sie im Kontext tatsächlich folgender Ereignisse zu interpretieren, weil das das Messkonstrukt verändern würde (vgl. Abschnitt 2.4 zur Abgrenzung der professionellen Unterrichtswahrnehmung von Reflexion.). Daher umfassen die Intervalle mindestens zwei, maximal 23 und im Schnitt 6,6 Sekunden.

Die Art, wie die Distraktoren entwickelt wurden, trägt zur Inhaltsvalidität des PUPiL-Instruments und seiner Interpretationen bei. Die Inhalte und Struktur der Antwortmöglichkeiten wurden aus Aussagen von Testpersonen abgeleitet, die in einer performanznahen Erhebung ihre professionelle Unterrichtswahrnehmung zeigen konnten. Damit repräsentieren die Items nur Inhalte, die sich auf für die (zukünftigen) Testpersonen Relevantes beziehen. Es wäre allerdings möglich, dass sich bei der Analyse von Aussagen einer anderen Stichprobe als der in Entwicklungsstudie II auch andere Inhalte und Strukturen für die Distraktoren ergeben hätten.

Das fertig programmierte PUPiL-Instrument samt Videos und entwickeltem Fragebogen wird um einen Fragebogen zu demographischen Daten, zur Selbsteinschätzung zu den in den Videos vorkommenden Themen sowie Skalen aus dem fachdidaktischen Wissenstest (Gramzow, 2015) ergänzt. Die Validierung des Tests wird in der Hauptstudie abgeschlossen.

7 Hauptstudie

In diesem Kapitel werden Methode, Ergebnisse und Interpretation der Hauptstudie präsentiert, in der das PUPiL-Instrument zur abschließenden Validierung eingesetzt wurde. Ziel der gesamten Studie ist ein validiertes Messinstrument zur Erfassung einer physikspezifischen professionellen Unterrichtswahrnehmung von angehenden Physiklehrkräften, welches aus Videos und einem teils geschlossenen Fragebogen besteht. Dieses soll durch Außenkriterien wie das fachdidaktische Wissen und durch die Prüfung konstruktimmanner Annahmen validiert werden (vergleiche H1 und H2).

7.1 Methode

Im Folgenden werden das PUPiL-Instrument, die Stichprobe, das Vorgehen während der Erhebung und das Vorgehen der Datenanalyse beschrieben.

7.1.1 Das PUPiL-Instrument

Das PUPiL-Instrument ist auf eine Testzeit von 90 Minuten ausgelegt. Der Kern des PUPiL-Instruments sind je Testperson zwei der drei Videos aus der Entwicklungsstudie I mit dem in der Entwicklungsstudie II entwickelten geschlossenen Fragebogen. Jedes Video wird durch ein permutierendes Test-Design von circa zwei Dritteln der Testpersonen gesehen.

In der Entwicklungsstudie II hat sich gezeigt, dass der Einsatz eines Übungsvideos (zur Verkehrssicherheit) die Testpersonen unterstützt, indem es einen Trainingseffekt erzeugt. Die Testpersonen üben den Umgang mit dem PUPiL-Test an einer analogen Situation. Die Distraktoren des Fragebogens (Ordered Multiple-Choice mit Single Select) im PUPiL-Test sind dabei entweder falsch in Bezug auf die Aufgabenstellung, behandeln statt Schülervorstellungen und Erkenntnisgewinnung Aspekte wie Classroom Management oder andere didaktische Aspekte wie Sprache im Physikunterricht. Oder aber die Distraktoren sind auf inhaltlicher Ebene falsch, betreffen also Schülervorstellungen oder Erkenntnisgewinnungsprozesse, sind aber kontrafaktisch.

Vor diesem Teil des Tests werden zusätzlich Fragen zu demographischen Daten erhoben. Es werden unter anderem Informationen zu Lehrerfahrungen gemessen an Erfahrungen mit Nachhilfeunterricht, Schulpraktika und Lehraufträgen sowie Studium erfragt.

Der Test schließt mit einer Selbstauskunft zu den in den Videos behandelten Themen und

Skalen zu *fachdidaktischen Konzepten* und *Experimentieren* aus dem fachdidaktischen Wissenstest des Projektes Profile-P (Gramzow, 2015) ab. Bezüglich der Selbstauskunft über die Themen werden die Testpersonen gefragt, mit welchen der folgenden Themen sie sich in ihrer Schulzeit, beziehungsweise universitären Ausbildung befasst haben. Zu den verschiedenen Themen⁴⁰ mussten die Testpersonen sich auf einer vierstufigen Likertskala selbst einschätzen.

7.1.2 Die Stichprobe

Für die Beantwortung der Forschungsfrage und die Testung der aufgestellten Hypothesen ist eine pseudo-längsschnittliche Stichprobe mit Studierenden verschiedener (naturwissenschaftsverbundener) Fächer angezeigt (vergleiche Tabelle 7.1), weil somit die Fragen beantwortet können, ob sich die professionelle Unterrichtswahrnehmung im Laufe des Studiums (H2) und in Abhängigkeit vom Fach (H1) entwickelt. Das bedeutet, dass Studierende aller Semester zum gleichen Zeitpunkt den Test bearbeiten. So kann zwar über einen tatsächlichen Verlauf einer Entwicklung der physikspezifischen professionellen Unterrichtswahrnehmung keine direkte Aussage getätigt werden, dennoch kann aus der Analyse von Medianunterschieden eine mögliche Entwicklung geschlussfolgert werden.

Für die Testung der Hypothese H2⁴¹ wurden sowohl Geographielehramtsstudierende als auch Sachunterrichtsstudierende in die Stichprobe integriert. Es wären auch andere Fächer möglich gewesen, aber vor allem bei naturwissenschaftsfernen Fächern liegt die Vermutung nahe, dass diese aufgrund fehlenden physikspezifischen Professionswissens zu Bodeneffekten führen würden. Es wurden außerdem Vollfachphysikstudierenden in die Erhebung miteinbezogen, um zu verstehen, welchen Anteil Fachwissen und fachdidaktisches Wissen an der Fähigkeit Physikunterricht professionell wahrzunehmen haben. Allerdings konnten nur fünf Studierende akquiriert werden, welche nicht repräsentativ für alle Vollfachphysikstudierenden sein dürften. Insbesondere weil sie sich freiwillig zur Erhebung gemeldet haben, ist mit einer sogenannten Positivauswahl zu rechnen. Ein besonders gutes Abschneiden dieser Teilstichprobe würde demnach nicht verwundern. Die Physiklehramtsstudierenden im dritten Semester haben bezüglich des physikalischen Fachwissens und physikdidaktischen Wissens die gleiche Ausbildung wie die Vollfachphysikstudierenden dieser Studie, weshalb für die Analysen der Daten die Gruppe der Vollfachphysikstudierenden den Physiklehramtsstudierenden zugeordnet wurde.

Um für die anvisierte IRT-Skalierung einen Stichprobenumfang zu erhalten, der in der Literatur (Hartig & Frey, 2013) als ausreichend gehandelt wird, wurde die Erhebung an

⁴⁰Diese Themen waren bezüglich des Physikwissens Energie, Elektrostatik, Teilchenmodell, Teilchenbewegung, Planung von Experimenten, Hypothesenbildung und Auswertung von Daten. Zum didaktischen Wissen wurde die Selbsteinschätzung bezüglich Bildungsstandards für den Physikunterricht, Schülervorstellungen im Physikunterricht, Konzeptwechselstrategien, Experimentieren in Physikunterricht, Classroom Management, Zusammenhang von Sprache und dem Lernen von Physik sowie Planung von Physikunterricht eingeholt.

⁴¹Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften verändert sich im Studienverlauf signifikant.

drei Hochschulen in Deutschland durchgeführt. Auf diese Weise konnte eine Stichprobe mit $N = 160$ Testpersonen zusammengestellt werden. Das bedeutet, dass jedes Video im Schnitt von 100 Testpersonen angesehen wird. Dabei sind zwei Drittel der Stichprobe an der Hochschule 1, 20 Studierende an der Hochschule 2 und 35 Studierende an der Hochschule 3 getestet worden.

Circa 77,5% der Studierenden in der Stichprobe befinden sich im Bachelor und 22,5% der Studierenden im Master ihres Studiums.

Für das Physiklehramt untypisch, gibt es eine Mehrheit von weiblichen Testpersonen, sodass es zu einer Verteilung von 63,7% (weiblich) zu 36,3% (männlich) kommt. Dieser hohe Anteil an Studentinnen ist nur für den Studiengang Sekundarstufe 1 typisch (vgl. Düchs & Ingold, 2017, S. 29). Dass die Mehrheit der Testpersonen weiblich ist, kann auch mit den Fächern zusammenhängen, welche neben dem Physiklehramt erhoben wurden: Circa ein Fünftel der Testpersonen studierte Geographielehramt und mehr als ein Drittel befand sich im Sachunterrichtslehramt. Der Anteil an Studentinnen ist im Lehramt Physik der Sekundarstufe I höher als der Anteil von Studenten (ebd.).

Tabelle 7.1: Beschreibung der Stichprobe der Hauptstudie. *LAPhy* bedeutet Lehramt Physik, *LAGeo* bedeutet Lehramt Geographie, *LASU* bedeutet Lehramt Sachunterricht.

Geschlecht	w = 63,7%		m = 36,3%
Bachelor/Master	Bachelor = 77,5%		Master = 22,5%
Hochschule	Hochschule 1 = 65,6%	Hochschule 2 = 12,5%	Hochschule 3 = 21,9%
Fächer	LAPhy = 45,3%	LAGeo = 20,5%	LASU = 34,2%

7.1.3 Die Erhebung zur Validierung

Die Erhebung zur Validierung des Instruments dauerte 90 Minuten und unterschied sich somit von zukünftigen Erhebungen mit dem PUPIL-Instrument. Für die Validierung wurden beispielsweise zusätzlich demographische Daten und weitere Tests zum professionellen Wissen der Testpersonen erhoben.

Da der Test online basiert aufgebaut ist, erhielt jede Testperson unabhängig von der Testleitung dieselben Aufforderungen. Bei Abweichungen im Ablauf, weil beispielsweise das Programm fehlerhaft operierte, notierte die Testleitung ausführliche Beschreibungen der Abweichung. Anschließend wurde durch die Testentwicklerin entschieden, inwiefern die betroffenen Testpersonen den Test fortführen oder neu beginnen sollten – sofern sie dazu bereit waren.

Zu Beginn des Tests wurden die Testpersonen gebeten, Angaben zu demographischen Daten zu machen (siehe für den Fragebogen Anhang E.1 und E.2). Die Items fragten unter anderem nach der Anzahl absolvierter Hochschulseмester, dem angestrebten Abschluss (beispielsweise Lehramt an Gymnasien oder Berufsschulen), den studierten Fächern und

bisherigen Abschlüssen, um einen möglichen Zusammenhang von Fortschritt und Art des Studiums mit der professionellen Unterrichtswahrnehmung überprüfen zu können. Des Weiteren sollten die Testpersonen angeben, wie viel Lehrerfahrung sie durch (a) Nachhilfe, (b) Unterrichtspraktika mit Schul- oder Fachbezug und (c) Lehraufträgen an Schulen gesammelt hatten. Diese Angabe hat geholfen, Zusammenhänge von (nicht angeleiteter) Lehrerfahrung und physikspezifischer professioneller Unterrichtswahrnehmung zu prüfen. Nachdem die Testpersonen Fragen zu demographischen Daten beantwortet haben, begannen sie den Test mit der Übung zum Umgang mit den Videos. Dort sahen die Testpersonen zunächst ein Video, wie eine Person mit dem Übungsvideo interagiert, damit sie erfuhren, was von Ihnen erwartet wurde. Anschließend erhielten die Testpersonen folgende Aufgabe, anhand derer sie das entsprechende Verhalten der Person im Übungsvideo imitieren und somit üben konnten:

„Jetzt sind Sie an der Reihe! Sie sehen sich nun das Verkehrsvideo an.
Halten Sie das Video immer dann an, wenn Sie etwas in Bezug auf Verkehrssicherheit von
Fahrradfahrern und/oder Fußgängern wahrnehmen.
Wenn Sie das Video angehalten haben, geben Sie bitte wahrheitsgemäß das von Ihnen wahrgenommene Ereignis an.
Halten Sie das Video besser häufiger an, als zu selten.“

Mit „Start“ können Sie jetzt beginnen.“



Beobachtung

Was ist Ihnen aufgefallen?

Tragen Sie Ihre Beobachtungen in das Feld ein und klicken Sie dann erneut auf Start, um das Video fortzusetzen.



PROFALE
PROFESSIONELLES LEHRERHANDELN ZUR
FÖRDERUNG FACHLICHEN LERNENS

Übung zum Pausieren



Start

Stopp

Abbildung 7.1: Interface bei der Übung zum Umgang mit den Videos. Die Testpersonen sahen ein Video mit Straßenverkehr in Hamburg und hatten nach Anhalten des Videos die Möglichkeit, einen Freitext über die wahrgenommene Situation zu verfassen.

Bei der Übung hatten die Testpersonen analog zu den Unterrichtsvideos die Möglichkeit, das Video anzuhalten, wenn sie etwas in Bezug auf Verkehrssicherheit der Radfahrenden

oder von Personen, die zu Fuß unterwegs sind, wahrgenommen haben. Anschließend sollten sie entweder in einem freien Textfeld (siehe Abbildung 7.1) oder per Ankreuzen eines Multiple-Choice-Items (siehe Abbildung 7.2) ihre Wahrnehmung mitteilen. Nach Ende des Videos wurde den Testpersonen per Text auf dem Bildschirm mitgeteilt, dass die Übung vorbei war. Diese Übung diente – wie in der Entwicklungsstudie II dargestellt – einem geübteren Umgang mit den Videos.

Wenn die Testpersonen auf „Weiter“ klickten, gelangten sie zu der Aufgabenstellung, mit welcher sie die Videos des Tests bearbeiten sollten. Diese Aufgabenstellung lautete – ebenfalls wie bereits in der Entwicklungsstudie II – folgendermaßen:

„Sie sehen nun zwei Videos, in denen Klassensituationen zu sehen sind. Stellen Sie sich vor, dass Sie die verantwortliche Lehrkraft im Klassenraum sind.

1. Drücken Sie immer dann auf STOPP, wenn Sie etwas Lernrelevantes wahrnehmen in Bezug auf:
 - Vorstellungen der Schüler_innen über physikalische Begriffe und Konzepte (Beispiel: Schülervorstellungen zur Lichtbrechung)
 - Physikalische Erkenntnisgewinnungsprozesse der Schüler_innen
2. Wenn Sie STOPP drücken, können Sie sich damit auf mehr oder auch weniger gelingende Lernsituationen beziehen. Zögern Sie nicht die Videos zu stoppen, immer wenn Sie etwas Entsprechendes wahrnehmen. Wenn Sie gestoppt haben, werden Sie aufgefordert Ihre Wahrnehmung entweder knapp zu begründen oder Ihnen wird eine Frage gestellt.

Für den PUPiL-Test hatten die Testpersonen die Aufgabe, das Video an wahrgenommenen Stellen unmittelbar anzuhalten und anschließend ihre Wahrnehmung anhand der Distraktoren und dem Attraktor, die in der Entwicklungsstudie II entwickelt wurden, als Multiple-Choice Single Select zu benennen oder ihre Beobachtung als offene Antwort zu geben, wenn keine der Antwortoptionen passend erschien (siehe Abbildung 7.3).

Hielten die Testpersonen das Video an Stellen an, in denen sich keine validierte kritische Situation befand, wurde ein freies Textfeld geöffnet, in dem sie ihre Wahrnehmung anhand der oben genannten Aufgabe 2 beschreiben konnten. Mit diesem Vorgehen wurde eine Reaktion der Testpersonen anvisiert, von welcher besser auf die eigentlichen Wahrnehmungsprozesse geschlossen werden kann als mit klassischen Paper-Pencil-Tests.

Abschließend bearbeiteten die Testpersonen zwei Skalen des fachdidaktischen Wissenstests aus dem Projekt Profile-P (Gramzow, 2015). Diese beiden Skalen bestanden aus 22 Items bezüglich Wissen über Konzepte und Konzeptbildung sowie Wissen über Experimentieren im Physikunterricht (vgl. entsprechende Skalen bei Gramzow, 2015).

Diese Skalen wurden ausgewählt, weil die Items Wissen zu Schülervorstellungen und Erkenntnisgewinnung der Testpersonen abdecken. Dieses Wissen wird theoriegeleitet (vgl. 4.1) als mögliche Voraussetzung für die physikspezifische professionelle Unterrichtswahrnehmung der verwendeten Videos gehandelt, weil das Wissen eben die Aspekte der Schülervorstellungen und Erkenntnisgewinnung betrifft. Das Ergebnis des fachdidaktischen Wissenstests wurde zur Überprüfung der Validität des PUPiL-Tests erhoben.

Übung zum Pausieren

Was haben Sie ahrgenommen? Bitte kreuzen Sie **EINE** Antwort an.

- ☐ Der Radweg ist uneben.
- ☐ Die Autos fahren zu schnell an der Kreuzung vorbei.
- ☐ Die Fahrradfahrerin fährt auf dem Fußgängerüberweg.
- ☐ Fußgänger gehen über rot.
- ☐ Die Fahrradfahrerin mit dem beigen Mantel ist zu langsam.
- ☐ Keine der Optionen, sondern ...



Start

Stopp

Weiter

Abbildung 7.2: Interface bei der Übung zum Umgang mit den Videos. Die Testpersonen sehen ein Video mit Straßenverkehr in Hamburg und haben nach Anhalten des Videos die Möglichkeit, ihre Wahrnehmung über eine geschlossene Frage mitzuteilen.

Was haben Sie wahrgenommen? Bitte kreuzen Sie **EINE** Antwort an.

- ☐ Die Schülerin äußert Beobachtungen zu ihrem Experiment und wertet diese aus.
- ☐ Der Schüler im Hintergrund hat seinen Kopf auf dem Tisch liegen und arbeitet nicht.
- ☐ Die Schülerin stellt sich vor, dass sich im warmen Wasser mehr Farbstoffe gelöst haben.
- ☐ Die Schülerin und der Schüler sind am Ende der Arbeitsphase immer noch diszipliniert.
- ☐ Der Schüler nutzt den Vergleich mit Schokolade, um sich sprachlich zu entlasten.
- ☐ Keine der Optionen, sondern ...

Test zu Wahrnehmung - Unterrichtsvideo 1



Start

Stopp

Weiter

Fortschritt: 49%

Abbildung 7.3: Interface zu einem angehaltenen Video 1 während einer kritischen Situation. Die Wahrnehmung der Testperson sollte über das Ankreuzen der gegebenen Distraktoren mitgeteilt werden.

7.1.4 Die Kodierung für die Analyse der Daten

Die Daten aus der Erhebung wurden durch das Online-Programm teils numerisch, teils nominal kodiert. Für die Analyse der Daten wurden folgende Kodierungen vorgenommen:

- **Video:** Für jede Testperson wurde angegeben, welches Video diese nicht angesehen hat. Die Variable wurde mit Video_x bezeichnet und konnte die Werte 1, 3 und 4 annehmen.
- **Demographische Daten:**
 - *Fach:* Da Studierende verschiedener Fächer an der Erhebung teilgenommen haben, wurde die Variable „Fach“ aufgenommen. Sie kann die Werte 1 – Physik, 2 – Lehramt Geographie, 3 – Lehramt Sachunterricht, 4 – Lehramt Physik annehmen. Auch wenn hier eine Hierarchie impliziert scheint, wird den Fächern eine Gleichwertigkeit zugewiesen, indem Fach als nominale Variable behandelt wurde.
 - *Studienfortschritt:* Bezüglich der Variable Studienfortschritt, hat sich gezeigt, dass eine Unterscheidung in Bachelor und Master in Hinblick auf Entwicklungsverläufe sinnvoller ist als in einzelnen Semester. Die Studierenden waren sich häufig nicht sicher, in welchem Fachsemester oder Hochschulsesemester sie waren. Daher bezeichnet die ordinale Variable die verschiedenen Ausbildungsstufen mit 1 – im Bachelorstudium und 2 – im Masterstudium.
 - *Studiengang:* In den verschiedenen Studiengängen gibt es zu verschiedenen Zeiten und in verschiedener Häufigkeit fachdidaktische Förderung durch Seminare, Vorlesungen und Praktika. Der angestrebte Studienabschluss wurde daher als nominale Variable „Studiengang“ mit 0 – Bachelor Physik, 1 – Lehramt für Sonderpädagogik, 2 – Lehramt an Berufsschulen, 3 – Lehramt für Sekundarstufe I und 4 – Lehramt an Gymnasien kodiert. Auch hier wurden die verschiedenen Ausprägungen der Kodierung gleichwertig in einer nominalen Variable behandelt.
 - *Hochschule:* Um mögliche Zusammenhänge mit der Ausbildungsstätte nachweisen zu können, wurde der Standort der Erhebung mit „Uni“ bezeichnet. Diese nominale Variable erhält die Ausprägungen 1 – Hochschule 1, 2 – Hochschule 2, 3 – Hochschule 2.
- **physikdidaktisches Wissen:** Die Daten zum physikdidaktischen Wissen wurden entlang des Kodierschemas des Projektes Profile-P kodiert (siehe Gramzow, 2015, S. 244-274). Die Items wurden jeweils mit „fd“ und der Itemnummer im Originaltest benannt (zum Beispiel: fd14).
- **professionelle Unterrichtswahrnehmung:** Die kritischen Situationen im Test wurden jeweils mit „vi_x.y“ benannt, wobei x das Video bezeichnet (also 1, 3 oder 4) und y die Situation (von 1 bis maximal 15). Die Wahl der Antwortmöglichkeit bei

geschlossenen Fragen wurde entlang der Level der Ordered Multiple-Choice-Items ordinal kodiert, wie in Unterunterabschnitt 6.2.3.1 zur Auswertung der Entwicklungsstudie II.2 bereits begründet. Wenn ein Video bei einer kritischen Situation nicht angehalten wurde, ist dies mit 0 – nicht wahrgenommen kodiert worden. Die Wahl der Distraktoren wurde mit 1 – kontrafaktisch, 2 – Classroom Management, 3 – anderer didaktischer Aspekt und 4 – Attraktor kodiert.

Die Wahl eines Distraktors oder Attraktors kennzeichnet, dass eine Testperson die kritische Situation erkannt hat (demnach wurde Perception des PID-Modells erreicht) und diese interpretiert, indem sie angibt, welcher Aspekt als lernrelevant wahrgenommen wurde. Da sich bereits in der Entwicklungsstudie II gezeigt hat, dass die Testpersonen selten einen Handlungsplan nennen, Perception und Interpretation aber nicht getrennt gemessen werden konnte, wurde das Fassen eines Handlungsplans empirisch von den Teilprozessen Perception und Interpretation getrennt. Dies widerspricht zwar der theoretischen Fassung des Konzepts der professionellen Unterrichtswahrnehmung, erwies sich aber im Lichte der Datenlage als sinnvoll (vgl. Kapitel 6 zur Entwicklungsstudie II.2). Das bedeutet, dass für die Teilprozesse Perception und Interpretation zusammen ein Code vergeben wurde und für Decision Making ein weiterer (siehe nächster Stichpunkt).

Wie in Abschnitt 6.2 zur Entwicklungsstudie II.2 erklärt, mussten Zeitintervalle gesetzt werden, in denen den Testpersonen bei Anhalten des Videos eine zu der kritischen Situation passende geschlossene Frage vorgelegt wird. Um möglichst viele Daten für eine Analyse der physikspezifischen professionellen Unterrichtswahrnehmung zu erhalten, wurden aber auch alle Antworten in Form von Freitexten, die drei Sekunden vor oder nach diesem Intervall gegeben wurden, entlang des obigen Kodierschemas eingestuft. Wenn die Antworten keine unmittelbar angrenzende kritische Situation betrafen, wurden diese ignoriert.

Die qualitativen Daten, die kritische Situationen betrafen, wurden von drei Kodierenden eingestuft. Dafür wurden erst zehn Prozent des Datenmaterials – bestehend aus den Freitexten der Testpersonen – kodiert. Es ergaben sich für die verschiedenen Kategorien folgende Übereinstimmungskoeffizienten (vgl. Krippendorff, 2011, SV – Schülervorstellungen, EG – Erkenntnisgewinnung, CM – Classroom Management, adA – anderer didaktischer Aspekt)⁴²: (a) $\alpha_{SV} = 0.774$, (b) $\alpha_{EG} = 0.793$, (c) $\alpha_{CM} = 0.735$, (d) $\alpha_{adA} = 0.475$. In einer anschließenden Diskussion wurden Differenzen besprochen und gegebenenfalls Änderungen an der Kodierung vorgenommen. Außerdem wurden Konkretisierungsvorschläge für das Kodierschema erarbeitet und angewendet. Nach der Diskussion ergaben sich folgende Übereinstimmungen für das selbe Datenmaterial: (a) $\alpha_{SV} = 0.914$, (b) $\alpha_{EG} = 0.797$, (c) $\alpha_{CM} = 0.840$ und (d) $\alpha_{adA} = 0.727$. Um die Daten möglichst objektiv bewerten zu können, wurde anschließend das ganze Datenmaterial bezüglich der Freitexte zu den kritischen Situationen der Testpersonen mit dem überarbeiteten Kodierschema durch die drei Kodierenden eingestuft ((a) $\alpha_{SV} = 0.713$, (b) $\alpha_{EG} = 0.727$, (c) $\alpha_{CM} = 0.832$, (d)

⁴²Die Berechnungen wurden durch eine Internetseite durchgeführt, die nach Skalenniveau und Anzahl der Kodierenden differenziert (<http://dfreelon.org/utlis/recalfront/>)

$\alpha_{adA} = 0.645$). Auch nach dieser Kodierung hat eine Diskussion zu einer Verbesserung der InterCoderübereinstimmung geführt: (a) $\alpha_{SV} = 0.949$, (b) $\alpha_{EG} = 0.969$, (c) $\alpha_{CM} = 0.977$, (d) $\alpha_{adA} = 0.919$.

Wie ebenfalls in der Entwicklungsstudie II beschrieben, gab es zu einer Situation (Situation 10 in Video 4) nicht ausreichend Datenmaterial, sodass keine geschlossene Frage entwickelt werden konnte. Alle schriftlichen Kommentare der Testpersonen zu dieser Situation wurden durch obiges Kodierschema kodiert und auf diese Weise mit in die Analyse der Daten integriert. Dieses Vorgehen sichert, dass alle in der Entwicklungsstudie I validierten kritischen Situationen das Potential behalten, in das PUPiL-Instrument einzufließen. Sollte Situation 10 aus Video 4 durch die Analysen als aufklärender Bestandteil des PUPiL-Instruments bestätigt werden, könnten aus den Antworten der Testpersonen in der Hauptstudie Distraktoren für Folgestudien entwickelt werden.

- **Fassen eines Handlungsplans:** Wie in Abschnitt 6.2 zur Entwicklungsstudie II.2 erläutert, wird das Fassen eines Handlungsplans nicht als Teil der professionellen Unterrichtswahrnehmung kodiert. Stattdessen werden die Handlungspläne in Form von Aussagen, welche die Testpersonen in einem freien Textfeld getätigt haben, entlang eines Kodierschemas eingestuft (vergleiche Tabelle 7.2 und 7.3). Hierbei wird unterschieden, ob ein Handlungsplan gefasst wurde, ob dieser inhaltlichen Bezug zu der erkannten und interpretierten Situation hat und ob Begründungen für diesen angeführt werden. Das höchstmögliche Level bezüglich Fassen eines Handlungsplans ist erreicht, wenn ein begründeter, zur interpretierten Situation passender Handlungsplan genannt und von Alternativen abgegrenzt wird. Der Übereinstimmungskoeffizient κ_n nach Kodierung von zehn Prozent des Datenmaterials durch zwei unabhängig Kodierende betrug $\kappa_n = 0.773$ und nach Diskussion der Kodierung $\kappa_n = 0.993$. Das Kodierschema wurde der Diskussion folgend angepasst.
- **Selbsteinschätzungen:** Die Selbsteinschätzungen der Testpersonen zu Lehrerfahrung wurden gemäß Länge und Häufigkeit der jeweiligen Erfahrungen mit Punkten bewertet. Je länger und häufiger die Tätigkeiten ausgeführt wurden, desto mehr Punkte erhielten die Testpersonen.

Das subjektiv eingeschätzte Wissen zu den in den Videos behandelten Themen wurde entlang der Likertskalen (von einem bis vier Punkten) bewertet. Anschließend wurde zu jedem Oberthema (Schulphysik, Universitätsphysik, Didaktik) jeweils eine Summe über die Punkte der Testpersonen berechnet.

Tabelle 7.2: Kodierschema zum Fassen eines Handlungsplans in der Hauptstudie – Teil 1.

Titel	Code-Definition	Ankerbeispiel
0 – kein Handlungsplan	Es wurde nichts geschrieben. ODER Es wurde eine nicht verständliche Zeichenabfolge geschrieben. ODER Es wurde die wahrgenommene Situation paraphrasiert.	„Die Schülerin erzählt von ihren eigenen Erfahrungen und wagt sich daran als erste eine Vermutung aufzustellen wie sich der Tee im Wasser verhält.“
1 – „falschen“ Handlungsplan nennen	Es wurde ein Handlungsplan oder mehrere Handlungspläne genannt aber OHNE Bezug zur wahrgenommenen Situation.	„Ich würde auf die Schülerin hinten rechts einwirken, welche den Anschein machen, nicht voranzukommen.“ (In der kritischen Situation ging es aber um ein nicht angemessenes Verständnis von Hypothesen einer anderen Schülerin.)
2 – Handlungsplan nennen	Es wurde ein Handlungsplan oder mehrere Handlungspläne genannt OHNE Begründung aber MIT Bezug zur wahrgenommenen Situation.	„Der These der Schülerin, dass der Tee sich auch in Eiswürfeln langfristig lösen würde, kann im folgenden Unterricht nachgegangen werden beziehungsweise kann diese berücksichtigt werden. Beim Endergebnis der Unterrichtsstunde kann auf diese These noch einmal Bezug genommen werden und die These eventuell überarbeitet werden.“

Tabelle 7.3: Kodierschema zum Fassen eines Handlungsplans in der Hauptstudie – Teil 2.

Titel	Code-Definition	Ankerbeispiel
3 – Handlungsplan begründen	Es wurde ein Handlungsplan oder mehrere Handlungspläne genannt MIT mindestens einer Begründung, die zu mindestens einer der genannten Handlungspläne passt.	„Ich würde mich hier möglichst wenig einmischen, da die zwei Kinder konzentriert und sachlich arbeiten. Sie hangeln sich korrekt am Handlungsfaden entlang.“
4 – Abgrenzen zu Alternativen	Es wurde ein Handlungsplan oder mehrere Handlungspläne genannt MIT passender Begründung und Abgrenzung zu Alternativen.	„Die Lehrkraft sollte sich Zeit nehmen verschiedene Beispiele anzuhören, denn die Klasse ist motiviert und möchte etwas beitragen. Wenn nur wenige und vielleicht immer die gleichen dran genommen werden, sinkt die Motivation der anderen SuS.“

Nach der Kodierung der Daten wurden diese quantitativ ausgewertet. Zunächst wurde eine IRT-Skalierung (gemäß der Item-Response-Theory) der Daten mittels Partial-Credit-Modell (vgl. zu IRT-Skalierung Strobl, 2010) zur physikspezifischen professionellen Unterrichtswahrnehmung vorgenommen. Diese diente der Auswahl konstruktkonformer Items anhand Modellpassung und Itemschwierigkeit sowie der Einschätzung der Personenfähigkeit, um aus diesen Items ein valides Instrument zusammenstellen zu können. Um diese Auswahl besser nachvollziehen zu können, folgt nun ein Exkurs zum Partial-Credit-Modell der IRT-Skalierung.

Exkurs IRT-Skalierung Mit der sogenannten Item-Response-Theory (IRT) werden Modelle aufgestellt, die zwischen latenten Personenfähigkeiten und beobachtbaren Daten einen logistischen Zusammenhang voraussetzen (vgl. Irtel, 1996, S. 46). Außerdem bauen IRT-Modelle auf indirekten Messungen durch Items in einem Test auf (ebd.). Als eine der größten Stärken zählt, dass durch die IRT-Skalierung – anders als in der klassischen Testtheorie – die Lösungswahrscheinlichkeiten auf Grundlage von Personenfähigkeiten θ_i einer Person i und Itemschwierigkeit δ_j eines Items j geschätzt werden (ebd.). Es wird demnach berücksichtigt, dass Testpersonen mit einer höheren Fähigkeit schwere Aufgaben mit einer höheren Wahrscheinlichkeit lösen als Testpersonen mit einer niedrigeren Fähigkeit.

Neben diesen Vorteilen werden an die zu modellierenden Daten Bedingungen gestellt. Daten können beispielsweise nur in ein IRT-Modell überführt werden, wenn die einzelnen Items voneinander linear stochastisch unabhängig sind (vgl. Strobl, 2010, S. 16-20). Die Lösungswahrscheinlichkeit für ein Item darf demnach nur von der Personenfähigkeit der Testperson abhängen und nicht vom Lösen vorheriger Items.

Für manche IRT-Modelle (beispielsweise das Raschmodell) müssen die Items außerdem die Voraussetzung erfüllen, dass sie die Testpersonen alle gleich gut diskriminieren (vgl. Strobl, 2010, S. 20-23). Analysen von Personenfähigkeiten sind demnach unabhängig von den Aufgaben anhand derer man die Testpersonen vergleichen möchte. Ist diese Annahme für einen Datensatz nicht erfüllt, kann kein Raschmodell gerechnet werden. Es gibt aber Alternativen (Strobl, 2010).

An dieser Stelle wird nur auf das Partial-Credit-Modell näher eingegangen, da dieses zur Analyse der Daten verwendet wurde. Um die Analyse beurteilen zu können, werden im Folgenden die wesentlichen Punkte des Modells beschrieben.

Das Partial-Credit-Modell erlaubt es, auf die Fähigkeit der Personen zu schließen, deren Performanz durch das Bearbeiten von Aufgaben mit mehrstufigen Antwortkategorien erhoben wird (zum Beispiel Tests oder Fragebögen). Voraussetzung für das Partial-Credit-Modell sind ordinale Daten. Die Antwortkategorien sollten also in irgendeiner Form geordnet sein (vgl. Stüber, 2016,

S. 15). Bei einem ordinalen Skalenniveau erhalten Testpersonen durch eine Beurteilung ihrer Antworten beispielsweise einen Score, der je nach gewählter Antwortkategorie höher oder niedriger sein kann.

Für das Partial-Credit-Modell ergibt sich je Kategorie k eines Items j eine Modellgleichung, in die Parameter für Personenfähigkeit und Itemschwierigkeit eingehen. Die Itemschwierigkeit wird entlang der verschiedenen zu erreichenden Niveaus (durch die Antwortkategorien) beschrieben. Beantwortet eine Testperson i ein Item j mit den Antwortkategorien $0, \dots, m_j$, wobei m_j die Anzahl der Kategorien von Item j ist, wird durch das Partial-Credit-Modell die bedingte Wahrscheinlichkeit berechnet, dass die Antwort der Testperson i in die Kategorie c fällt, bedingt durch die Annahme, dass die Antwort entweder in Kategorie $c - 1$ oder c fällt (vgl. auch für die Modellgleichungen Strobl, 2010, S. 55):

$$P(u_{ij} = c | u_{ij} \in \{c - 1, c\}, \theta_i, \delta_{jc}) = \frac{e^{\theta_i - \delta_{jc}}}{1 + e^{\theta_i - \delta_{jc}}}$$

Der Parameter δ_{jc} beschreibt die Schwierigkeit für das Erreichen der jeweiligen Antwortkategorie c von Item j . Um die Forderung nach lokal stochastischer Unabhängigkeit zu erfüllen, wird diese Gleichung folgendermaßen in eine Gleichung für unbedingte Wahrscheinlichkeit überführt (vgl. Strobl, 2010, S. 55):

$$P(u_{ij} = c | \theta_i, \delta_{j1}, \dots, \delta_{jm_j}) = \frac{e^{\sum_{k=0}^c (\theta_i - \delta_{jk})}}{\sum_{l=0}^{m_j} e^{\sum_{k=0}^l (\theta_i - \delta_{jk})}}$$

Dabei wird definiert, dass

$$\sum_{k=0}^0 (\theta_i - \delta_{jk}) = 0, \sum_{k=0}^l (\theta_i - \delta_{jk}) = \sum_{k=1}^l (\theta_i - \delta_{jk})$$

gilt. Die Modellgleichung ist eine logistische Funktion und wird für jede Kategorie eines Items berechnet. Für jedes Item werden die $m_j + 1$ Graphen der Modellgleichungen in *einem* System dargestellt. Man nennt diesen Graphen *Category Characteristic Curve* (CCC, siehe Abbildung 7.4). Diese Kurven zeigen an, wie hoch die Wahrscheinlichkeit für eine Person mit Fähigkeit θ_i ist, mit ihrer Antwort eine bestimmte Kategorie zu treffen. Für Testpersonen mit niedriger Personenfähigkeit ist die Wahrscheinlichkeit hoch die Kategorie mit dem niedrigsten Score zu wählen. Diese Wahrscheinlichkeit sinkt mit steigendem θ_i . In der ordinalen Reihenfolge steigen die Wahrscheinlichkeiten für die Kategorien zwischen der niedrigsten und höchsten Kategorie bis zu einem Maximum an und fallen anschließend wieder ab. Ab einer bestimmten Personenfähigkeit ist die höchste Kategorie am wahrscheinlichsten. Zu jeder Personenfähigkeit θ_i addieren sich die Wahrscheinlichkeiten für alle Kategorien zusammen zu dem Wert 1.

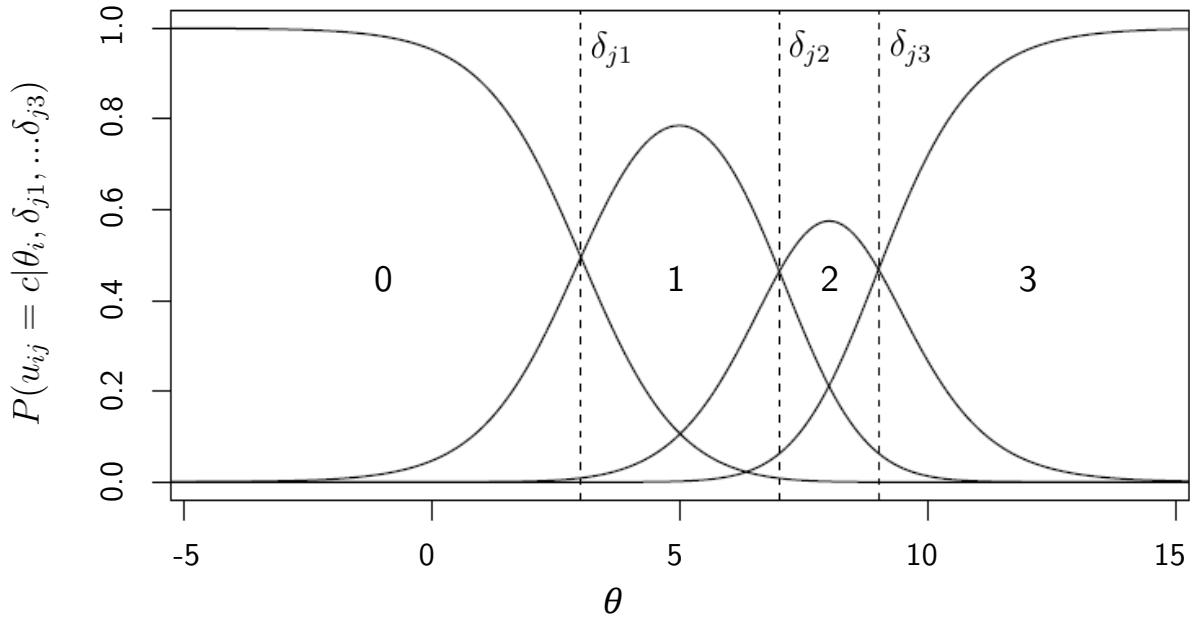


Abbildung 7.4: Der Graph ist ein Beispiel für eine Category Characteristic Curve bei Strobl (2010, S. 58).

Wie in Abbildung 7.4 an den gestrichelten Linien zu sehen ist, können auch die Schnittpunkte der Wahrscheinlichkeitsverteilungen eingezeichnet werden. Diese Schnittpunkte bezeichnet man als Schwellenparameter. Sie bilden die Kategorienschwierigkeiten β_{jk} ab (vgl. Strobl, 2010, S. 56-57).

An der Position der Schnittpunkte kann geprüft werden, ob die vorher festgelegte Ordnung der Kategorien durch die erhobenen Daten bestätigt wird. In diesem Fall müssen die Schnittpunkte zweier aufeinanderfolgender Kategorien geordnet sein: $\delta_{j1} < \delta_{j2} < \dots < \delta_{jm_j}$. Wenn diese Reihenfolge nicht eingehalten wird, sollte entschieden werden, ob einige Kategorien zusammengezogen werden müssen oder nicht modellkonforme Items aus dem Test entfernt werden müssen.

Eine Spezialform des Partial-Credit-Modells ist das sogenannte Raschmodell, welches nur zwei Kategorien (0 und 1) umfasst. Die Modellgleichung vereinfacht sich dann zu:

$$P(U_{ij} = 1 | \theta_i, \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}}$$

– Ende des Exkurses –

Die Bedingung der lokal stochastischen Unabhängigkeit wird durch die Daten aus dem PUPiL-Instrument erfüllt, da jede Testperson unabhängig von der Wahrnehmung vorheriger Situationen an jeder Stelle im Video erneut die Möglichkeit hat, eine kritische Situation adäquat wahrzunehmen. Nach Auswahl geeigneter Items wurde geprüft, ob die Daten normalverteilt sind, was auf Grund der Stichprobenzusammensetzung (vgl. Tabelle 7.1) allerdings nicht zu erwarten war. Diese Information gibt Aufschluss darüber,

ob weitere Analysen der intervallskalierten Daten parametrisch oder nichtparametrisch durchgeführt werden sollten. Nachdem die Daten zur physikspezifischen professionellen Unterrichtswahrnehmung nach eben erklärter Methode ausgewertet wurden, wurden die EAP-Reliabilität des fachdidaktischen Wissenstests überprüft, indem – wie ursprünglich für den Test vorgesehen (Gramzow, 2015) – eine Raschanalyse durchgeführt wurde. Durch die Raschanalyse des physikdidaktischen Wissenstests liegen außerdem Daten der Personenfähigkeit in Bezug auf das fachdidaktische Wissen vor, die später zur Überprüfung der Hypothesen H1 und H2 genutzt werden können.

Um die Forschungsfrage zu beantworten, wurden (a) Varianzanalysen unter Berücksichtigung der Variablen Fach, Studienfortschritt, Studiengang, Hochschule sowie Video berechnet und (b) Korrelationen zu den Variablen physikdidaktisches Wissen und Fassen eines Handlungsplans betrachtet. Ergibt sich ein signifikanter Einfluss der unter (a) aufgelisteten Variablen, stützt dies sowohl die in Bezug zur Forschungsfrage aufgestellte Hypothese H1, dass sich die professionellen Unterrichtswahrnehmung Physiklehramtsstudierender und Geographielehramtsstudierender unterscheidet, als auch Hypothese H2, dass sich die professionelle Unterrichtswahrnehmung im Laufe des Studiums verändert. Die Zusammenhänge bezüglich (b) könnten helfen, nachzuweisen, ob der PUPiL-Test die professionelle Unterrichtswahrnehmung valide misst, da theoretisch Zusammenhänge zum physikdidaktischen Wissen und der Fähigkeit, Handlungspläne zu generieren, angenommen werden. Alle Berechnungen wurden mit dem Computerprogramm R und der Oberfläche RStudio (RStudio Team, 2015) oder SPSS 26 durchgeführt (IBM Corp., Released 2017).

Da es sich bei dem Konstrukt der physikspezifischen professionellen Unterrichtswahrnehmung – wie in den Ausführungen zum Stand der empirischen Forschung gezeigt – um ein noch nicht ausreichend erforschtes Konstrukt handelt, werden Signifikanzniveaus p für die Berechnung von Zusammenhängen mit Perception, Interpretation und Decision Making mit folgender Bedeutung versehen (orientiert an Bortz & Schuster, 2010, S. 101): (a) $p \leq 0,1$ zeigt das die Testergebnisse zur Signifikanz tendieren, (b) $p \leq 0,05$ zeigt, dass das Testergebnis statistisch signifikant ist und (c) $p \leq 0,01$ bedeutet, dass das Testergebnis statistisch hoch signifikant ist.

In den folgenden Kapiteln werden die Ergebnisse der Hauptstudie präsentiert und jeweils direkt anschließend interpretiert und diskutiert. Zunächst erfolgt ein Bericht über die IRT-Skalierungen des entwickelten Tests zur Erfassung der physikspezifischen professionellen Unterrichtswahrnehmung. Anschließend folgt die IRT-Skalierung des physikdidaktischen Wissenstests. Weiter werden zur Beantwortung der Forschungsfrage Ergebnisse quantitativer Analysen zu Korrelationen zwischen professioneller Unterrichtswahrnehmung und fachdidaktischem Wissen sowie der Fähigkeit, einen Handlungsplan zu fassen, dargestellt. Außerdem werden Varianzanalysen bezüglich der Variablen Fach, Studienfortschritt, Studiengang und Video präsentiert.

7.2 Der PUPiL-Test

Dieses Kapitel ist unterteilt in die Ergebnispräsentation (a) der IRT-Skalierung der Daten zur physikspezifischen professionellen Unterrichtswahrnehmung (Unterabschnitt 7.2.1) sowie (b) der Varianzanalysen mit unabhängigen Variablen wie das Fach, Studienfortschritt etc. und (c) die Interpretation sowie Diskussion dieser Ergebnisse (Unterabschnitt 7.2.3).

7.2.1 Ergebnisse der IRT-Skalierung

Die IRT-Skalierung der Daten zur physikspezifischen professionellen Unterrichtswahrnehmung wurde mit dem R-Package TAM (Robitzsch, 2019) durchgeführt. Für die erhobenen Daten ist es sinnvoll, von einem Partial-Credit-Modell auszugehen (Strobl, 2010). Das Partial-Credit-Modell des Packages TAM wird mit einer Einschätzung der sogenannten Maximum Marginal Likelihood (ebd., S. 33) in R mit dem Befehl *tam.mml()*, berechnet. Durch das Programm wurden – wie auch theoretisch hergeleitet und durch die Distraktoren intendiert – fünf Kategorien erkannt. Die Personenfähigkeiten werden zum einen in Form von einem ganzzahligen Score ausgegeben. Zu jedem dieser Items und Kategorien gibt das Programm zusätzlich Itemschwierigkeiten und Personenfähigkeiten aus. In Tabelle 7.4 sind die Analysehandlungen und deren Zwecke sowie Ergebnisse dargestellt.

Tabelle 7.4: Analysehandlungen und ihre Zwecke sowie Ergebnisse.

Schritt	Analysehandlung: Zweck	Stichproben- größe	Anzahl der Items	Anzahl der Kategorien	Anzahl der konformen Items
1	Ausgangsposition	160	38	5	?
2	IRT-Analyse: Exkludieren von Testpersonen mit weniger als 5% der maximalen Punktzahl	160	38	5	?
3	IRT-Analyse mittels PCM: Suche konformer Items	151	38	5	0
4	Rekodieren der Daten: Erhöhen der konformen Items	151	38	3	19
5	Finale IRT-Analyse mittels PCM	151	19	3	19, EAP-Reliability = 0.698

Anhand des Scores wurden zunächst all diejenigen Testpersonen von weiteren Analysen exkludiert, die weniger als fünf Prozent der in dem Test erreichbaren Gesamtpunktzahl erreicht haben. Wenn Testpersonen nur einen so geringen Score erhalten, ist das ein Hinweis darauf, dass sie die Aufgabe nicht richtig verstanden haben oder Probleme in der Bedienung des Online-Programms hatten. Als Folge kam es zu einem Ausschluss von neun Testpersonen, sodass die neue Stichprobe für alle folgenden Berechnungen $N = 151$ Testpersonen umfasst.

In einer zweiten Analyse wurden die gleichen Berechnung wie oben nur anhand der $N = 151$ Testpersonen durchgeführt. Zusätzlich zu den Scores wird zu jeder der vom Programm berechneten fünf Kategorien außerdem je Item ausgegeben, welche Personenfähigkeit in Form einer reellen Zahl θ (Schwellenparameter) eine Testperson aufweisen muss, um den zu dieser bestimmten Kategorie gehörenden Distraktor mit einer größeren Wahrscheinlichkeit zu wählen, als einen Distraktor der nächst niedrigen Kategorie. Diesen Fähigkeitswert nennt man *Threshold*. Graphisch werden diese Informationen in sogenannten *Categorical Characteristic Curves (CCC)* dargestellt. Für jede Kategorie ergibt sich ein Funktionsgraph, der angibt mit welcher Wahrscheinlichkeit (unabhängige Variable) Personen mit einer bestimmten Fähigkeit (abhängige Variable) den Distraktor dieser Kategorie gewählt haben (vgl. beispielsweise Abbildung 7.5 mit drei Kategorien).

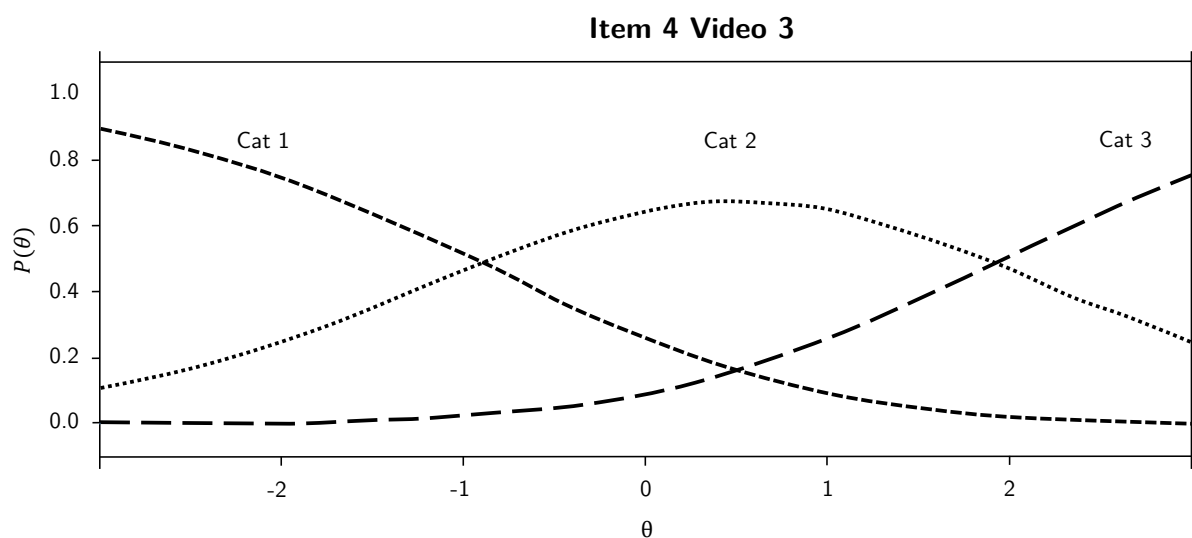


Abbildung 7.5: Categorical Characteristic Curce für Item 4 aus Video 3.

Da die Distraktoren theoretisch hergeleitet und dem Forschungsinteresse dieser Studie gemäß gestuft wurden, gibt diese Information Auskunft über die durch das PUPiL-Instrument gemessene physikspezifische professionelle Unterrichtswahrnehmung der untersuchten Testpersonen. Wenn die Items in Form eines Ordered Multiple-Choice-Designs in der Lage sind, die physikspezifische professionelle Unterrichtswahrnehmung valide zu messen, dann müssten die Thresholds für die Kategorien gemäß der Level ansteigen. Das Threshold von Level 1 (kontrafaktischer Distraktor) muss demnach einen kleineren Wert haben als das von Level 2 (Distraktor zu Classroom Management), und so weiter. Das führt dazu, dass die Schnittpunkte der Funktionsgraphen von aufeinanderfolgenden Kategorien auch in dieser Reihenfolge in den CCCs der Partial-Credit-Modell konformen Items auftreten. Im Beispiel in Abbildung 7.5 ist das der Fall. Der Funktionsgraph der Kategorie 1 (Cat 1) trifft auf den Funktionsgraphen von Kategorie 2 (Cat 2), bevor dieser den Graphen von Kategorie 3 (Cat 3) schneidet. Wird diese Reihenfolge der Thresholds in einer CCC nicht dargestellt, ist das dazugehörige Item nicht konform mit dem angesetzten Partial-Credit-Modell. Die Analyse der $N = 151$ Testpersonen und 38 Items ergab eine

EAP-Reliabilität⁴³ von 0.849. Allerdings weist keines der 38 Items die anvisierte Reihenfolge der Thresholds für ein Partial-Credit-Modell mit fünf Kategorien auf. Die Struktur der Kategorien orientiert sich an der Construct-Map (vgl. Unterunterabschnitt 6.2.3.1 zu den Ergebnissen der Entwicklungsstudie II.2). Die Ergebnisse der Analyse zeigen allerdings, dass diese eher theoretische Strukturierung die Daten nicht abzubilden vermag. Daher wurden alternative, theoretisch sinnvolle Aufteilungen für die Kategorisierung der Antworten abgeleitet (vgl. dafür zum Beispiel Gramzow, 2015; Strobl, 2010) und im gleichen Vorgehen geprüft, ob die Items mit einem Partial-Credit-Modell konform sind⁴⁴. Dabei hatte oberste Priorität, bei gleicher Anzahl an Partial-Credit-Modell konformen Items möglichst viele Kategorien aufrecht zu erhalten, um den Informationsgewinn bezüglich der professionellen Unterrichtswahrnehmung möglichst hoch zu halten:

- Es ist denkbar, dass Personen, die Kontrafaktisches wahrnehmen, eine ähnliche Personenfähigkeit aufweisen, wie Testpersonen, die eine kritische Situation gar nicht wahrgenommen haben. Daher wurde betrachtet, wie viele der 38 Items mit einem Partial-Credit-Modell konform sind, wenn die Kategorien 0 (nicht wahrgenommen) und 1 (Kontrafaktisches wahrgenommen) zusammengezogen werden. Die Items haben dann nur noch vier Kategorien. Nach der Modellierung ergibt sich ein ähnlicher Reliabilitätswert wie zuvor (EAP-Reliabilität: 0.793). Es zeigt sich aber auch, dass keines der Items mit diesen vier Kategorien zu einem Partial-Credit-Modell passt.
- Da eine physikspezifische professionelle Unterrichtswahrnehmung gemessen werden soll, scheint es auch sinnvoll, die Kategorie 3 (anderer didaktischer Aspekt) und 4 (Schülervorstellungen oder Erkenntnisgewinnung) zu einer Kategorie zusammenzuziehen. Diese beschreibt dann, dass fachdidaktisch Relevantes wahrgenommen wurde. Es bleiben wiederum vier Kategorien zum Einspeisen in das Partial-Credit-Modell. Auch bei dieser Berechnung weicht der Wert für die Reliabilität des Modells kaum von den vorherigen ab. Allerdings erweisen sich auch hier alle Items als nicht konform zum angelegten Partial-Credit-Modell.
- Eine strengere Kodierung der Antworten aller Testpersonen kodiert das nicht Wahrnehmen mit 0, das Wählen eines Distraktors mit 1 und das Wählen des Attraktors mit 2. Die Testpersonen erhalten demnach nur volle Punktzahl, wenn der Kern der kritischen Situation erkannt wurde, jede andere Art von Wahrnehmung würde aber mit einem Punkt honoriert werden. Bei diesem Modell bleiben nur drei Kategorien erhalten. Die Analyse der Daten auf Grundlage dieser drei Kategorien führt zu einem Reliabilitätswert von 0.796 für die EAP-Reliabilität. Bezüglich der Partial-Credit-Modell konformen Items ergeben sich sechs Items für Video 1, acht Items für Video 3 und fünf Items für Video 4, deren Schnittpunkte gemäß der Abfolge der Kategorien geordnet sind. Insgesamt erfüllen 19 der 38 Items das Partial-Credit-Modell mit drei Kategorien.

⁴³Die Reliabilität wird ähnlich wie in der klassischen Testtheorie berechnet. Die EAP-Reliabilität gleicht dabei am ehesten Cronbachs Alpha und wird wie dieser interpretiert (Gramzow, 2015).

⁴⁴Alle Daten sowie R-Scripts sind digital angehängt, sodass diese Schritte numerisch und graphisch nachvollzogen werden können.

Weitere Umkodierungen führten nicht zu mehr als 19 Items. Mit diesen 19 Items wurde die IRT-Skalierung mittels Partial-Credit-Modell erneut durchgeführt. Der Wert für die EAP-Reliabilität ist auf 0.698 gesunken.

Durch die Betrachtung der Schwellenparameter können Aussagen über die Testschwierigkeit getroffen werden. Diese sind zum einen in Abbildung 7.6 und zum anderen in Tabelle 7.7 dargestellt. Die mit RStudio und dem Befehl *wrightmap()* erzeugte Wrightmap in Abbildung 7.6 zeigt auf der linken Seite die Verteilung der Testpersonen (der unterste Balken repräsentiert circa 1,7 Testpersonen). Die Logit-Skala auf der rechten Seite steht sowohl für die Personenfähigkeit der Testpersonen, als auch für die Kategorienschwierigkeit. Entlang der X-Achse sind die verbleibenden 19 Items mit aufsteigenden Schwierigkeitsparameter der ersten Kategorie dargestellt. Die untere Kurve repräsentiert demnach die Personenfähigkeit, die eine Testperson aufweisen muss, um bezüglich des jeweiligen Items die erste Schwelle zu übertreten. Das entspricht dem Antwortverhalten, in dem die Testpersonen ein Video anhalten, aber einen Distraktor wählen (Cat 1 in der Abbildung 7.6). Die obere Kurve gibt dahingegen an, welche Personenfähigkeit eine Testperson aufweisen muss, um das Item den Erwartungen gemäß zu lösen und somit das Video anzuhalten und den jeweiligen Attraktor zu wählen.

Modellkonforme Items In der Abbildung 7.6 und in Tabelle 7.7 ist abzulesen, dass es einige Items gibt, die die Fähigkeit der Testpersonen, Physikunterricht physikspezifisch wahrzunehmen, besser von einer eher generischen professionellen Unterrichtswahrnehmung differenzieren (beispielsweise 3.13, 3.9, 3.10 1.7, 4.1, 4.5, 3.11) als andere Items. Mit besser oder schlechter differenzieren ist gemeint, wie weit die Schwierigkeit der beiden Kategorien 1 und 2 voneinander entfernt liegen. Umfasst die Differenz mehr als einen aber nicht alle Balken in der Wrightmap, wird hier von einer guten Differenzierung gesprochen. Ist die Differenz zu gering oder zu hoch, wird von einer schlechten Differenzierung gesprochen. Für die Items 1.3, 1.1, 4.2, 1.5, 1.8, 4.7, 4.12, 3.12, 3.6 und 4.14 liegen die Schwellenparameter eng aneinander. Sobald diese Situationen innerhalb der Videos erkannt werden, ist eine adäquate Lösung der entsprechenden Items eher wahrscheinlich. Inwiefern dies plausibel ist, wird in Unterabschnitt 7.2.3 diskutiert.

Außerdem weisen zwei Items hohe Unterschiede zwischen den Schwellenparametern auf. Für Item 3.4 ist eine sehr geringe Personenfähigkeit nötig, um zu erkennen, dass in der Situation zu dem Item etwas der Wahrnehmung einer Lehrkraft bedarf, ohne dabei den gewünschten fachdidaktischen Fokus in der Wahl der Antwortmöglichkeit zu setzen ($\theta = -0,909$). Um jedoch den physikspezifischen Kern der kritischen Situation erkennen zu können, ist eine sehr hohe Personenfähigkeit nötig ($\theta = 1,917$). Das Item 3.2 wird nur von circa der Hälfte der Testpersonen auf niedrigster Stufe gelöst und erfordert zusätzlich eine Personenfähigkeit von $\theta = 2,584$ für das korrekte Lösen des Items. Mögliche Gründe für diese verschiedenen Differenzierungsgrade werden unten in Unterabschnitt 7.2.3 diskutiert.

Insgesamt zeigt die Verteilung des Antwortverhaltens der Testpersonen, dass ungefähr die Hälfte aller Items für die Testpersonen – selbst für eine professionelle Unterrichtswahr-

nehmung mit generischem Fokus – eher anspruchsvoll sind.

Die Infit-Werte von Items beschreiben die durchschnittlichen quadrierten Residuen und sollten zwischen 0.8 und 1.2 liegen (Bond & Fox, 2007). Alle Items des Tests erfüllen diese Voraussetzung.

Tabelle 7.5: Charakterisierung der modellkonformen Items: SV – Schülervorstellungen, EG – Erkenntnisgewinnung; 0 – nicht adäquat, 1 – adäquat; „fachlich angemessen“ gibt an, ob die Vorstellungen oder Erkenntnisgewinnung der Lernenden physikalisch angemessen sind.

Item	Schwierigkeit (Cat2)	Aspekt	fachlich angemessen
1.3	leicht	SV	ja
1.1		SV	nein
1.8		EG	ja
1.5	mittel	SV	nein
1.7		SV	nein
3.10	leicht	EG	ja
3.13		EG	nein
3.11	mittel	SV	nein
3.9		SV	nein
3.12	schwer	SV	ja
3.6		EG	ja
3.4		EG	ja
3.2		SV	nein
4.2	leicht	SV	ja
4.1		SV	nein
4.12	mittel	SV	nein
4.7		SV	ja
4.5	schwer	SV	ja
4.14		SV	ja

Fünf der Items mit einer geringeren Kategorienschwierigkeit (für Cat 1) gehören zu Video 3, drei Items sind dem ersten Video zuzuordnen und nur zwei Items sind Teil von Video 4. Im Vergleich dazu zählen drei Items von Video 4, drei Items von Video 3 und zwei Items von Video 1 zu den anspruchsvolleren Items.

Zu diesen Vergleichen folgen am Ende dieses Kapitels weitere Ausführungen.

Schwierigkeit der modellkonformen Items Werden die Items nach der Personenfähigkeit sortiert, die für das Absolvieren der zweiten Kategorie – also für das korrekte Lösen des Items – notwendig ist, können die Items wie in Tabelle 7.5 charakterisiert werden. Dazu wurden die ersten sechs der 19 Items als „leicht“, die nächsten sechs als „mittel“ und die letzten sieben Items als „schwer“ eingestuft. Demnach sind drei Items von Vi-

deo 1 leicht, betreffen eine inadäquate Schülervorstellung zum Lösen von Tee und je eine adäquate Schülervorstellung und Erkenntnisgewinnung (Aufstellen einer Vermutung). Etwas schwerer sind hingegen die zwei Items, welche inadäquate Schülervorstellungen zum Schmelzen von Tee zeigen. Über Plausibilität dieser Befunde wird ebenfalls in Unterabschnitt 7.2.3 diskutiert.

Video 3 hat nach dieser Sortierung zwei leichte Items zu einer gelungenen und einer weniger gelungenen Erkenntnisgewinnung (adäquate Nutzung von Informationen aus den Medien und Aufstellen einer nicht überprüfbaren Hypothese). Nur ein Item ist mittelschwer und zeigt eine physikalisch inadäquate Schülervorstellung (Blitze haben etwas mit Kälte zu tun). Die restlichen fünf Items von Video 3 sind schwer und betreffen zwei physikalisch inadäquate Schülervorstellungen, eine adäquate Schülervorstellung (Blitze und Kälte, Elektrizität als Kleben, Elektrizität durch Ladungen) und zwei gelungene Erkenntnisgewinnungen (physikalisch adäquates Beschreiben, adäquate Nutzung von Informationen aus den Medien).

Video 4 beinhaltet nur kritische Situationen zu Schülervorstellungen. Ein Item ist dabei leicht und zeigt eine physikalisch adäquate Schülervorstellung zur Energieumwandlung. Drei Items sind mittelschwer und behandeln zwei physikalisch inadäquate und eine adäquate Schülervorstellung (Kraft ist Energie, Energieumwandlung ist Energieumwandlung, Energie wird nicht verbraucht, sondern umgewandelt). Die verbleibenden zwei Items von Video 4 sind schwer und behandeln physikalisch adäquate Schülervorstellungen zur Energieumwandlung.

Bei Betrachtung der Verteilung der Aspekte auf das ganze PUPiL-Instrument ergeben sich 14 kritische Situationen zu Schülervorstellungen und fünf kritische Situationen zu Erkenntnisgewinnung. Von diesen Items zeigen zehn gelungenes Lernen zu den Aspekten (sechs zu Schülervorstellungen, vier zu Erkenntnisgewinnung) und neun handeln von weniger gelungenem Lernen (acht zu Schülervorstellungen, eine zu Erkenntnisgewinnung).

Nicht modellkonforme Items Bei der Betrachtung der Items, die *nicht* konform mit einem passenden Partial-Credit-Modell sind, zeigen sich weitere explorative Erkenntnisse über die physikspezifische professionelle Unterrichtswahrnehmung und ihre Messbarkeit. Die Ergebnisse lassen sich ähnlich wie zu den Modell konformen Items in einer Tabelle zusammenfassen. Die nicht modellkonformen Items sind mindestens mittelschwer, aber überwiegend schwer (vgl. Tabelle 7.6). Es war demnach zu erwarten, dass die meisten dieser Items keine Varianz erzeugen und daher nichts zur Aufklärung der professionellen Unterrichtswahrnehmung beitragen würden. In Video 1 (Brown'sche Molekularbewegung) konnten vier Items nicht zur Messung des Konstrukts verwendet werden. Alle betreffen kritische Situationen, in denen Lernende Unterstützung bezüglich Erkenntnisgewinnung bedürfen. Zwei von diesen Items beziehen sich auf ein inadäquates Verständnis von (naturwissenschaftlichen) Hypothesen (Item 1.2 und 1.4, siehe für die Drehbücher Unterabschnitt 5.1.2). In Item 1.2 werden nicht überprüfbare Aussagen von Lernenden als Hypothesen verfasst.

Tabelle 7.6: Charakterisierung der nicht modellkonformen Items: SV – Schülervorstellungen, EG – Erkenntnisgewinnung; 0 – nicht adäquat, 1 – adäquat; „fachlich angemessen“ gibt an, ob die Vorstellungen oder Erkenntnisgewinnung der Lernenden physikalisch angemessen sind.

Item	Schwierigkeit (Cat1)	Aspekt	fachlich angemessen
1.2	mittel	SV	nein
1.4		EG	nein
1.6		EG	nein
1.9	schwer	EG	nein
3.7		EG	ja
3.1		EG	ja
3.3	schwer	SV	nein
3.5		EG	ja
3.8		EG	ja
3.14		EG	ja
3.15		EG	ja
4.3	mittel	SV	nein
4.4		SV	nein
4.6		SV	ja
4.13		SV	nein
4.8	schwer	SV	nein
4.9		SV	ja
4.10		SV	ja
4.11		SV	nein

Die Schülerin in Item 1.4 beschreibt, dass für Hypothesen aufgeschrieben wird, was in einem Experiment zu beobachten war. In der Mitte des Videos können zwei Lernende beobachtet werden, die für die Durchführung ihres Experiments weitere Materialien gebrauchen könnten. Sie sind sich dessen nicht bewusst und führen das Experiment daher ohne Hilfsmittel durch. Auch dieses Item (1.6) ist nicht modellkonform. Das letzte Item des Videos (1.9), welches nicht in das PUPiL-Instrument aufgenommen wurde, zeigt zwei Lernende, die die beobachteten Ergebnisse ihres Experiments verallgemeinern, ohne dazu ausreichend Evidenz anzuführen. Video 3 (Elektrizität) hatte ursprünglich die meisten Items. Von diesen sind allerdings sieben Items nicht konform mit dem Partial-Credit-Modell. Sechs dieser Items betreffen kritische Situationen, in denen Lernende eher erfolgreich lernen. Es werden adäquate Bezüge zu Alltagserfahrungen hergestellt (3.1, 3.7 und 3.8), Phänomene beschrieben (3.5), eigene Hypothesen aufgestellt und durch eigene Versuche getestet (3.13, 3.15). Lediglich eine der exkludierten kritischen Situationen behandelt eine inadäquate Schülervorstellung: Elektrisch geladene Objekte *kleben* aneinander.

Abbildung 7.6: Wrightmap für eine physikspezifische professionelle Unterrichtswahrnehmung. Die Items sind nach dem ersten Schwellenparameter sortiert.

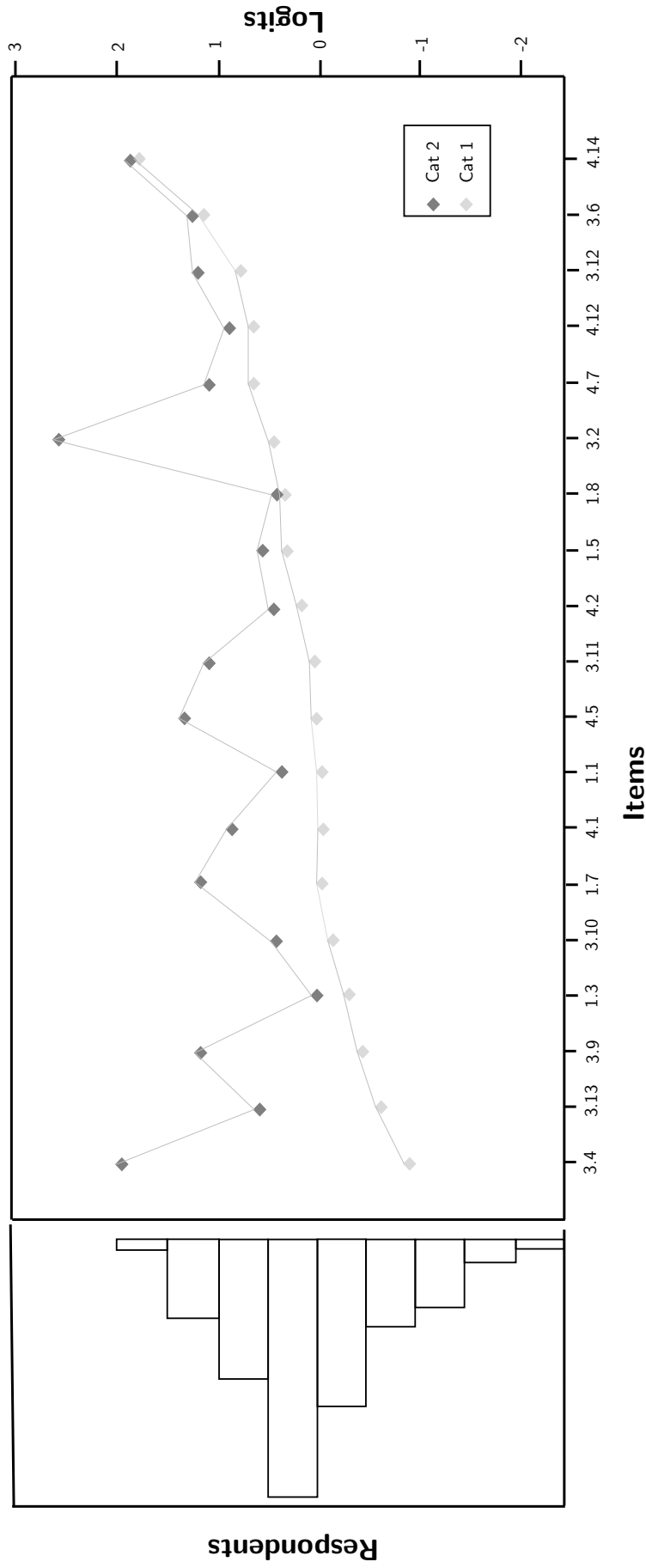


Tabelle 7.7: Thresholds und Fehler der Kategorien(auf drei Stellen nach dem Komma gerundet).

Item	Cat 1 xsi	se.xsi	Cat 2 xsi	se.xsi
1.1	-0,026	0,225	0,342	0,238
1.3	-0,302	0,240	-0,019	0,224
1.5	0,285	0,219	0,561	0,256
1.7	-0,032	0,218	1,133	0,281
1.8	0,315	0,220	0,356	0,247
3.2	0,441	0,215	2,584	0,521
3.4	-0,909	0,234	1,917	0,330
3.6	1,117	0,231	1,255	0,368
3.9	-0,453	0,225	1,142	0,275
3.10	-0,117	0,226	0,361	0,241
3.11	0,041	0,217	1,110	0,289
3.12	0,769	0,221	1,222	0,338
3.13	-0,611	0,238	0,548	0,238
4.1	-0,028	0,214	0,819	0,251
4.2	0,148	0,215	0,445	0,237
4.5	0,026	0,210	1,292	0,285
4.7	0,652	0,211	1,071	0,294
4.12	0,658	0,211	0,902	0,281
4.14	1,805	0,259	1,865	0,523

Das vierte Video handelt von Unterricht zur Energieumwandlung. Acht der 14 kritischen Situationen werden für die Messung der physikspezifischen professionellen Unterrichtswahrnehmung nicht verwendet, weil sie nicht modellkonform sind. In vier der Items wäre zu erkennen, dass die Lernenden sich eher vorstellen, dass Energie verbraucht als umgewandelt wird. Diese Vorstellung ist hinderlich für das Verstehen von Energieumwandlung (vgl. beispielsweise Duit et al., 2018, S. 167). In einem weiteren Item äußert eine Schülerin eine Vorstellung von stofflicher Energie, die von einem Objekt auf ein anderes Objekt übertragen werden kann. Inwiefern diese Vorstellung physikalisch angemessen ist, wäre zu diskutieren. Die übrigen drei Items behandeln Situationen, in denen sich adäquate Schülervorstellungen zeigen oder wenigstens anbahnen. Wie auch die modellkonformen Items, erfüllen die nicht modellkonformen Items die Voraussetzung, dass ihre Infit-Werte zwischen 0.8 und 1.2 liegen.

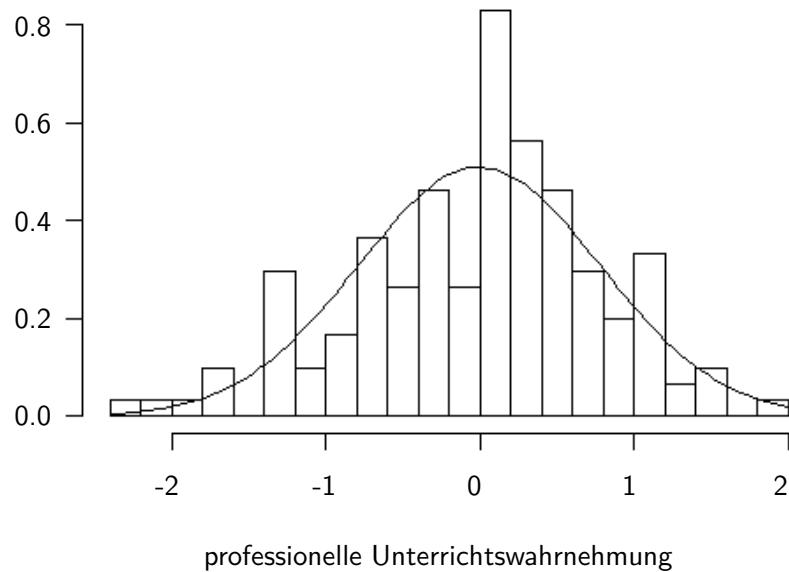


Abbildung 7.7: Normalwahrscheinlichkeitsplot für die Personenfähigkeit physikspezifische professionelle Unterrichtswahrnehmung.

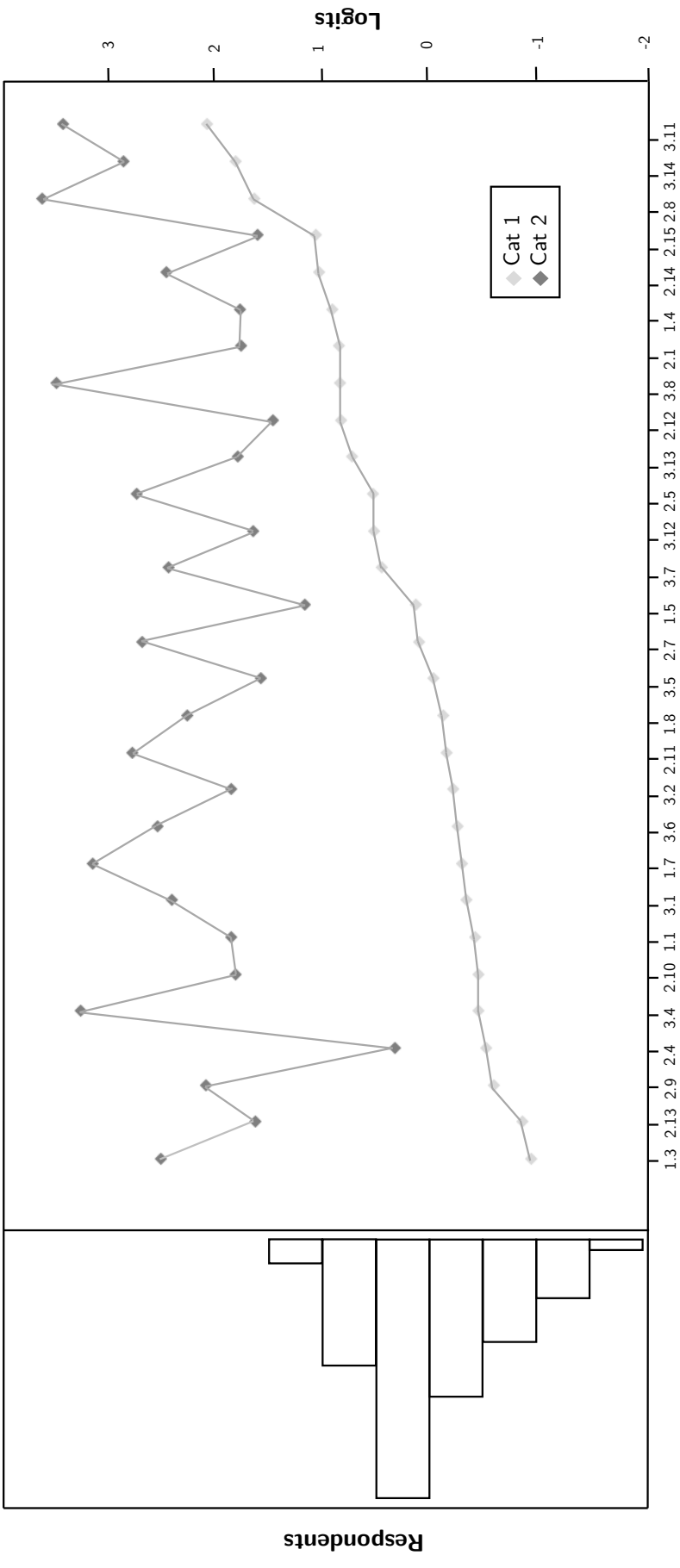
Prüfung auf Normalverteilung Um herauszufinden, ob weitere Analysen parametrisch oder nicht-parametrisch gerechnet werden müssen, wurden die Daten der IRT-Skalierung auf Normalverteilung überprüft. In einem ersten Schritt wurde graphisch mit einem Histogramm auf Normalverteilung überprüft (in Rstudio mit den Befehlen `hist(data1)`). Wie in der dazugehörigen Abbildung 7.7 zu erkennen ist, weichen die Daten vor allem für niedrige und sehr hohe Werte der Personenfähigkeit von einer Normalverteilung ab. Auch ein Shapiro-Wilk-Test ergab $W = 0.985$ mit einem p-Wert von 0.108. Da der p-Wert ungefähr 0.1 ist, wird für weitere Analysen davon ausgegangen, dass die Daten nicht normalverteilt sind. Alle folgenden Berechnungen mit den Daten (Personenfähigkeit, Itemschwierigkeit) werden daher nicht-parametrisch durchgeführt.

Rekodierung: Classroom Management Fokus Neben den Rekodierungen, die einen eher physikspezifischen Fokus beibehalten und für die Itemauswahl analysiert wurde, ist eine Rekodierung hin zu einer Classroom Management fokussierten professionellen Unterrichtswahrnehmung möglich. Innerhalb dieser explorativen Studie ist es interessant, zu verstehen, wie sich die Personenfähigkeiten der Testpersonen und die Kategorienschwierigkeiten des Tests verändern, wenn der (theoretisch gelegte) Fokus der professionellen Unterrichtswahrnehmung verschoben wird. Für diese Rekodierung erhielten die Testpersonen zwei Punkte, wenn sie eine der Classroom Management Antwortmöglichkeiten wählten, einen Punkt für alle anderen Antwortmöglichkeiten und keinen Punkt, wenn sie das Video an den entsprechenden Stellen nicht angehalten haben. In Abbildung 7.8 ist die Wrightmap der Daten des Tests nach der Rekodierung abgebildet.

Tabelle 7.8: Thresholds und Fehler für eine Classroom Management fokussierte professionelle Unterrichtswahrnehmung (auf drei Stellen nach dem Komma gerundet).

Item	Cat 1 xsi	se.xsi	Cat 2 xsi	se.xsi
1.1	-0,397	0,222	1,723	0,319
1.3	-0,881	0,234	2,332	0,377
1.4	0,846	0,222	1,651	0,382
1.5	0,121	0,218	1,081	0,282
1.7	-0,278	0,216	2,928	0,518
1.8	-0,115	0,216	2,106	0,379
3.1	0,786	0,220	1,639	0,381
3.4	-0,491	0,240	0,302	0,228
3.5	0,498	0,216	2,550	0,519
3.7	0,091	0,213	2,495	0,467
3.8	1,525	0,258	3,371	1,011
3.9	-0,555	0,224	1,942	0,343
3.10	-0,419	0,222	1,682	0,319
3.11	-0,144	0,214	2,587	0,467
3.12	0,772	0,220	1,365	0,347
3.13	-0,809	0,236	1,518	0,291
3.14	0,963	0,227	2,290	0,520
3.15	0,990	0,226	1,498	0,382
4.1	-0,318	0,211	2,240	0,376
4.2	-0,206	0,211	1,729	0,317
4.4	-0,421	0,211	3,032	0,516
4.5	-0,027	0,210	1,471	0,299
4.6	-0,242	0,209	2,365	0,399
4.7	0,414	0,208	2,266	0,430
4.8	0,782	0,215	3,246	0,720
4.11	1,930	0,279	3,186	1,011
4.12	0,484	0,209	1,534	0,331
4.13	0,678	0,211	1,669	0,360
4.14	1,684	0,257	2,659	0,722

Abbildung 7.8: Wrightmap für eine Classroom Management fokussierte professionelle Unterrichtswahrnehmung. Die Items sind nach dem ersten Schwellenparameter sortiert.



Diese Wrightmap und die Tabelle 7.8 zeigen, dass 29 Items mit diesem Partial-Credit-Modell konform sind. Die EAP-Reliabilität beträgt 0.78 und ist somit höher als die des physikspezifischen Partial-Credit-Modells. Die Verteilung der Personenfähigkeiten ist nach einem Shapiro-Wilk-Test nicht normalverteilt ($W = 0.932, p < 0.001$). Nur zwei der Items, die für das PUPiL-Instrument nun final ausgewählt wurden, sind nicht Teil dieser Itemauswahl für eine Fokussierung auf Classroom Management des Instruments. Diese Items betreffen die Situation 2 und 6 aus Video 3. Item 3.2 behandelt die Schülervorstellung, dass elektrostatische Anziehung das gleiche wie Kleben sei, die Klasse ist währenddessen ruhig und anscheinend konzentriert. Situation 3.6 hingegen beinhaltet eine adäquate Beschreibung einer Experimentierbeobachtung und auch während dieser kritischen Situation ist es im Klassenraum ruhig. Neben den Vorteilen dieser Rekodierung wie der höheren Anzahl an modellkonformen Items und einer höheren EAP-Reliabilität kann man den Kategorienschwierigkeiten in Tabelle 7.8 in Verbindung mit der Verteilung der Testpersonen gemäß Personenfähigkeit in Abbildung 7.8 entnehmen, dass dieser Test auf eine Classroom Management fokussierte professionelle Unterrichtswahrnehmung schwerer wäre, was aber auch an der Aufgabenstellung für die Testpersonen liegt. Diese ist auf eine physikspezifische professionelle Unterrichtswahrnehmung ausgelegt.

Nur für drei Items wären Testpersonen überhaupt in der Lage, den kritischen (nicht fachspezifischen) Kern der Situation zu erkennen. Für 26 Items sind dafür aber alle Testpersonen fähig zu erkennen, dass die kritischen Situationen der Items ihre professionelle Wahrnehmung bedürfen. Es war allerdings nicht die Aufgabe der Testpersonen, Aspekte des Classroom Managements wahrzunehmen, daher sind diese Befunde nicht beunruhigend.

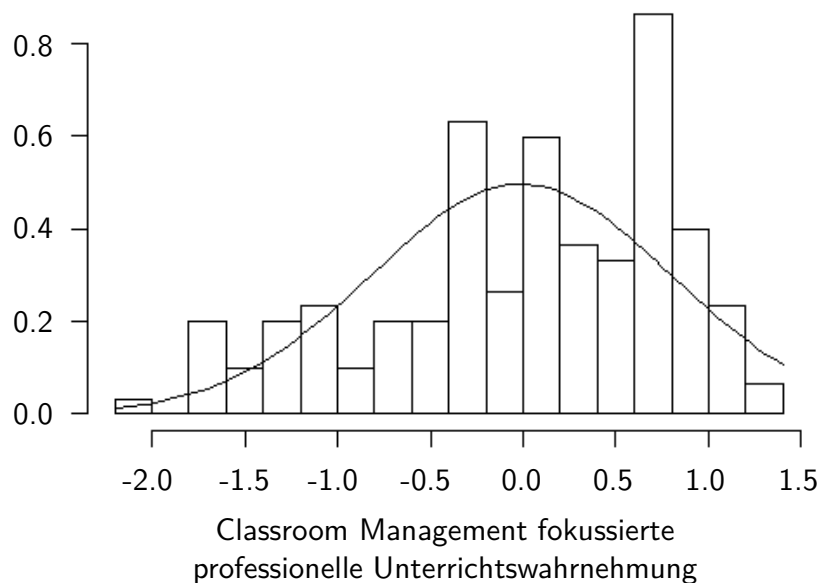


Abbildung 7.9: Normalwahrscheinlichkeitsplot für die Personenfähigkeit Classroom Management fokussierte professionelle Unterrichtswahrnehmung.

Bezüglich der Daten einer Classroom Management bezogenen professionellen Unterrichtswahrnehmung ergibt sich ebenso wie für eine physikspezifische professionelle Unterrichtswahrnehmung keine Normalverteilung (vgl. Abbildung 7.9).

7.2.2 Abhängigkeit der physikspezifischen professionellen Unterrichtswahrnehmung von weiteren Variablen

Neben der Analyse der Items wurde untersucht, ob und inwiefern die physikspezifische professionelle Unterrichtswahrnehmung – gemessen an diesen Items – von den Variablen Fach, Studienfortschritt, Studiengang, Hochschule und nicht gesehenes Video abhängt. Die Ergebnisse dieser Untersuchung sind in Tabelle 7.9 gemeinsam mit den Ergebnissen der nächsten beiden Abschnitte dargestellt (siehe zunächst unterer Teil der Tabelle).

Tabelle 7.9: Zusammenfassung der Zusammenhänge bezüglich der physikspezifischen professionellen Unterrichtswahrnehmung.

Professionelle Unterrichtswahrnehmung	Variable		Signifikanz
Korrelationen			
	θ_{FD}	$\tau = 0.113$	0.049
	Decision Making	$\tau = 0.370$	<0.001
	Selbsteinschätzung Uniphysik	$\tau = 0.153$	0.007
	Selbsteinschätzung Unididaktik	$\tau = 0.145$	0.011
Kruskal-Wallis			
	Fach	$\chi^2 = 7.592$	0.023
	Physik ~ Geographie	$z = 2.753$	0.018 adj.
	Studienfortschritt	$\chi^2 = 0.009$	0.925
	Studiengang	$\chi^2 = 7.800$	0.050
	Lehramt für Gymnasien ~ Lehramt an Berufsschulen	$z = 2.572$	0.061 adj.
	Hochschule	$\chi^2 = 2.079$	0.354
	Video_x	$\chi^2 = 2.235$	0.327

Varianzanalysen mit Hilfe des Kruskal-Wallis-Tests mit den unabhängigen Variablen Fach, Studienfortschritt, Studiengang und Hochschule geben somit weitere Informationen, um Zusammenhänge zwischen den in der Erhebung inkludierten Variablen aufzuklären und

somit Aussagen über die Validität des PUPiL-Instruments und der daraus resultierenden Interpretationen zu treffen.

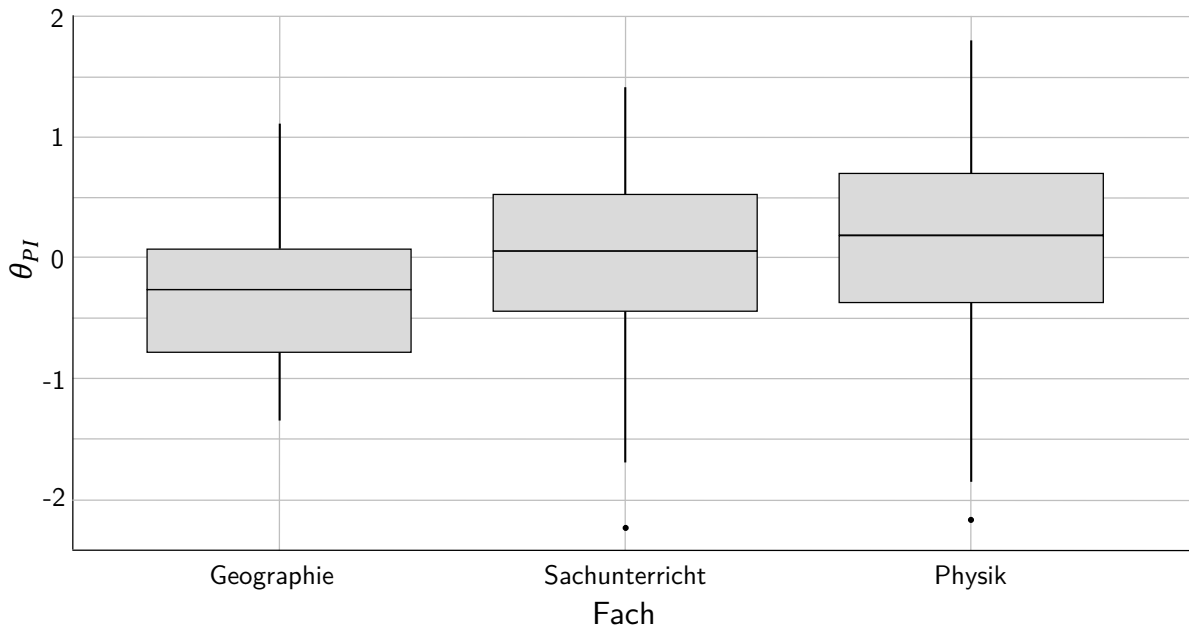


Abbildung 7.10: Boxplots zur Varianzanalyse von der physikspezifischen professionellen Unterrichtswahrnehmung mit Hilfe des Fachs.

Nicht signifikant werden die Gruppenvergleiche zum Studienfortschritt (Bachelor oder Master), zur besuchten Hochschule während der Erhebung und zum Video, welches die Testpersonen nicht gesehen haben.

Bezüglich des studierten Fachs (für das Lehramtsstudium) ergibt sich insgesamt ein signifikanter Unterschied ($z = 2.753$, $p = 0.023$ adjustiert). Auf Gruppenvergleichsebene bedeutet das einen signifikanten Medianunterschied zwischen den Fächern Geographie und Physik (siehe Abbildung 7.10)⁴⁵.

Ein weiterer Unterschied besteht zwischen den Gruppen Lehramt für Gymnasien und Lehramt für Berufsschulen bezüglich der Variable Studiengang (siehe Abbildung 7.11). Damit ist die physikspezifische professionelle Unterrichtswahrnehmung tendenzielle statistisch signifikant abhängig von dem Studiengang der Testpersonen. Es ergab sich eine Teststatistik von $z = 2.572$ und ein adjustierter p-Wert von 0.061.

Bezüglich der Variablen Erfahrung und physikalisches Schulwissen aus der Selbsteinschätzung ergeben sich keine signifikanten Zusammenhänge.

Es besteht allerdings eine hoch signifikante Korrelation zwischen der professionellen Unterrichtswahrnehmung und der Selbsteinschätzung zum physikalischen Wissen aus der Hochschule (vgl. auch im Streudiagramm Abbildung 7.12). Der Zusammenhang ist mit $\tau = 0.153$ allerdings schwach ($p = 0.007$).

⁴⁵Für die Berechnung der Varianzanalyse wurde nach dem Kruskal-Wallis-Test post-hoc ein Dunn-Test durchgeführt. Die Signifikanz wurde dann anhand des adjustierten p-Wertes bestimmt.

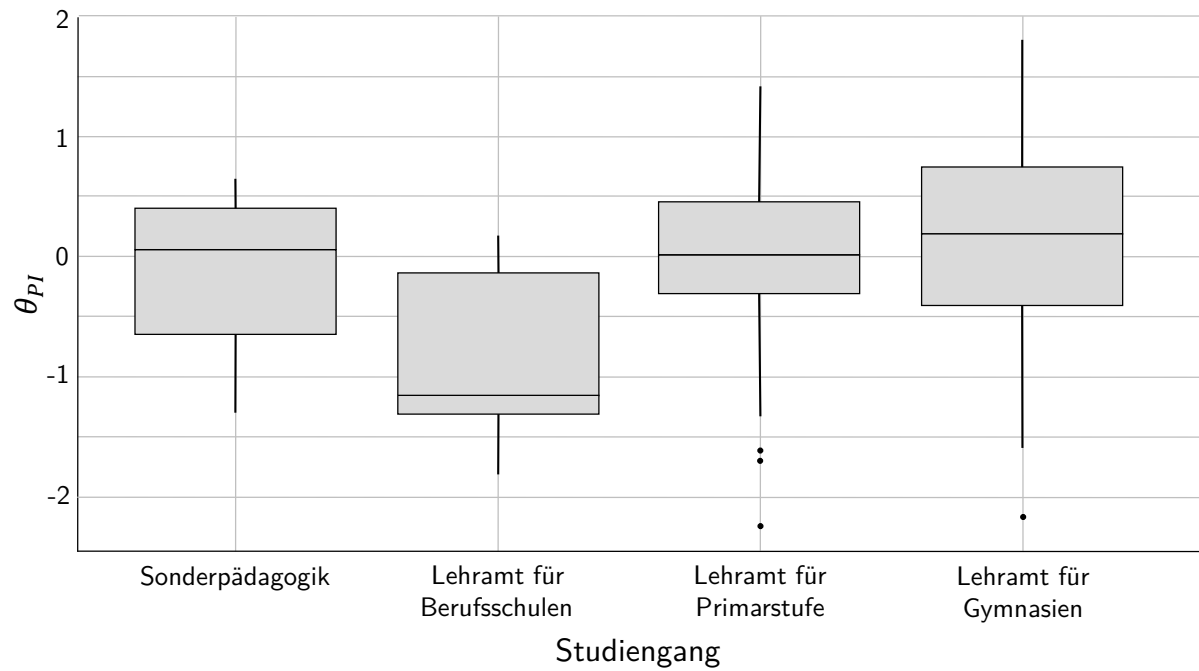


Abbildung 7.11: Boxplots zur Varianzanalyse von der physikspezifischen professionelle Unterrichtswahrnehmung mit Hilfe des Studiengangs.

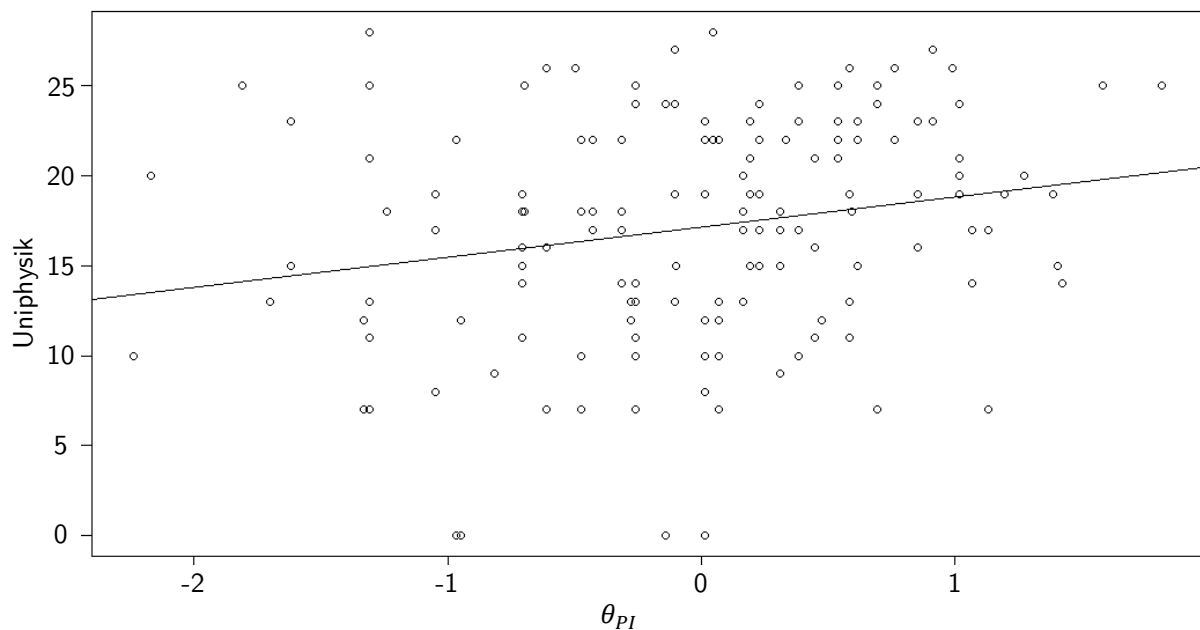


Abbildung 7.12: Streudiagramm zum physikalischen Wissen an der Hochschule (Y-Achse) und zur physikspezifischen professionellen Unterrichtswahrnehmung (X-Achse) mit einer Ausgleichsgeraden.

Des Weiteren wurde ein signifikanter, aber schwacher Zusammenhang zwischen der Selbsteinschätzung der Testpersonen bezüglich des erworbenen didaktischen Wissens in der Hochschule und der physikspezifischen Unterrichtswahrnehmung festgestellt ($\tau = 0.145$, $p =$

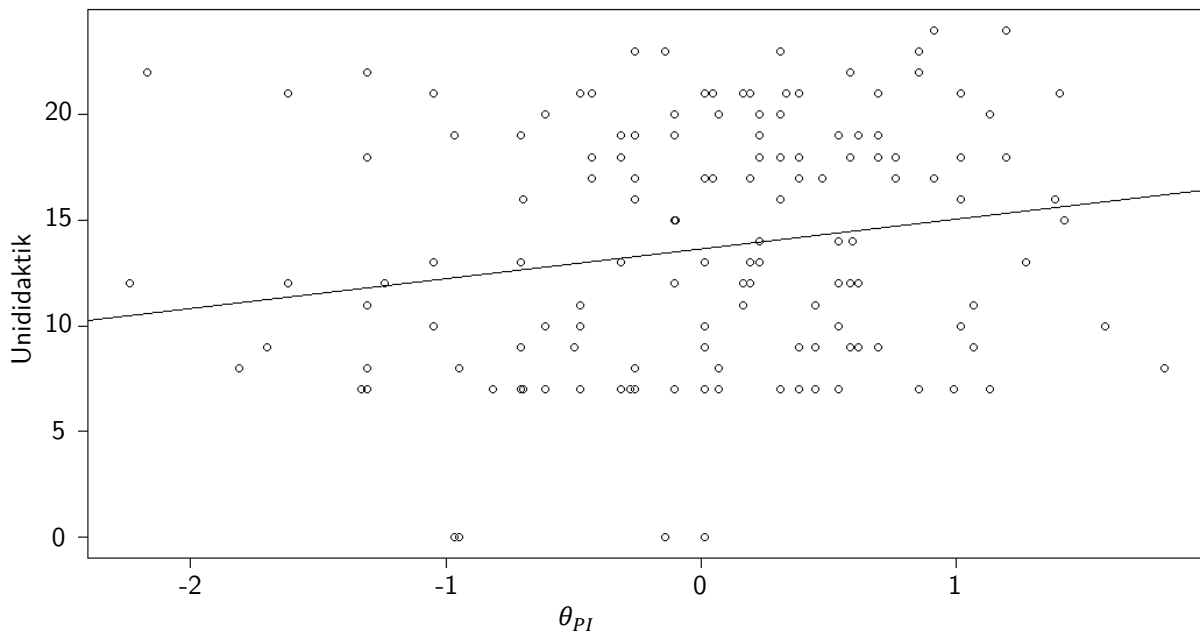


Abbildung 7.13: Streudiagramm zum physikdidaktischen Wissen an der Hochschule (Y-Achse) und zur physikspezifischen professionellen Unterrichtswahrnehmung (X-Achse) mit einer Ausgleichsgeraden.

0.011, vergleiche auch im Streudiagramm Abbildung 7.13). Eine Analyse, ob der Studiengang diesen Zusammenhang moderiert, ergab keine signifikanten Ergebnisse.

7.2.3 Interpretation und Diskussion

In diesem Kapitel werden die oben dargestellten Ergebnisse interpretiert. Die Analyse der Daten hat zum einen zur Aufklärung der Frage nach Validität beigetragen und zugleich explorative Einblicke in die professionellen Unterrichtswahrnehmung der Testpersonen gegeben.

Die Datengrundlage wurde als erstes um die Daten von Testpersonen verringert, die weniger als fünf Prozent der erreichbaren Punkte erhalten haben. Damit wurden 5,6% aller Testpersonen ausgeschlossen. Mögliche Gründe für das unterdurchschnittliche Abschneiden der Testpersonen könnte sein, dass sie nicht motiviert waren, diesen Test zu absolvieren, dass sie tatsächlich nur über eine gering ausgeprägte professionelle Unterrichtswahrnehmung verfügen oder den Test abgebrochen haben, ohne dies der Testleitung mitzuteilen. Es wird womöglich nicht jeder studierenden Person deutlich, warum es sinnvoll ist, solche Instrumente zu entwickeln und es daher einer konzentrierten Bearbeitung bedarf. Als Folge könnten sie weniger motiviert in die Erhebungssituation eingestiegen sein. Wie beispielsweise Hattie (2014) gezeigt hat, hat die Motivation einen Einfluss auf die Leistung von Lernenden, womit sich mögliches schlechtes Abschneiden der Testpersonen auch erklären ließe.

Die Daten der Hauptstudie wurden einem (Partial-Credit-) Modell angepasst, um ein geschlossenes Instrument entwickeln zu können. Selbstverständlich könnten die Daten auch mit anderen Methoden in Hinblick auf die professionelle Unterrichtswahrnehmung untersucht werden. Dies geschah in dieser Arbeit allerdings nur insoweit, wie es zur Beurteilung der Validität des PUPiL-Instruments beitragen konnten.

Das ursprünglich theoretisch gewählte, aus fünf Kategorien bestehende Modell zur Beschreibung der Daten hat sich empirisch als nicht haltbar erwiesen. Die Schwellenparameter der Items liegen nicht in der korrekten Reihenfolge vor. Erst wenn die mittleren Kategorien (kontrafaktischer, anderer didaktischer und Classroom Management Aspekt) zusammengezogen werden, ergibt sich eine konsistente Skala für einen Test, der aber für die betrachtete Stichprobe relativ schwer ist. Dafür, dass sich die theoretischen Annahmen, dass sich die professionelle Unterrichtswahrnehmung im Verlauf des Studiums entwickelt und sich eine Physikspezifität zeigt, nicht bestätigt haben, gibt es mehrere hypothetische Gründe:

- (a) Die professionelle Unterrichtswahrnehmung ist entweder vollständig oder gar nicht in der Art ausgebildet. Daher kann das Instrument keine fachspezifischen Abstufungen nachweisen. Die Parameter der Kategorienschwierigkeiten unterstützen dabei eher die Vermutung, dass die professionelle Unterrichtswahrnehmung noch *nicht ausreichend* ausgebildet ist. Es vollzieht sich demnach anscheinend keine Progression der professionellen Unterrichtswahrnehmung während des Studiums. Allerdings ließe sich die Fähigkeit in unterrichtsnahen Situationen trainieren.
- (b) Die professionelle Unterrichtswahrnehmung lässt sich nicht in fachspezifische und nicht-fachspezifische Anteile unterteilen.
- (c) Das PUPiL-Instrument dient nicht zur Unterscheidung der fachspezifischen und nicht-fachspezifischen Anteile der professionellen Unterrichtswahrnehmung.

Mit den vorliegenden Daten kann nicht erklärt werden, welche dieser Gründe zutreffen. Allerdings kann in Folgestudien überprüft werden, ob und wie einzelne Gründe ausgeschlossen oder bestätigt werden können:

- zu (a) Wenn eine mangelnde Progression der professionellen Unterrichtswahrnehmung tatsächlich ein Grund für die Ergebnisse dieser Studie ist, sollte eine Stichprobe untersucht werden, deren Testpersonen hinsichtlich der professionellen Unterrichtswahrnehmung mithilfe von wirksamen Förderangeboten gefördert wurden. Des Weiteren sollte der Test mit einer Stichprobe von erfahrenen Lehrkräften durchgeführt werden, weil diese auf Grund ihrer Lehrerfahrungen eine ausgeprägtere (aber nicht zwingend fachspezifischere) professionelle Unterrichtswahrnehmung haben müssten (vgl. Abschnitt 3.1 zum Experten-Novizen-Vergleich). Zeigt sich in den Daten dann ein ähnliches Bild, wie in dieser Studie, ist das Instrument nicht sensitiv für diesen plausiblen Effekt, der in Studien von beispielsweise Carter et al. (1988) oder Oser et al. (2010) bereits nachgewiesen wurde.

- zu (b) Möglicherweise existieren, aber konfundieren die verschiedenen Aspekte der professionellen Unterrichtswahrnehmung, wenn sich Studierende eher komplexe Unterrichtsvideos ansehen, wie sie in dieser Studie verwendet wurden. Wie bereits in Abschnitt 3.1 zum Experten-Novizen-Vergleich beschrieben, könnte es Studierende überfordern, mit komplexen und realitätsnahen Unterrichtsvideos zu arbeiten (Kramer et al., 2017). Um herauszufinden, ob die professionelle Unterrichtswahrnehmung in fachspezifische und generische Aspekte unterteilbar ist, sollte ein Instrument entwickelt werden, in dem je kritischer Situation ganz deutlich ausschließlich der eine (fachspezifische) oder der andere (generische) Aspekt der professionellen Unterrichtswahrnehmung angeregt wird. Diese Items wären dann aber weniger authentisch.
- zu (c) Das PUPiL-Instrument ist allerdings mit der Reduktion auf drei Kategorien in der Lage, zumindest die sehr grobe Kategorie „Schülervorstellungen und Erkenntnisgewinnung“ von allen anderen wahrzunehmenden Aspekten zu unterscheiden. Dabei bleibt es, wie oben gezeigt, für einen Großteil der Testpersonen schwer, die kritische Situation physikspezifisch wahrzunehmen. Folgt man der Theorie von Situations-typen und Handlungsschemata nach Wahl (2001) ließe sich der Befund dadurch erklären, dass Testpersonen mit Lehrerfahrung und Wissen ihre Aufmerksamkeit eher erfahrungsgeleitet und nicht aufgabengeleitet auf Unterrichtsaspekte fokussiert haben. Es sei daran erinnert, dass die Lehramtsstudierenden selber circa 12 Jahre Erfahrungen mit Unterricht (nämlich aus der Perspektive der Lernenden) gesammelt haben. Unter diesen Umständen kommt es dann dazu, dass die Erfahrungen einen größeren Einfluss auf die Auswahl der Kategorien zu den kritischen Situationen haben als die Aufforderung, physikspezifisch wahrzunehmen. Das PUPiL-Instrument besteht aber grundlegend aus den Videos und der Aufgabenstellung gefolgt von den geschlossenen Fragen zu jeder kritischen Situation. Beachten die Testpersonen die Aufgabenstellung nicht, kann das PUPiL-Instrument nicht zu validen Interpretationen führen. Um sichere Aussagen über die Validität des PUPiL-Instruments aus diesen Befunden ableiten zu können, müssten die Testpersonen befragt werden, was sie in ihrer professionellen Unterrichtswahrnehmung gelenkt hat – die Aufgabenstellung oder externe Faktoren. Dies ist nachträglich nicht mehr möglich.

Dass der Test zu schwer scheint, kann wiederum mehrere Gründe haben. Auch hier ließe sich vermuten, dass sich die professionelle Unterrichtswahrnehmung im Studienverlauf noch nicht entwickelt, was möglicherweise an einer fehlenden oder nicht wirksamen Förderung liegt. Diese Vermutung kann unter anderem durch die Berechnung des Zusammenhangs zwischen der Personenfähigkeit θ der professionellen Unterrichtswahrnehmung und dem Studienfortschritt der Testpersonen überprüft werden. Wie die entsprechende Varianzanalyse aber gezeigt hat, ist die Fähigkeit, Physikunterricht professionell wahrzunehmen, offenbar nicht abhängig davon, ob die Testpersonen im Bachelor- oder Masterstudium sind. Es hängt vielmehr davon ab, welches Fach und in welchem Studiengang sie Lehramt studierten. Dies ist ein Indiz dafür, dass die Testpersonen die Aufgabenstellung ernst genommen haben, weil es sonst den Fächerunterschied nicht gäbe. Das heißt, dass

für das Wahrnehmen relevante Inhalte und Fähigkeiten über den Studienverlauf keine systematische Progression erfahren, Lerngelegenheiten sich eher unsystematisch über das gesamte Studium verteilen und professionelle Unterrichtswahrnehmung während der universitären Lehramtsausbildung nur in geringem Maße gefördert wird.

Zum anderen könnte eine Ursache für die eher geringen Werte in den Personenfähigkeiten auch ein fehlendes oder nur gering ausgeprägtes fachdidaktisches Wissen sein. Nach der Kompetenzmodellierung von Blömeke, Gustafsson & Shavelson (2015) müsste es einen Zusammenhang zwischen dem Professionswissen und damit auch mit dem fachdidaktischen Wissen und der professionellen Unterrichtswahrnehmung geben. Diese Vermutung wird in einem eigenen Abschnitt durch Berechnungen von Korrelationen überprüft (vgl. Abschnitt 7.3 weiter unten).

In der Ergebnispräsentation wurde bereits beschrieben, wovon die ausgeschlossenen und beibehaltenen Items handeln. Es soll nun interpretiert werden, warum einige Items durch das dreistufige Partial-Credit-Modell beschrieben werden können und die anderen nicht. Wie oben bereits gezeigt, hat Video 3 die meisten eher schweren Items (vgl. Tabelle 7.5). Es könnte sich die Frage stellen, ob die Personenfähigkeit θ der Testpersonen von den Videos abhängt, die sie gesehen haben. Jede Testperson hat nur zwei der drei Videos gesehen und bearbeitet. Wie allerdings in den Ergebnissen präsentiert, wird eine Varianzanalyse bezüglich der Gruppenvariable Video_x nicht signifikant. Die Fähigkeit, Physikunterricht professionell wahrzunehmen, war in diesem Test nicht abhängig von der Kombination an Videos, die die Testpersonen bearbeitet haben. Das erklärt allerdings noch nicht, ob es nicht doch bestimmte Schülervorstellungen oder Prozesse der Erkenntnisgewinnung gibt, deren Wahrnehmung Testpersonen leichter oder schwerer fällt.

Bei der Betrachtung der einzelnen Videos fällt auf, dass häufig die Wahrnehmung genau eines Aspekts des Unterrichtsgeschehens zum Ausschluss der Items aus dem Video geführt hat. In Video 1 sind es die Vermutungen, welche von den Lernenden unter dem Deckmantel der Hypothesen aufgestellt werden. Die entsprechenden Items erfüllen das ursprünglich angelegte Partial-Credit-Modell nicht. Möglicherweise ist den meisten Studierenden selbst nicht bekannt, was eine Hypothese von einer Vermutung unterscheidet und nehmen in diesen kritischen Situationen darum eher Aspekte des Classroom Managements oder andere didaktische Aspekte wahr.

Von Video 3 werden vor allem die Items exkludiert, in denen Lernende erfolgreiche Erkenntnisgewinnung zeigen (durch das Formulieren von Beobachtungen oder das Aufstellen von Analogien). Da dies für Video 1 nicht gilt, kann vermutet werden, dass die professionelle Unterrichtswahrnehmung themenspezifisch ist. Das bedeutet, dass in einigen Themengebieten eher inadäquates Lernverhalten der Lernenden wahrgenommen wird, weil die Wahrnehmenden das Wissen über diese haben und in anderen Themengebieten die inadäquaten Aspekte nicht wahrgenommen werden, weil die Wahrnehmenden selbst nicht über das Wissen verfügen oder es nicht anwenden können (vgl. beispielsweise Treagust, Chittleborough & Mamiala, 2002; Wider, Schwichow & Mikelskis-Seifert, 2019). Zentral in Video 4 ist beispielsweise die Schülervorstellung des Energieverbrauchs. Der Großteil der exkludierten Items betrifft eben die kritischen Situationen, in denen Lernende eine

Energieverbrauchsvorstellung äußern oder von dieser ausgehend die physikalisch adäquatere Vorstellung einer Umwandlung entwickeln.

Zu den Items, die in das PUPiL-Instrument aufgenommen wurden, gehören sowohl gelingende als auch weniger gelingende Lernsituationen. Die Vermutung, dass die Wahrnehmung von Lehrkräften häufiger auf nicht gelingende Lernprozesse fällt, kann demnach für die Stichprobe der vorliegenden Studie nicht bestätigt werden. Allerdings ist dies auch kein Widerspruch zu der Annahme, dass (erfahrene) Lehrkräfte eher auf atypische und weniger gelingende Lernprozesse aufmerksam werden. Erfahrene Lehrkräfte haben – anders als die Testpersonen in der vorliegenden Studie – viel Unterrichtserfahrung und Erwartungen an das Ergebnis der Unterrichtsstunde, welche ihre professionelle Unterrichtswahrnehmung beeinflussen.

In der Wrightmap (siehe Abbildung 7.6) ist zu erkennen, dass einige Items die Testpersonen besser differenzieren als andere Items. Zur Erinnerung: Mit besser oder schlechter differenzieren ist gemeint, wie weit die Kategorienschwierigkeiten der beiden Kategorien 1 und 2 voneinander entfernt liegen. Es gibt Item-Paare, die eine ähnliche Kategorienschwierigkeit für die Kategorie 2, also das physikspezifische professionelle Unterrichtswahrnehmen haben, aber verschieden gut differenzieren. Mit Blick auf die geschlossenen Fragen kann dies für einige Items mit der Art der Distraktoren erklärt werden. Für das nicht physikspezifische Wahrnehmen (Kategorie 1) von Item 3.4 benötigt eine Testperson wie oben beschrieben nur eine relativ niedrige Personenfähigkeit, während sie für die physikspezifische professionelle Unterrichtswahrnehmung eine sehr hohe Fähigkeit benötigen. Für Item 4.14 benötigen die Testpersonen für die Wahl des Attraktors eine ähnlich hohe Personenfähigkeit wie bei Item 3.4, jedoch ist die notwendige Personenfähigkeit für die Wahl eines Distraktors nicht viel niedriger. Das bedeutet, dass die kritische Situation zu Item 4.14 für die Stichprobe insgesamt schwer professionell wahrzunehmen ist. Der Unterschied zwischen Item 3.4 und 4.14, der vermutlich zu diesen unterschiedlich guten Differenzierungen der Items führt, liegt in der Art der Distraktoren. In der kritischen Situation zu Item 3.4 geschieht vieles parallel. Dies wird auch in den Antwortmöglichkeiten widergespiegelt: Lernende scheinen zu schlafen, der aktive Schüler hat das Wort seines Vorredners unterbrochen, drückt sich dabei aber fachsprachlich korrekt aus. Der Kern der kritischen Situation ist die Wahl an Informationen, die der Schüler zum Verstehen eines elektrostatischen Phänomens heranzieht (Erkenntnisgewinnung). Item 4.14 handelt dagegen von einer kritischen Situation, in der ein Schüler Reibungskraft mit Energie gleichsetzt. Während seiner Aussage sind die anderen Lernenden ruhig und aufmerksam, lediglich der aktive Schüler benutzt Gestik, um seine Ausführungen zu untermauern.

Wenn Komplexität als Grad der Differenziertheit und Integriertheit verstanden wird, bedeutet das für komplexen Unterricht eine hohe Anzahl an Ereignissen, „die durch ein dichtes Netz von Beziehungen miteinander verbunden werden“ (vgl. Martin, 1988, S. 294-295). In den Unterrichtsvideos zeichnet sich das durch die gleichzeitigen und teils verknüpften Aktivitäten vieler Lernender sowie gegebenenfalls viele simultan auftretende Lernhindernisse aus. Mit diesem Verständnis von Komplexität könnte geschlussfolgert werden, dass die kritische Situation 4.14 mit einem geringen Grad an Komplexität ausgestattet ist, wohingegen die kritische Situation 3.4 eher komplex ist. Es stellt sich die Frage, ob es

Lehrkräften in einer eher komplexen kritischen Situation wie 3.4 nicht schwerer fallen müsste, ihre Aufmerksamkeit auf fachliches Lernen zu lenken. Was bestimmt die Schwierigkeit eines Items? Wie anspruchsvoll der Attraktor ist oder wie groß beziehungsweise interessant die Ablenkungen sind? Über diese Fragen kann derzeit nur gemutmaßt werden. Daher werden in dieser Arbeit keine Antworten auf diese Fragen gegeben.

Ähnliche Diskussionen lassen sich für das Item 3.13 (gut differenzierend) im Vergleich zu den eher schlecht differenzierenden, aber bezüglich des physikspezifischen Wahrnehmens gleich schwierigen Items 3.10, 1.1, 4.2, 15. und 1.8 führen. In der kritischen Situation zu Item 3.13 wird von zwei Lernenden in Partnerarbeit eine nicht überprüfbare Hypothese formuliert. Die Distraktoren lenken die Aufmerksamkeit der Testpersonen auf die gute Arbeitsteilung zwischen den beiden Lernenden, auf die sprachlichen Schwierigkeiten, die der Schüler beim Lesen der Aufgaben hat und auf einen Schüler, der aufgrund fehlender Optionen alleine arbeiten muss. Diese Situation scheint in Hinblick auf simultan auftretende Lernhindernisse komplex zu sein. Die Situationen zu den Items 3.10 und 1.8 sind nach dieser Unterscheidung ähnlich komplex und haben eine ähnliche Schwierigkeit (für Details zu den Distraktoren siehe Anhang D in der Entwicklungsstudie II). Die anderen drei Situationen, deren Kategorienschwierigkeit (von beiden Kategorien in etwa gleich) ähnlich zu der Schwierigkeit von Kategorie 2 des Items 3.13 ist, sind nach der obigen Definition von Komplexität allerdings weniger komplex. Dass die Kategorienschwierigkeiten der beiden Kategorien bei diesen drei kritischen Situationen nah aneinander liegen, ist nachvollziehbar, da in den entsprechenden Situationen nicht viel Ablenkendes geschieht. Warum diese aber nicht leichter sind als eine Situation, in der mehrere Vorkommnisse simultan die Aufmerksamkeit der Testpersonen fordern, ist fraglich. Halten Testpersonen die Videos an den entsprechenden Stellen an, wählen sie auch sehr wahrscheinlich den Attraktor.

In zukünftigen, ähnlichen Studien könnte die Unklarheit über den Einfluss von Komplexität damit beleuchtet werden, dass sehr ähnliche kritische Situationen bewusst in abgestufter, durch ein Expertenurteil gesicherter Komplexität entwickelt und die Wahrnehmung dieser untersucht wird.

In Abschnitt 4.2 zur Validität wurde erklärt, dass für eine hohe Validität der Interpretation der Daten notwendig ist, dass die verschiedenen Aspekte des Konstrukts repräsentativ auf die Items verteilt sind. Nach der Reduktion auf drei Kategorien und Auswahl von 19 Items ist das Verhältnis von Items zu Schülervorstellungen und zu Erkenntnisgewinnung nicht repräsentativ. Diese Tatsache muss man bei der Beurteilung der Validität des Tests mit in Betracht ziehen. Um das Verhältnis möglichst repräsentativ zu halten, hätten noch mehr Items in die Entwicklungsstudie I eingehen müssen. Dann bliebe nach der Durchführung und Analyse tatsächlich eine Auswahl an Items bestehen. In der vorliegenden Studie mussten alle verbleibenden Items übernommen werden, um ausreichend viele Items und somit Informationen über die Testpersonen zu erhalten. Mit diesen 19 Items erreicht das PUPiL-Instrument eine EAP-Reliabilität von 0.698. Diese Reliabilität ist akzeptabel, da es sich beim PUPiL-Instrument um eines der ersten Instrumente handelt, dass eine große Nähe zur unterrichtlichen Handlung aufbaut. In dieser Studie wurde ein größerer Fokus auf die Sicherung der Validität gelegt, um in der Lage zu sein, wirk-

lich die physikspezifische professionelle Unterrichtswahrnehmung messen zu können. Die Genauigkeit der Messung kann von dem PUPiL-Instrument ausgehend verbessert werden.

Bei der Verlagerung des kritischen Kerns einer wahrzunehmenden Situation auf Aspekte des Classroom Managements ist zu erkennen, dass mehr Items mit dem angesetzten Partial-Credit-Modell mit drei Kategorien (0 – nicht wahrnehmen, 1 – physikspezifische Distraktoren, 2 – Classroom Management Attraktor) konform sind als mit dem physikdidaktischen Fokus. Das bedeutet, dass es mehr Items aus den ursprünglich 38 Items gibt, bei denen fähigere Personen tatsächlich auch einen der zwei Attraktoren, also die höher bewertete Kategorie (Classroom Management Aspekte wahrnehmen) wählen. Die Schwellen der Kategorienschwierigkeit sind für mehr Items in erwarteter Form angeordnet als es für die physikspezifische professionelle Unterrichtswahrnehmung der Fall ist. Dies indiziert, dass die professionelle Unterrichtswahrnehmung möglicherweise eher auf Situationen gerichtet wird, in denen Classroom Management Aspekte lernrelevant werden. Der Befund, dass mehr Items mit einem Partial-Credit-Modell konform sind, wenn ein Classroom Management Fokus gesetzt wird, wird allerdings verzerrt durch die Tatsache, dass immer zwei Antwortmöglichkeiten zu Classroom Management vorhanden waren. Die Ratewahrscheinlichkeit für den Attraktor mit einem Classroom Management Fokus ist demnach höher, als die mit einem physikspezifischen Fokus. Außerdem ist es nicht für alle Classroom Management Aspekte notwendig, aufmerksam hinzuhören und die (fachlichen) Zusammenhänge zu verstehen. Häufig reicht es aus hinzusehen, um solche Situationen wahrzunehmen und adäquat zu handeln.

Da in der Erhebung explizit um die Wahrnehmung von Schülervorstellungen und Erkenntnisgewinnung gebeten wurde, ist das Ergebnis der Datenanalyse mit einem Classroom Management Fokus – dass die Personenfähigkeit gering ausfällt – nicht verwunderlich. Um all diese möglichen Einflüsse auf das Ergebnis der Analysen verstehen zu können, müsste der Test erneut durchgeführt werden. Dieses Mal allerdings mit je einer Antwortmöglichkeit zu jedem Aspekt. Die Ratewahrscheinlichkeit wäre dann für jede Antwortmöglichkeit gleich groß. Möglicherweise würden dann mehr Items mit dem Partial-Credit-Modell konform sein, das aus drei Kategorien besteht und einen physikspezifischen Attraktor voraussetzt. Inwiefern die durch das Instrument gemessene Fähigkeit der professionellen Unterrichtswahrnehmung mit dem physikdidaktischen Wissen zusammenhängt, wird nach einer Zusammenfassung in den nächsten Abschnitten berichtet.

7.2.4 Zusammenfassung

Nach der Analyse der Daten zur physikspezifischen professionellen Unterrichtswahrnehmung konnte das Ordered Multiple-Choice-Design mit vier Level (4 – physikspezifische Aspekte wahrnehmen, 3 – andere didaktische Aspekte wahrnehmen, 2 – Classroom Management Aspekte wahrnehmen, 1 – kontrafaktische Aspekte wahrnehmen, vgl. Unterunterabschnitt 6.2.3.1 zu den Ergebnissen der Entwicklungsstudie II.2) nicht umgesetzt werden. Das kann zum einen an einer fehlenden Förderung während der universitären

Lehrerbildung liegen⁴⁶ oder zum anderen daran, dass die professionelle Unterrichtswahrnehmung nicht in fachspezifische und weniger fachspezifische Anteile zerfällt, beziehungsweise das PUPiL-Instrument nicht in der Lage ist, diese Unterteilung zu messen. Die Daten wurden einem Partial-Credit-Modell mit drei Kategorien angepasst (0 – kritische Situation nicht wahrnehmen, 1 – einen Distraktor wählen, 2 – den Attraktor wählen) und sind nicht normalverteilt. Die EAP-Reliabilität von 0.698 ist für die Entwicklung eines Instruments, in dem mit der Erhebungssituation versucht wird, nah an die Ausführung der Fähigkeit der professionellen Unterrichtswahrnehmung zu gelangen, akzeptabel. Es erfüllen 19 der 38 geschlossenen Items die Bedingungen, die das Partial-Credit-Modell an die Daten stellt. Interpretationen der Analysen haben ergeben, dass die Testpersonen durch sieben der 19 Items gut differenziert werden. Die anderen 12 Items differenzieren weniger gut, zwei von diesen weisen eine hohe Differenz zwischen den benötigten Personenfähigkeiten für Kategorie 1 und Kategorie 2 auf. Eine Ursache für diese unterschiedliche Differenzierung durch die Items ist die induktive Distraktorenentwicklung, die der unterschiedlichen Komplexität der Situationen zu den Items nicht ausgleicht.

Diejenigen Items, die tendenziell als schwer deklariert wurden, verteilen sich ungefähr gleich auf die Aspekte Erkenntnisgewinnung und Schülervorstellungen.

Items, die aus dem PUPiL-Instrument entfernt wurden⁴⁷, betreffen je Video unterschiedliche Aspekte. Für Video 1 sind es hauptsächlich Items zu Situationen, in denen Erkenntnisgewinnungsprozesse nicht gelingen. Anders als bei Video 1 wurden aus Video 3 sechs Items ausgeschlossen, die gelingende Erkenntnisgewinnungsprozesse beschreiben. Nur eine Situation zu einer physikalisch nicht angebrachten Schülervorstellung wird exkludiert. Bezüglich Video 4 werden sowohl Items mit physikalisch adäquaten als auch inadäquaten Schülervorstellungen ausgeschlossen.

Das PUPiL-Instrument kann durch Umkodierung auch zur Untersuchung einer Classroom Management fokussierten professionellen Unterrichtswahrnehmung dienen. Unter den bisherigen Umständen der Videos, Distraktoren und Aufgabenstellung enthält das Instrument dann 29 Items, die aber auf Grund der Aufgabenstellung mit Fokus auf Erkenntnisgewinnung und Schülervorstellungen überwiegend zu schwer für die Stichprobe sind. Es konnte nachgewiesen werden, dass die physikspezifische professionelle Unterrichtswahrnehmung vom studierten Fach abhängt. Die Fähigkeit der Gruppe Lehramt Physik unterscheidet sich signifikant von der Fähigkeit der Gruppe Lehramt Geographie. Des Weiteren hat der Studiengang Einfluss auf die fach- und situationsspezifische Fähigkeit, solange sich der Anteil der Facetten von Professionswissen in den Ausbildungsgängen unterscheidet (Sonderpädagogik versus Lehramt an Gymnasien).

⁴⁶Das ist im Sinne einer universitären Ausbildung zunächst ausreichend.

⁴⁷Das bedeutet nicht, dass auch die Videosequenz entfernt wurde. Die Testpersonen erhalten lediglich kein geschlossenes Item nach Anhalten des Videos an diesen Stellen.

7.3 Der fachdidaktische Wissenstest

Wie in Abschnitt 2.2 zum PID-Modell, in Unterabschnitt 2.3.1 zur eigenen Konzeptualisierung und in Unterabschnitt 4.4.3 zur Operationalisierung des Konstrukts bereits erläutert, hängt die professionelle Unterrichtswahrnehmung von den Dispositionen der Lehrkraft ab. Zu diesen Dispositionen zählt unter anderem das Professionswissen der Lehrkräfte. Um Erkenntnisse über den Zusammenhang von Professionswissen und professioneller Unterrichtswahrnehmung zu gewinnen, wurde dem PUPiL-Test ein Test zur Ermittlung des physikdidaktischen Wissens⁴⁸ nachgeschaltet. Dieser Test wurde in der Dissertation von Yvonne Gramzow entwickelt und 2015 mit der Dissertationsschrift veröffentlicht (Gramzow, 2015). In diesem Kapitel werden die verwendeten Items mit den entsprechenden Kennwerten vorgestellt. Anschließend wird der untersuchte Zusammenhang der Ergebnisse des physikdidaktischen Wissenstests mit den Ergebnissen des PUPiL-Tests berichtet. Diese Information gibt weitere Hinweise auf die Konstruktvalidität des PUPiL-Instruments.

Der Test zum physikdidaktischen Wissen umfasst die vier empirisch überprüften fachdidaktischen Facetten *Experimente*, *fachdidaktische Konzepte*, *Instruktionsstrategien* und *Schülervorstellungen* (vgl. Gramzow, 2015, S. 110-111)⁴⁹. Wenn Items aus den Facetten für die Betrachtung des Zusammenhangs von physikdidaktischem Wissen und professioneller Unterrichtswahrnehmung verwendet werden sollen, müssen die ganzen Skalen der entsprechenden Facette verwendet werden, um die Gütekriterien des Tests aufrechtzuerhalten.

Die Facette *Experimentieren* spannt einen Teil des physikdidaktischen Wissens auf, der sich von Wissen über Experimente hin zu Wissen über Erkenntnisgewinnung erstreckt (ebd., S. 110). Darunter fallen außerdem Wissen über Schülervorstellungen und Schülerfehler bezüglich „der Planung und Durchführung von Experimenten“ (ebd.), sowie Wissen über einen adäquaten Umgang mit „unerwünschten Experimentierergebnissen im Unterricht“ (ebd.). Diese Facette war für die Betrachtung des Zusammenhangs von physikdidaktischem Wissen und professioneller Unterrichtswahrnehmung nützlich, weil die physikspezifische professionelle Unterrichtswahrnehmung für das PUPiL-Instrument unter anderem auf Aspekte der Erkenntnisgewinnung eingeengt wurde (vgl. Abschnitt 4.4 zur Operationalisierung).

Die *fachdidaktischen Konzepte* umfassen als Facette das Wissen über „theoretische Überlegungen und Hintergründe fachdidaktischer Konzepte“ (ebd.) wie beispielsweise Elementarisierung, didaktische Rekonstruktion und Ansätze zu Schülervorstellungen. Auch diese Facette schien hilfreich zu sein, um Informationen zum Zusammenhang des physikdidaktischen Wissens und der physikspezifischen professionellen Unterrichtswahrnehmung zu gewinnen, weil letztere neben Erkenntnisgewinnung auf das Erkennen von Schülervorstellungen eingeschränkt wurde. Auch wenn es eine Facette mit dem Namen Schülervorstellungen

⁴⁸Zum Professionswissen gehören zwar auch noch das Fachwissen und das allgemeinpädagogische Wissen (Shulman, 1987), diese wurden aber nicht erhoben, weil die Erhebung für die Entwicklung des PUPiL-Instruments zeitlich auf 90 Minuten begrenzt war.

⁴⁹Für die Konzeptualisierung von fachdidaktischem Wissen siehe Gramzow (2015, S. 97-103).

lungen gibt, umfasst die Facette fachdidaktische Konstrukte vor allem Items, die an den Aspekt des Umgangs mit Schülervorstellungen anschlussfähig sind.

Der Test zum physikdidaktischen Wissen beinhaltet bezüglich der Facette *Schülervorstellungen* nur Items, die nach konkreten Schülervorstellungen und ihren Diagnosen fragen (ebd.) und damit nicht notwendigerweise Bestandteil des PUPiL-Tests sind. Daher wurde diese Facette nicht in den Wissenstest zur Bestimmung des Zusammenhangs zur professionellen Unterrichtswahrnehmung mit aufgenommen.

Die letzte Facette *Instruktionsstrategien* umfasst Wissen über Beispiele (zu physikalischen Themen, Gebieten oder Begriffen) und über Verwendung von Darstellungen (ebd.). Es wurden keine Items dieser Facette verwendet, weil sie nicht im direkten Zusammenhang mit den Aspekten im Unterricht stehen, die von angehenden Lehrkräften im PUPiL-Test wahrgenommen werden sollten.

Insgesamt wurden demnach die Skalen zu den Facetten *Experimente* und *fachdidaktische Konzepte* zur Betrachtung des Zusammenhangs von physikdidaktischem Wissen und physikspezifischer professioneller Unterrichtswahrnehmung herangezogen. Fünf der 12 Items betreffen die Facette Experimentieren. Sie umfassen die konkreten Themen Funktionen von Experimenten im Physikunterricht (Item 3⁵⁰), Probleme beim Experimentieren (Item 10), Fehler beim Experimentieren (Item 12), motivierende Elemente von Experimenten (Item 16) und adäquate Messfehlerdiskussion in Experimentierphasen (Item 22) (vgl. für Items und ihre Kodierung Gramzow, 2015, S. 229-276).

Die anderen sieben Items fokussieren auf die Facette *fachdidaktische Konstrukte*. Die Testpersonen müssen ihr Wissen zu Ursachen von Schülervorstellungen (Item 4), Elementarisierung (Item 5), Unterschieden zwischen belastbaren und nicht belastbaren Schülervorstellungen (Item 11), Konzeptwechselstrategien (Item 14), didaktischer Rekonstruktion (Item 19), Planen von Unterricht (Item 23) und dem Unterschied von Schülervorstellungen und Schülerfehlern zeigen (ebd.). Sechs Items werden dichotom kodiert (Item 3, 10, 11, 12, 16 und 22), zwei Items sind im Multiple-Choice-Design (Item 5 und 19) und vier weitere Items sind als mehrteiliges Item design (4, 14, 23 und 24, ebd., S. 276).

Die Skalen des physikdidaktischen Wissenstests wurden den Testpersonen vorgelegt, nachdem sie den Test zur professionellen Unterrichtswahrnehmung bearbeitet hatten, damit sie in ihrer Wahrnehmung nicht bereits durch die behandelten Themen auf gewisse Aspekte in den Unterrichtsvideos gelenkt werden konnten. Die Daten wurden, wie in der Entwicklung des Instruments von Gramzow (2015), mit Hilfe der Item-Response-Theory analysiert. Als Ergebnis liegen Itemschwierigkeitsparameter δ_{FD} und Parameter der Personenfähigkeit θ_{FD} vor. Die Personenfähigkeit θ_{FD} wurde verwendet, um Zusammenhänge mit den Personenfähigkeiten θ_{PI} des PUPiL-Tests zu berechnen.

7.3.1 Ergebnisse der IRT-Skalierung

Der physikdidaktische Wissenstest ist zwar von allen 151 berücksichtigten Testpersonen begonnen worden, einige Items wurden aber nur von 147 Testpersonen beantwortet, weil

⁵⁰Die Itemnummer ist bezogen auf die Aufgabenreihenfolge in Testheft A bei Gramzow (2015, S. 229-276).

die Testzeit abgelaufen war. Wenn die Testpersonen – abgesehen von den 4 Personen, die den Test nicht beenden konnten – ein Item nicht beantwortet hatten, erhielten sie dafür null Punkte. Die IRT-Skalierung der Daten ist in Tabelle 7.10 und in der Wrightmap in Abbildung 7.14 dargestellt. Die EAP-Reliabilität des gesamten Tests (bestehend aus den Skalen Experimente und fachdidaktische Konzepte) beträgt 0.624. Bei der Entwicklung des physikdidaktischen Wissenstests lag die EAP-Reliabilität für die Skala Experimente ursprünglich bei 0.74 und für fachdidaktische Konzepte bei 0.76, wobei die Stichprobenzahl bei $N = 216$ lag.

Der Wrightmap (Abbildung 7.14) ist zu entnehmen, dass die Items für viele Testpersonen einfach zu lösen war, weil die Bereiche der Personenfähigkeit, in denen viele Testpersonen liegen (links), einen großen Teil der Items abdecken. Lediglich die Itemschwierigkeiten des Items fd23 Cat3 liegt über der höchsten Personenfähigkeit der Stichprobe. Dafür sind die Items fd3, fd4 Cat1, fd4 Cat2, fd5_1, fd5_4 und fd5_6 für die Mehrheit der Testperson leicht zu lösen. Einige Testpersonen haben allerdings eine geringe Personenfähigkeit, sodass für diese Testpersonen das Lösen aller Items schwierig ist. Des Weiteren ist auffällig, dass die Items dem Trend nach schwieriger werden, je weiter hinten sie im Test vorkommen, je mehr das Arbeitsgedächtnis der Testpersonen also bereits belastet wurde.

Tabelle 7.10: Itemschwierigkeiten der Skalen zum physikdidaktischen Wissen; Cat 1, Cat 2 und Cat 3 bezeichnen die verschiedenen Kategorien bei den Aufgaben im Multiple-Choice-Design.

Item	fd3	fd4	fd5_1	fd5_2	fd5_3	fd5_4	fd5_5	fd5_6	fd10_1	fd10_2	fd11
Cat 1	-0,640	-2,514	-1,638	-0,369	0,550	-1,452	0,249	-1,165	0,875	0,550	1,013
Cat 2		-1,228									
Cat 3											
Item	fd12	fd14	fd16	fd19_1	fd19_2	fd19_3	fd19_4	fd19_5	fd22	fd23	fd24
Cat 1	0,014	1,632	0,308	-0,250	0,102	0,645	-0,015	-0,250	1,013	1,084	1,837
Cat 2		0,282								0,160	-0,113
Cat 3										1,676	

Die Items, welche als mehrteilige Items designt sind, füllen in Tabelle 7.10 zwei beziehungsweise drei Zeilen. Ähnlich wie bei den Ergebnissen des PUPiL-Tests werden diese Items durch ein Partial-Credit-Modell analysiert. Für die Stichprobe in der Entwicklung der Skalen zum physikdidaktischen Wissenstest erfüllten die Thresholds die Voraussetzung, dass sie aufsteigend sortiert sein müssen (vgl. Gramzow, 2015, S. 17, 138). Wie Tabelle 7.10 zu entnehmen ist, gilt dies aber nicht für die Stichprobe im PUPiL-Test. Bezüglich Item fd14, fd23 und fd24 sind die Thresholds (teilweise) vertauscht.

Da die Reliabilität und Validität des Tests aber nur für alle Items aus den Skalen berechnet wurden und der Test somit nur mit vollständigen Skalen reliabel und valide ist, werden diese Items in der Berechnung der Personenfähigkeit berücksichtigt.

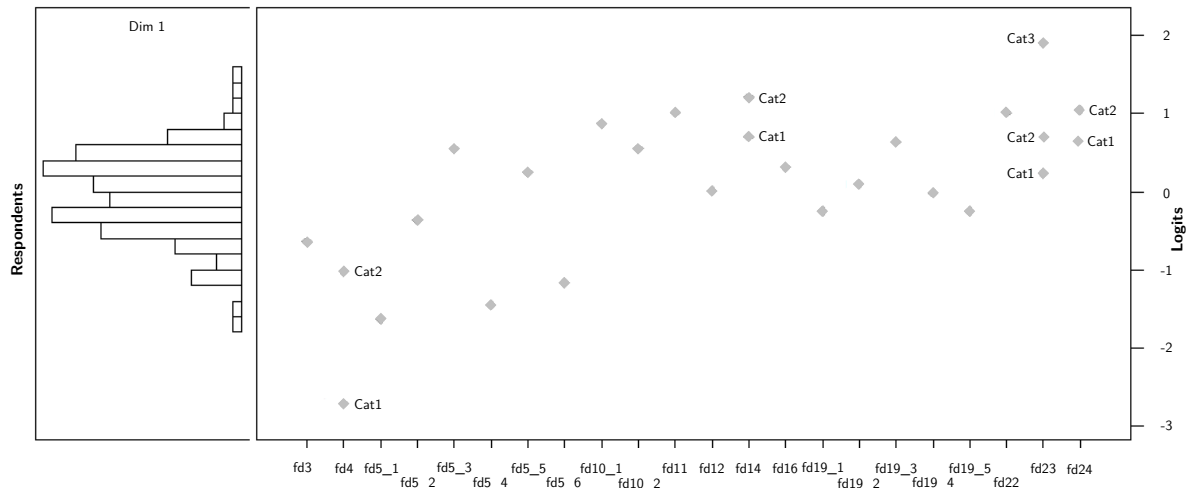


Abbildung 7.14: Wrightmap für die Items zu den Skalen des physikdidaktischen Wissens.

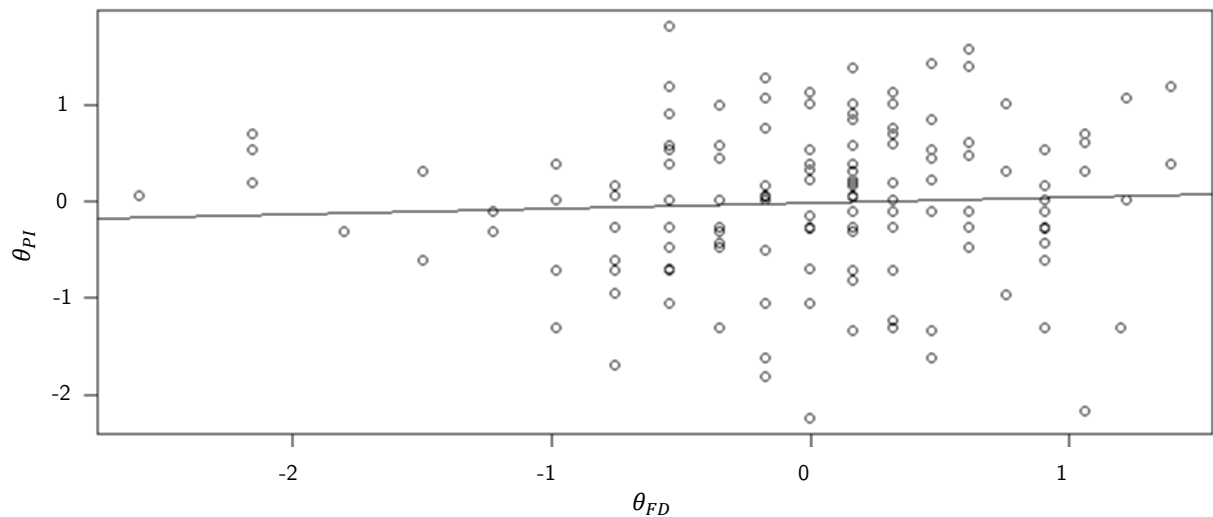


Abbildung 7.15: Streudiagramm zum physikdidaktischen Wissen (Y-Achse) und zur physikspezifischen professionellen Unterrichtswahrnehmung (X-Achse) mit einer Ausgleichsgeraden.

Zusammenhang zwischen dem physikdidaktischen Wissen und der physikspezifischen professionellen Unterrichtswahrnehmung Um den Zusammenhang zwischen dem physikdidaktischen Wissen und der physikspezifischen professionellen Unterrichtswahrnehmung zu untersuchen, wurde eine nichtparametrische Korrelation (Kendalls τ) berechnet. Das Streudiagramm in Abbildung 7.15 zeigt einen positiven Zusammenhang zwischen diesen beiden Fähigkeiten. Der Korrelationstest ergibt eine Teststatistik von $z = 1.5968$ mit $p = 0.110$, womit die Korrelation nur tendenziell signifikant ist. Der Korrelationskoeffizient $\tau = 0.093$ deutet auf einen nur schwachen korrelativen Zusammenhang.

Abhängigkeit des physikdidaktischen Wissens von weiteren Variablen Varianzanalysen mit Hilfe des Kruskal-Wallis-Tests mit den unabhängigen Variablen Fach, Studien-

fortschritt, Studiengang und Hochschule und der abhängigen Variable physikdidaktisches Wissen geben weitere Informationen, um Zusammenhänge der in der Erhebung inkludierten Variablen aufzuklären (siehe Tabelle 7.11).

Tabelle 7.11: Zusammenfassung der Zusammenhänge bezüglich des physikdidaktischen Wissens.

Fachdidaktisches Wissen	Variable		Signifikanz
Korrelationen			
	θ_{PI}	$\tau = 0.093$	0.110
	Decision Making	$\tau = 0.100$	0.090
	Selbsteinschätzung Unididaktik	$\tau = 0.192$	0.001
Kruskal-Wallis			
	Fach	$\chi^2 = 12.843$	0.002
	Physik ~ Geographie	$z = 2.464$	0.041 adj.
	Physik ~ Sachunterricht	$z = 3.301$	0.002 adj.
	Studienfortschritt	$\chi^2 = 8.867$	0.003
	Master ~ Bachelor	$z = 2.978$	0.003 adj.
	Studiengang	$\chi^2 = 5.869$	0.118
	Hochschule	$\chi^2 = 3.675$	0.1592

Bezüglich des studierten Fachs (für das Lehramtsstudium) ergibt sich insgesamt ein hoch signifikanter Medianunterschied ($\chi^2 = 11.57$, $p = 0.003$)⁵¹. Auf Gruppenvergleichsebene bedeutet das einen signifikanten Medianunterschied zwischen den Fächern Geographie und Physik sowie einen hoch signifikanten Medianunterschied zwischen Physik und Sachunterricht. Der Medianunterschied zwischen Geographie und Sachunterricht ist nicht signifikant (siehe Abbildung 7.16⁵²). Da das studierte Fach sowohl auf das physikdidaktische Wissen als auch auf die physikspezifische professionelle Unterrichtswahrnehmung Einfluss nimmt, ließe sich vermuten, dass das Fach möglicherweise ein Moderator im Effekt vom physikdidaktischen Wissen auf die professionelle Unterrichtswahrnehmung ist. Das bedeutet, dass das fachdidaktische Wissen erst dann zu einer besseren professionellen Unterrichtswahrnehmung führt, wenn die Lehramtsstudierenden in dem jeweiligen Fach sozialisiert sind. Die Analysen ergaben allerdings, dass ein Interaktionsmodell, in dem die professionelle Unterrichtswahrnehmung durch das physikdidaktische Wissen – über das Fach moderiert – erklärt wird, nicht signifikant zur Aufklärung beiträgt. Des Weiteren ist der Interaktionseffekt (Fach x θ_{FDW}) nicht statistisch signifikant. Möglicherweise ist aber nicht das

⁵¹In der Entwicklung des physikdidaktischen Wissenstests wurde nur Testpersonen mit dem Unterrichtsfach Physik in die Stichprobe inkludiert, weshalb kein Vergleich zu den hier vorliegenden Daten gezogen werden kann.

⁵²Für die Berechnung der Varianzanalyse wurde nach dem Kruskal-Wallis-Test post-hoc ein Dunn-Test durchgeführt. Die Signifikanz wurde dann anhand des adjustierten p-Wertes bestimmt.

Fach selbst, sondern das Fachwissen ein Moderator des Zusammenhangs zwischen dem physikdidaktischen Wissen und der physikspezifischen professionellen Unterrichtswahrnehmung. Dieses Wissen wurde allerdings nicht erhoben, sodass keine Aussagen darüber abgeleitet werden können.

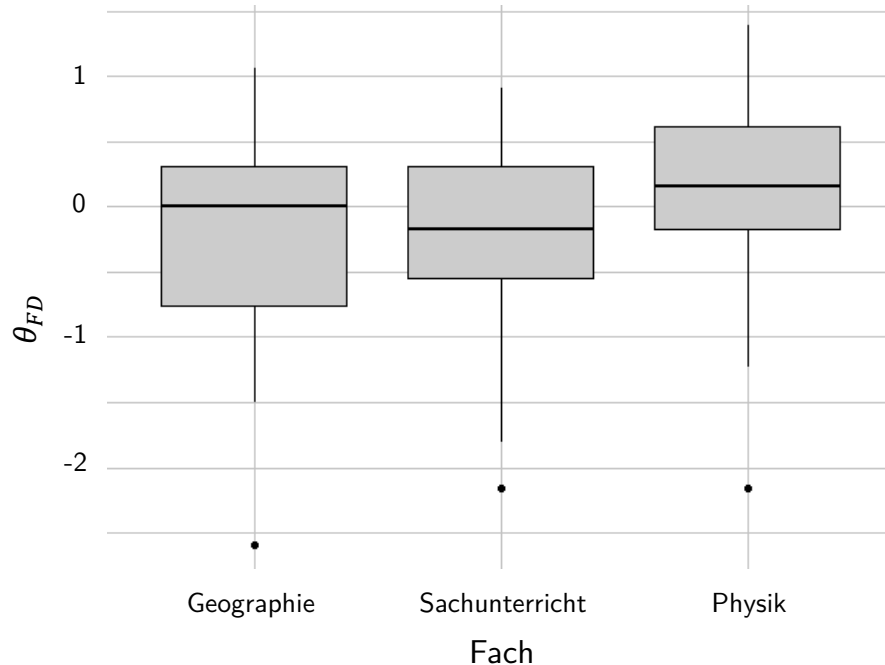


Abbildung 7.16: Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe des Fachs.

Bezüglich des Studienfortschritts (Bachelor oder Master) ergibt sich insgesamt ein signifikanter Medianunterschied ($\chi^2 = 6.02$, $p = 0.014$). Auf Gruppenvergleichsebene bedeutet das einen signifikanten Medianunterschied zwischen den Studierenden im Bachelor und denen im Master (siehe Abbildung 7.17). Studien wie COACTIV (Krauss et al., 2008) und FALKO (vgl. Krauss, Lindl & Schilcher, 2017, S. 229) haben gezeigt, dass sich anscheinend das Professionswissen der Studierenden im Laufe des Studiums und im weiteren Berufsleben entwickelt. Diese Befund kann durch die Daten der PUPiL-Erhebung bestätigt werden. Bezüglich des studierten Studiengangs (für das Lehramtsstudium) ergibt sich insgesamt ein fast signifikanter Medianunterschied ($\chi^2 = 6.75$, $p = 0.08$).

Auf Gruppenvergleichsebene bedeutet das am ehesten einen (signifikanten) Medianunterschied zwischen den Studiengängen Lehramt für Gymnasien und Lehramt für Primarstufe (siehe Abbildung 7.18).

Bezüglich der Hochschule ergibt sich insgesamt kein signifikanter Medianunterschied ($\chi^2 = 4.34$, $p = 0.114$). Dies ist auch in dem entsprechenden Boxplot zu sehen (vergleiche Abbildung 7.19)

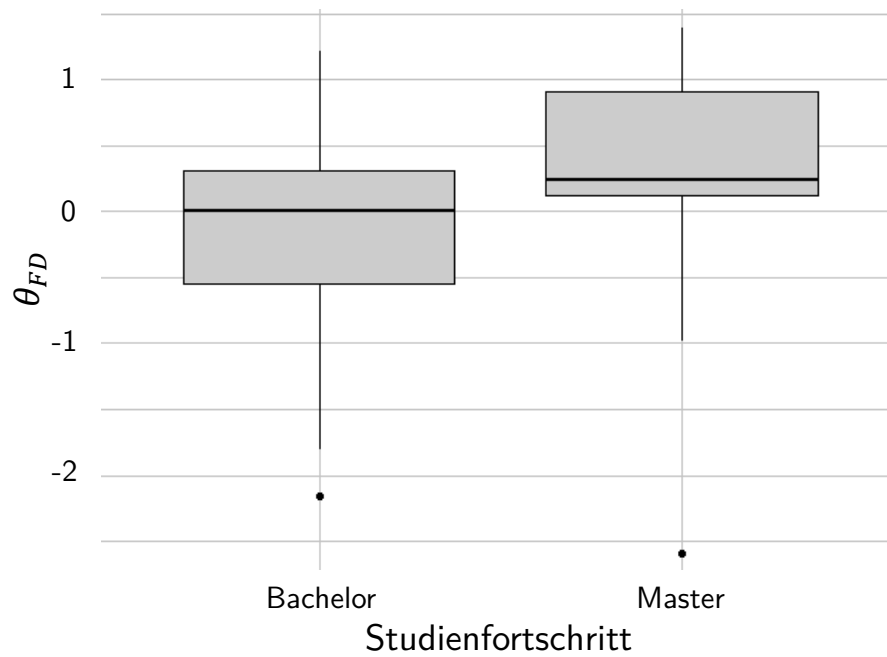


Abbildung 7.17: Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe des Studienfortschritts.

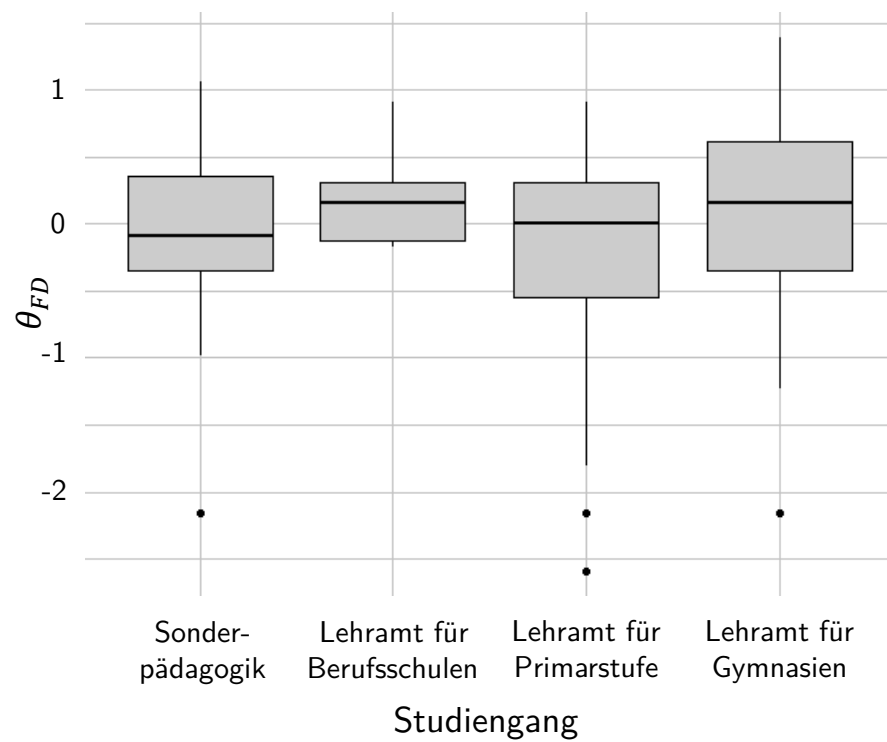


Abbildung 7.18: Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe des Studiengangs.

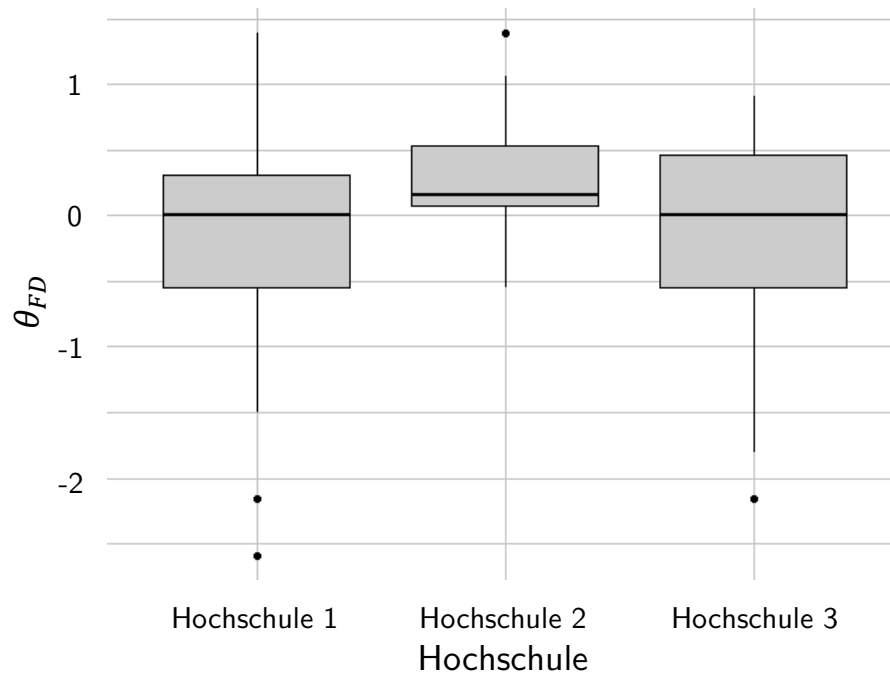


Abbildung 7.19: Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe der Hochschule.

7.3.2 Interpretation und Diskussion

In diesem Unterabschnitt werden die oben dargestellten Ergebnisse interpretiert. Die Analyse der Zusammenhänge des fachdidaktischen Wissens mit der physikspezifischen Unterrichtswahrnehmung und weiteren Variablen unterstützt die Beurteilung der Konstruktvalidität bezüglich möglicher Interpretationen der Ergebnisse des PUPiL-Tests.

Um die Testzeit für die Erhebung rund um den PUPiL-Test in einem vertretbaren Maß zu halten, wurden nur Skalen zu Facetten des physikdidaktischen Wissenstests ausgewählt, die auch im PUPiL-Test relevant waren. In diesem Fall waren das die Facetten Experimentieren und fachdidaktische Konzepte, da diese am besten zu den durch die Videos und die Aufgabenstellung fokussierten Aspekten Erkenntnisgewinnung und Schülervorstellungen passen. Auch die anderen Facetten enthalten Items, die in Hinblick auf den Zusammenhang von physikdidaktischem Wissen und professioneller Unterrichtswahrnehmung aufschlussreich wären. Auf Grund der begrenzt verfügbaren Zeit für die Erhebung wurden dennoch nur zwei passende Skalen ausgewählt. Nicht alle Items in diesen beiden Skalen sind für den PUPiL-Test relevant, weil sie nach Wissen fragen (vgl. dazu Gramzow, 2015, S. 63-91), das in dem Test nicht zur Anwendung kommt. Die folgenden Interpretationen sind daher zwar valide, aber nicht ohne Einschränkung gültig.

Anders als in der IRT-Skalierung von Gramzow (2015, S. 138) werden die Items mit steigender Itemnummer und folglich mit fortschreitender Testzeit vom Trend her immer schwieriger. Die Skalen des physikdidaktischen Wissenstests werden während der Erhe-

bung zum PUPiL-Instrument erst nach circa 70 Minuten Testzeit von den Testpersonen bearbeitet. Möglicherweise sind die Items, die weiter hinten im Test vorkommen, weniger leicht zu lösen, weil die Konzentration der Testpersonen mit steigender Zeit abnimmt. Um diesem Effekt vorzubeugen, hätten die Tests zu den beiden Fähigkeiten an verschiedenen Tagen administriert werden müssen. Dies war aber weder logistisch noch zeitlich umsetzbar, da für die Erhebung Seminarzeit der Studierenden verwendet wurde. Jede neue Testung bringt Vor- und Nachbereitungszeit mit sich, sodass bei zwei getrennten Testungen effektiv mehr als die 90 Minuten benötigt worden wären⁵³.

Beim Vergleich weiterer Kennwerte der Analyse (Reliabilität, Reihenfolge der Thresholds, Personenfähigkeiten) mit den Werten der Validierungsstudie von Gramzow (2015) fällt auf, dass die beiden Skalen des physikdidaktischen Wissenstests nicht sonderlich gut zu der Stichprobe des PUPiL-Tests passen. Das kann an der Stichprobengröße, aber auch an der geringen Varianz in der Stichprobe liegen (es wurden nicht alle Studierende der drei Hochschulen untersucht). Auch die nicht optimale Passung zwischen Test und Stichprobe muss bei folgenden Analysen und Interpretationen beachtet werden.

Der Grund für die Inklusion des physikdidaktischen Wissenstests ist – wie bereits erläutert – die Konstruktvalidierung des PUPiL-Instruments. Die nur tendenziell signifikant schwach positive Korrelation zwischen dem physikdidaktischen Wissen und der physikspezifischen professionellen Unterrichtswahrnehmung der Testpersonen bestätigt (im Rahmen der Gültigkeit aller Aussagen über die Skalen des physikdidaktischen Wissenstests) die theoretische Modellierung von professionellen Lehrerkompetenzen (siehe Abschnitt 2.2 zum PID-Modell). In dieser ist ein Zusammenhang zwischen Dispositionen und den situationsspezifischen Fähigkeiten implementiert. Das PID-Modell impliziert, dass die situationsspezifischen Fähigkeiten wie Erkennen, Interpretieren und das Fassen eines Handlungsplans von Wissen, Erfahrungen, Erwartungen, Orientierungen und vielem mehr abhängen (Blömeke, Gustafsson & Shavelson, 2015). Aus der Wahrnehmungspsychologie kann aber auch gelernt werden, dass die Wahrnehmung ihrerseits die kognitiven und affektiv-motivationalen Aspekte der Dispositionen beeinflussen kann (vgl. beispielsweise Christmann, 1999, S. 6–8). Da die Ergebnisse des PUPiL-Instruments diese theoretischen Annahmen ansatzweise stützen, kann geschlussfolgert werden, dass das Instrument eine situationsspezifische Fähigkeit wie die professionelle Unterrichtswahrnehmung misst. Wie fachspezifisch diese gemessen werden kann, wurde oben in Unterabschnitt 7.2.3 bereits diskutiert. Ein weiteres Indiz für die Validität des Tests ist die *schwache* Korrelation mit dem physikdidaktischen Wissen. Wäre die Korrelation stark, könnte vermutet werden, dass der PUPiL-Test ebenfalls eher fachdidaktisches Wissen als die professionelle Unterrichtswahrnehmung misst. Auf Grund der Beschaffenheit des Tests war zu erwarten, dass der Test nicht fachdidaktisches Wissen misst, weil noch weiteres Wissen (Fachwissen, allgemeindidaktisches Wissen), affektive Aspekte (Erwartungen und Erfahrungen)

⁵³Das PUPiL-Instrument wird zukünftig ohne begleitende Skalen oder Fragen zu demographischen Daten eingesetzt. Diese dienten lediglich der Validierung des Instruments. Dementsprechend nehmen zukünftige Testungen nicht mehr als 90 Minuten in Anspruch.

sowie situative und kontextuelle Fähigkeiten zum Lösen der Items notwendig waren. Die Korrelation mit dem physikdidaktischen Wissen könnte stärker sein, wenn der Test zum physikdidaktischen Wissen besser an das PUPiL-Instrument angepasst gewesen wäre. Im Rahmen des Dissertationsprojekts ist dies jedoch nicht möglich gewesen.

Es hat sich gezeigt, dass sich das physikdidaktische Wissen zwischen den Fächern unterscheidet. Von der Konzeption des Tests ausgehend war zu erwarten, dass die Geographielehramtsstudierenden und die Sachunterrichtsstudierenden sich nicht signifikant unterscheiden, weil ihr physikdidaktisches Wissen ähnlich von dem der Physiklehramtsstudierenden abweicht. Wäre die Physiklehramtsstudierenden nicht besser abschneiden als die anderen Studierenden und jede Person hätte unabhängig vom studierten Fach die gleiche Chance, die Aufgaben des physikdidaktischen Wissenstests zu lösen, wäre dies ein Indiz dafür, dass das Lehramtsstudium nicht ausreichend auf den fachspezifischen Beruf vorbereitet.

Es ist nicht verwunderlich, dass sich das physikdidaktische Wissen in der pseudolängsschnittlichen Stichprobe vom Bachelorstudium zum Masterstudium entwickelt. Die ersten fachdidaktischen Seminare finden im Bachelor statt, deren Vertiefungen liegen gemeinsam mit schulpraktischen Phasen im Master der Lehramtsausbildung. Die Geographielehramtsstudierenden sind in beiden Gruppen ungefähr gleich oft vertreten, sodass deren geringeres physikdidaktisches Wissen auf diesen Befund einen eher geringen Effekt haben dürfte. Alle Sachunterrichtsstudierenden befanden sich zum Zeitpunkt der Erhebung im Bachelorstudium und können entweder der Sonderpädagogik oder dem Lehramt für Primarstufe zugeordnet werden. Die Physiklehramtsstudierenden verteilen sich eher auf das Lehramt für Gymnasien oder Berufsschulen. Die Tendenz bezüglich des physikdidaktischen Wissens, die in Abbildung 7.18 abzulesen ist, wird auch durch den Kruskal-Wallis-Test bestätigt. Dieses Ergebnis war zu erwarten, da in den verschiedenen Studiengängen verschiedene Schwerpunkte gesetzt werden.

Das physikdidaktische Wissen ist über die verschiedenen Hochschulen, die in der Studie berücksichtigt wurden, nicht signifikant verschieden unter den Testpersonen ausgeprägt. Dieser Befund legt nahe, dass die Ausbildung in den drei Standorten auf einem ähnlichen Niveau abläuft. Allerdings ist anzumerken, dass an der Hochschule 3 ausschließlich Sachunterrichtsstudierende während des Bachelors befragt wurden. Dass das physikdidaktische Wissen dieser Gruppe an Studierenden mit dem Wissen der Studierenden an den anderen beiden Hochschulen mithalten kann, ist verwunderlich, weil dort zusätzlich Masterstudierende getestet wurden. Das Wissen bezüglich fachdidaktischer Konzepte und Experimente der Sachunterrichtsstudierenden aus Hochschule 3 ist erstaunlich hoch.

7.3.3 Zusammenfassung

Die Skalen des physikdidaktischen Wissenstests bezüglich der Facetten Experimentieren und fachdidaktische Konzepte passen insgesamt weniger gut zu der Stichprobe der Erhebung zur Erfassung der professionellen Unterrichtswahrnehmung. Das ist vor allem an

der niedrigen EAP-Reliabilität von 0.663 und den vertauschten Schwellen der Partial-Credit-Items festzumachen. Bei der Interpretation weiterer Analysen, in denen die Daten bezüglich des physikdidaktischen Wissenstests eingehen, muss diese Einschränkung daher immer mitgedacht werden.

Es existiert eine tendenziell signifikante aber schwache Korrelation zwischen der physikspezifischen professionellen Unterrichtswahrnehmung und dem physikdidaktischen Wissen ($\tau = 0.093$, $p = 0.110$). Diese schwache Korrelation stützt die Konstruktvalidität des PUPiL-Instruments, da sie den theoretisch vermuteten Zusammenhang tendenziell bestätigt. Die Korrelation ist aber schwach genug, um zu zeigen, dass der Test nicht reines fachdidaktisches Wissen misst.

Des Weiteren hängt das hier gemessene physikdidaktische Wissen hoch signifikant vom studierten Fach, signifikant vom Studienfortschritt (Bachelor oder Master), fast signifikant vom Studiengang und nicht signifikant von der untersuchten Hochschule ab.

7.4 Fassen eines Handlungsplans

Wie bereits in Unterabschnitt 7.1.4 zur Kodierung der Daten aus der Hauptstudie erläutert, wurde die Fähigkeit, für erkannte und interpretierte kritische Situationen einen Handlungsplan zu treffen, von der Messung der physikspezifischen professionellen Unterrichtswahrnehmung abgekoppelt, aber mit dem gleichen Instrument. Grund dafür waren empirische Ergebnisse in der Entwicklungsstudie II, die zeigen, dass es vor allem Studierenden schwer fällt, Handlungspläne zu formulieren, selbst wenn sie die erkannten Situationen adäquat interpretiert haben.

In diesem Abschnitt folgen deskriptive Berichte über das Verhalten der Testpersonen bezogen auf das Fassen eines Handlungsplans. Anschließend wird eine Korrelation des sogenannten Decision Makings mit dem physikdidaktischen Wissen berichtet.

Um die Interpretation der Daten auf Konstruktebene trotz Abkopplung des Decision Makings von den anderen beiden Dimensionen (Perception und Interpretation) möglichst gut validieren zu können, werden außerdem wieder Varianzanalysen zu den unabhängigen Variablen Fach, Studienfortschritt, Studiengang, Hochschule und zum nicht gesehenen Video berechnet.

7.4.1 Ergebnisse

Die Datengrundlage für die Analyse der Fähigkeit, zu einer erkannten und interpretierten Situation einen (lernförderlichen) Handlungsplan zu entwickeln und diesen zu äußern, sind die 151 verbliebenen Testpersonen, für welche entweder 12 kritische Situationen (für die Kombinationen aus Video 1 und 2 sowie Video 1 und 3) oder 14 kritische Situationen (für die Kombination aus Video 2 und 3) berücksichtigt wurden.

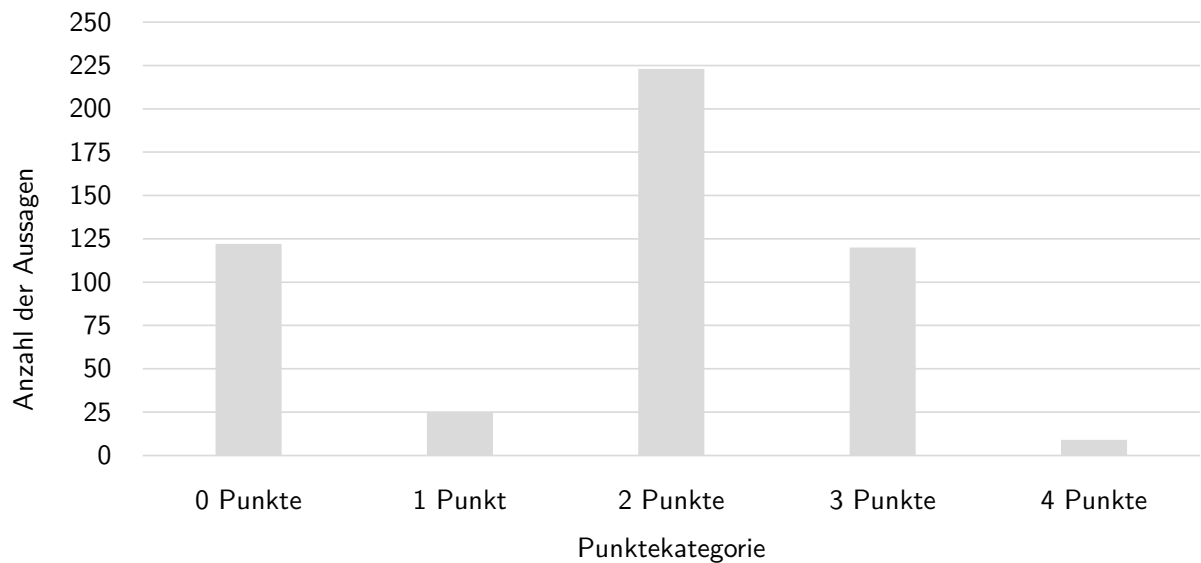


Abbildung 7.20: Verteilung der Antworten zum Decision Making in der Hauptstudie.

Die Antworten der Testpersonen wurden mit Punkten gemäß des in Unterabschnitt 7.1.4 skizzierten Kodierschemas versehen. Insgesamt wurden nur 499 Handlungspläne angegeben, während 1920 für die 151 Testpersonen und 19 Items möglich gewesen wären. Von den geäußerten Handlungsplänen erhielten 122 Pläne keinen Punkt, weil das Textfeld leer hinterlassen beziehungsweise mit einer unverständlichen Zeichenfolge versehen wurde oder die erkannten Situationen nur paraphrasiert wurden (vergleiche für die Verteilung der Punkte auf die Aussagen Abbildung 7.20). Für 25 Handlungspläne erhielten die Testpersonen einen Punkt, da diese Handlungspläne keinen Bezug zur erkannten kritischen Situation hatten. Zwei Punkte konnten an 223 Aussagen zu Handlungsplänen gegeben werden. Diese 223 Handlungspläne wurden mit Bezug zur kritischen Situation erstellt, aber nicht begründet. Mit drei Punkten konnten 120 Antworten bezüglich der Handlungspläne bewertet werden. Die Testpersonen, die für ihren Handlungsplan drei Punkte erhielten, haben ihren Handlungsplan mit Bezug zur erkannten Situation und mit Begründung für ihre geplantes Handeln aufgestellt. Neun Aussagen der Testpersonen über mögliches Handeln konnten mit vier Punkten versehen werden, weil die Testpersonen nicht nur einen passenden, begründeten Handlungsplan aufgestellt, sondern diesen auch von Alternativen abgegrenzt haben.

Insgesamt haben alle Testpersonen zusammen 867 Punkte erhalten. Im Schnitt erhielten die Testpersonen pro Aussage somit 1,7 Punkte. Das entspricht Handlungsplänen, die teils zur kritischen Situation passen, aber nicht begründet wurden. Tatsächlich haben aber sogar 36 Testpersonen keinen einzigen Punkt erhalten. Außerdem hat jede Testperson durchschnittlich 3,3 Handlungspläne über den gesamten Test verteilt gegeben.

Um herauszufinden, ob ein Zusammenhang zum physikdidaktischen Wissen besteht wie zwischen der physikspezifischen professionellen Unterrichtswahrnehmung und dem phy-

sikdidaktischen Wissen, wurde auch dazu eine Korrelationsrechnung durchgeführt. Des Weiteren wurde untersucht, ob es gewissen Gruppen leicht fiel, einen Handlungsplan zu erstellen und anderen nicht. Dazu wurden für die Fähigkeit, einen Handlungsplan aufzustellen Varianzanalysen mit den unabhängigen Variablen Fach, Studienfortschritt, Studiengang, Hochschule und dem nicht gesehenen Video durchgeführt. Für die folgenden Analysen wurde mit relativen Werte für die Fähigkeit, einen Handlungsplan aufzustellen gerechnet. Dieser relative Wert ergibt sich aus dem Quotienten der Punktzahl, die eine Testperson erhalten hat und der Anzahl an maximalen Punkten, die diese optimaler Weise hätte erhalten können.

Die Korrelation des Decision Makings mit dem physikdidaktischen Wissen tendiert nur zur Signifikanz ($p = 0.090$) und wäre außerdem nur schwach mit einem Korrelationskoeffizienten von $\tau = 0.100$ ⁵⁴. Die Varianzanalysen mit dem Kruskal-Wallis-Test haben keine signifikanten Abhängigkeiten von dem Studienfortschritt der Testpersonen, der Hochschule oder des nicht gesehenen Videos ergeben. Dieser Befund stimmt überein mit dem bezüglich der physikspezifischen professionellen Unterrichtswahrnehmung. Ebenfalls ähnlich zu den anderen beiden Teilprozessen der physikspezifischen professionellen Unterrichtswahrnehmung gibt es Hinweise darauf, dass die Teilstichproben bezüglich des Fachs signifikant aus anderen Populationen stammen. Der Test ergab ein $\chi^2 = 7.895$ mit einem p-Wert von 0.019. Ein post-hoc Test mit Adjustierung des p-Wertes präziserte, dass es einen signifikanten Unterschied zwischen den Physiklehramtsstudierenden und den Geographielehramtsstudierenden gibt. Auch bezüglich des belegten Studiengangs wurde das Ergebnis des Kruskal-Wallis-Tests signifikant mit $\chi^2 = 11.072$ und $p = 0.011$. Der post-hoc Test hat hier einen Unterschied zwischen den Studierenden des Lehramts für Gymnasien und der Sonderpädagogik ergeben.

Des Weiteren gibt es einen tendenziell signifikanten Zusammenhang des Decision Makings mit dem selbst eingeschätzten Wissen über die Themen in den Videos auf universitärem Niveau ($\tau = 0.105$, $p = 0.069$).

Die Ergebnisse dieser Berechnungen sind in Tabelle 7.12 zusammengefasst.

7.4.2 Interpretation und Diskussion

In diesem Unterabschnitt werden die oben dargestellten Ergebnisse interpretiert. Die Analyse der Zusammenhänge des Decision Makings mit dem physikdidaktischen Wissen und weiteren Variablen unterstützt die Beurteilung der Konstruktvalidität bezüglich möglicher Interpretationen der Ergebnisse des PUPiL-Tests.

Die deskriptiven Daten zur Fähigkeit, bezüglich erkannter und interpretierter kritischer Situationen einen Handlungsplan zu entwickeln, zeigen, dass nur circa 25% aller Möglichkeiten, einen Handlungsplan aufzustellen, genutzt wurden. Das liegt zum einen daran, dass nicht alle kritischen Situationen von den Testpersonen wahrgenommen wurden.

⁵⁴Zur Erinnerung: Auch für die Berechnung von Zusammenhängen mit der Personenfähigkeit, Handlungspläne aufzustellen, wird ein Signifikanzniveau von $p = 0.1$ als Tendenz zur Signifikanz verstanden.

Tabelle 7.12: Zusammenfassung der Zusammenhänge bezüglich des Decision Makings.

Decision Making	Variable		Signifikanz
Korrelationen			
	θ_{FD}	$\tau = 0.120$	0.040
	Selbsteinschätzung Uniphysik	$\tau = 0.105$	0.069
Kruskal-Wallis			
	Fach	$\chi^2 = 7.895$	0.019
	Physik ~ Geographie	$z = 2.484$	0.039 adj.
	Semester	$\chi^2 = 0.013$	0.909
	Studiengang	$\chi^2 = 11.072$	0.011
	Lehramt für Gymnasien ~ Sonderpädagogik	$z = 3.269$	0.007 adj.
	Hochschule	$\chi^2 = 0.852$	0.653
	Video_x	$\chi^2 = 0.823$	0.663

Zum anderen bestätigt diese Befundlage die Befunde aus Entwicklungsstudie II, auf deren Grundlage entschieden wurde, das Decision Making additiv zu dem Erkennen und Interpretieren zu erheben. Das gewählte Vorgehen wird durch diese Daten legitimiert. Allerdings könnten die Daten anders aussehen, wenn den Testpersonen – ähnlich wie für das Interpretieren – vorgefertigte Distraktoren vorgelegt worden wären. Gegebenenfalls sollte auf Basis dieser Ergebnisse erneut in eher offenen Formaten wie Interviews der Frage nachgegangen werden, wovon die Fähigkeit des Decision Makings beeinflusst wird.

Ebenfalls ein Viertel der Antworten indiziert, dass die Testpersonen nicht in der Lage sind, Handlungspläne zum Unterstützen des Lernens zu generieren. Insgesamt 36 Testpersonen erhalten keinen Punkt für ihre Handlungspläne oder generieren erst gar keine. Ein möglicher Grund für diesen hohen Anteil an nicht fähigen Personen bezüglich des Decision Makings ist, dass das Fassen eines Handlungsplans nicht obligatorisch war. Die Testpersonen konnten ohne einen Beitrag zu geben mit dem Anschauen der Videos fortfahren. Des Weiteren hat die Entwicklungsstudie II gezeigt, wie schwer es den Studierenden fällt, einen Handlungsplan zu generieren. Mit diesem Befund war daher zu rechnen. Die einzigen Unterrichtserfahrungen, die einige der Testpersonen der Stichprobe bisher hatten, sind ihre Erfahrungen aus Lernendenperspektive. Ein größerer Praxisbezug, beziehungsweise mehr und im Studium früher angesetzte Praxisphasen könnten dieser Tatsache entgegenwirken. Allerdings sollte ein größerer Praxisbezug nur als Ergänzung und nicht als Ersatz für die theoretische Ausbildung gedacht werden. Demnach wäre es nicht sinnvoll theoretische Ausbildungsinhalte durch praktische Phasen zu ersetzen. Stattdessen sollten die praktischen Arbeitsphasen immer wieder durch das theoretische Wissen geleitet werden. Dass 25% der Antworten keinen Bezug zur interpretierten kritischen Situation haben,

zeigt unter Umständen eine Unklarheit der Aufgabenstellung an. Die Testpersonen haben in diesen Fällen einen Handlungsplan gefasst, der nicht zum (benannten) Kern der kritischen Situation passt oder sogar eine ganz andere Situation erkannt. Folglich hat die Aufgabenstellung „Stellen Sie sich zu jeder Situation vor, Sie stünden neben den agierenden Lernenden, äußern Sie, wie Sie handeln würden.“ nicht zum gewünschten Handeln geführt. Möglicherweise haben aber auch die Distraktoren, die die Testpersonen zwar gelesen, aber nicht gewählt haben, zu den Handlungsplänen geführt.

Der Großteil der mitgeteilten Handlungspläne weist einen Bezug zur kritischen Situation auf, wird allerdings ohne Begründung generiert, welche durch die Aufgabenstellung aber auch nicht eingefordert wurde. Die Studierenden scheinen dabei ihren Intuitionen zu folgen. Das kann punktuell genau das Richtige sein, in vielen Fällen dürfte eine informierte Entscheidung aber zu einem besseren Lernerfolg der Lernenden führen. In der universitären Lehrerbildung sollte deshalb das Entscheiden mit Begründungen gefördert werden, damit die angehenden Lehrkräfte lernen, alle wesentlichen Merkmale einer Lernsituation in die Entscheidung über ihr anschließendes Handeln mit einzubeziehen.

Immerhin 25% der Handlungspläne beinhalten eine Begründung. Wie die Varianzanalyse mit der unabhängigen Variable „Studienfortschritt“ gezeigt hat, kann dies aber nicht mit dem Grad an Erfahrung beispielsweise durch Praktika festgemacht werden, weil diese Abhängigkeit nicht signifikant ist. Wissen über Evidenz für Begründungen ist demnach eher fach- und studiengangsabhängig. Dieser Befund zeigt wieder eine gewisse Abhängigkeit einer situationsspezifischen Fähigkeit mit (fachspezifischem) Wissen.

Der geringe Anteil an Handlungsplänen, die mit Alternativen verglichen werden, ist nicht überraschend. Für das Wissen über verschiedene mögliche Handlungspläne müssen die Personen zunächst Erfahrungen gesammelt haben, die ihnen zeigen, dass Handlungen, wie sie im Lehrbuch stehen, nicht zu jeder Lernsituation passen. Lernsituationen sind stark an einen Kontext geknüpft und bedürfen daher adaptiven Maßnahmen und auch in Lehrbüchern gibt es nicht zu jeder kritischen Situationen eine bestimmte Lösung.

Festzuhalten ist, dass im Schnitt eher Handlungspläne mit Bezug, aber ohne Begründung notiert wurden. Das ist wenig verwunderlich, weil den Testpersonen insgesamt die schulischen Erfahrungen fehlen, um evidenz- oder wissensbasierte Entscheidungen über potentiell Handeln im Unterricht zu treffen, welches möglicherweise lernförderlicher ist, als rein intuitives Handeln. Nichtsdestotrotz sollten die Hochschulen es als ihre Aufgaben sehen, den Studierenden durch mehr Praxisbezug und frühere Praxisphasen im Studium eine Grundlage für solche situationsspezifischen Fähigkeiten zu geben.

Insgesamt bestätigen diese Daten die Ergebnisse aus der Entwicklungsstudie II.

Bei der Betrachtung der (nicht) signifikanten Varianzanalysen mittels Kruskal-Wallis-Tests sind Parallelen zu den Ergebnissen bezüglich der physikspezifischen professionellen Unterrichtswahrnehmung festzustellen. Das Decision Making wird offenbar ebenfalls nicht signifikant von Studienfortschritt, Hochschule und Video_x beeinflusst. Dahingegen zeigt die Abhängigkeit von dem Fach, dass diese Fähigkeit vom professionellen Wissen der Testpersonen abhängt. Der Unterschied zwischen den Gruppen Physik und Geographie, zwi-

schen denen eine signifikanter Unterschied auftrat, ist zum einen das Fachwissen und zum anderen das fachdidaktische Wissen, welches sie im Laufe ihrer Lehramtsausbildung erlangen können. Ähnliches gilt für den signifikanten Unterschied zwischen den Teilstichproben zum Studiengang: Die Studierenden der Studiengänge Sonderpädagogik und Lehramt für Gymnasien unterscheiden sich im Wesentlichen in dem fachbezogenen professionellen Wissen, welches sie während des Studiums vermittelt bekommen. Die Korrelationen mit dem erhobenen physikdidaktischen Wissen und selbst eingeschätzten Fachwissen zeigen allerdings statistisch signifikante Zusammenhänge. Es scheint aber noch andere Dispositionen zu geben, die mit der Fähigkeit, überhaupt zu Entscheiden und diese Entscheidungen evidenzbasiert zu fällen, zusammenhängen. Bisher wurde häufig untersucht, wie die situationsspezifischen Fähigkeiten mit den kognitiven Dispositionen der Testpersonen zusammenhängen, aber weniger welche Rolle der affektiv motivationale Anteil der Dispositionen in der Ausprägung solcher Fähigkeiten einnimmt (vgl. Kapitel 3 zum Stand der empirischen Forschung). Wie sich aber in aktuellen Berichten auf Tagungen wie beispielsweise der European Science Education Research Association zeigt, gibt es bereits Bemühungen Aspekte wie die Selbstwirksamkeitserwartung in die empirischen Untersuchungen der professionellen Unterrichtswahrnehmung mit einzubeziehen (Honskusová & Rusek, 2011). In Zukunft sollten demnach neben dem Wissen und den Erfahrungen mit Unterricht auch affektiv motivationale Dispositionen der Studierenden gefördert werden. Auf diese Weise würde nicht nur die Fähigkeit, sondern auch die Bereitschaft der Studierenden, Unterricht professionell wahrzunehmen, um Lernprozesse zu fördern, unterstützt (Weinert, 2001a). Es bleibt in Zukunft noch zu untersuchen, ob und welche Zusammenhänge zwischen der professionellen Unterrichtswahrnehmung und dem fachdidaktischen Wissen, aber auch dem Fachwissen und der Fähigkeit, lernförderliche Handlungspläne zu generieren, bestehen.

7.4.3 Zusammenfassung

Die Analyse der Daten zum Decision Making weist darauf hin, dass die Annahme einer dreidimensionalen professionellen Unterrichtswahrnehmung bestätigt werden kann.

Außerdem konnten die Ergebnisse aus der Entwicklungsstudie II bestätigt werden. Die Fähigkeit, einen Handlungsplan zu generieren, der (fiktiven) Lernenden helfen soll, scheint bei Studierenden noch nicht ausgebildet zu sein. Sie fällen Entscheidungen häufiger, ohne Begründungen zur Unterstützung heranzuziehen und vergleichen ihre Wahl auch nicht mit alternativen Handlungen, die in den entsprechenden kritischen Situationen potentiell lernen unterstützen könnten.

Die Untersuchungen der Stichprobenunterschiede geben zusätzlich Anlass dazu, den Zusammenhang zu weiteren kognitiven aber auch affektiv motivationalen Dispositionen zu untersuchen.

8 Abschließende Zusammenfassung und Diskussion

In diesem Kapitel werden die zentralen Ergebnisse dieses Projekts zusammengefasst, diskutiert und in den Forschungsstand eingeordnet. Es werden die Fragen beantwortet, welche Erkenntnisse sich für die Physikdidaktik und die Bildungswissenschaften aus den Ergebnissen ergeben, welche Folgen diese Ergebnisse für die physikdidaktische Forschung sowie die Lehramtsausbildung haben und welche Forschungsdesiderate bezüglich der professionellen Unterrichtswahrnehmung bestehen bleiben.

8.1 Zusammenfassung

Die vorliegende Arbeit begann mit der Forschungsfrage, ob die professionelle Unterrichtswahrnehmung von angehenden Physiklehrkräften mithilfe eines auf Videostimuli basierenden Online-Instruments erfasst werden kann. Hypothesen, die zur Beantwortung dieser Frage untersucht wurden, lauten:

- H1: Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften und von (angehenden) Geographielehrkräften sowie (angehenden) Sachunterrichtslehrkräften unterscheidet sich signifikant voneinander.
- H2: Die professionelle Unterrichtswahrnehmung von (angehenden) Physiklehrkräften verändert sich im Studienverlauf signifikant.

Dazu wurde das Konstrukt der professionellen Unterrichtswahrnehmung zunächst als ein dreidimensionaler, situationsspezifischer und kontextabhängiger Prozess definiert und hinsichtlich der Wahrnehmung physikspezifischer Aspekte des Physikunterrichts begrenzt. Diese Definition ist wichtig, weil das Konstrukt die Grundlage für das Instrument gebildet hat. Ohne klare Definition der Fähigkeit, Unterricht professionell wahrzunehmen, könnten keine validen Aussagen aus der Studie bezüglich dieser Fähigkeit abgeleitet oder Verallgemeinerungen der Ergebnisse vorgenommen werden.

Die erste Dimension ist *Perception* als Richten der Aufmerksamkeit auf lernrelevante Situationen im Unterricht. Das Richten der Aufmerksamkeit kann dabei sowohl ein Top-Down-Prozess als auch ein Bottom-Up-Prozess sein. Dieser Teilprozess wird sowohl durch kognitive als auch durch affektiv-motivationale Dispositionen beeinflusst.

Die zweite Dimension wird durch die *Interpretation* der lernrelevanten Situation definiert. Die Interpretation reicht vom Erfassen der Situationsinhalte über das Erklären und Beur-

teilen hin zum in Beziehung Setzen aller Elemente der Situation. Auch dieser Teilprozess wird von den Dispositionen beeinflusst, hat allerdings auch Potential diese zu verändern. Die dritte Dimension ist das *Decision Making*, mit welcher Handlungen generiert werden, die an die wahrgenommene Situation anschließen können, um Lernen zu unterstützen.

Insgesamt sind die Situationen, in denen die professionelle Unterrichtswahrnehmung relevant wird, nicht planbar und es muss innerhalb kurzer Zeit gehandelt werden. Weil das professionelle Unterrichtswahrnehmen nicht planbar ist, ist die professionelle Unterrichtswahrnehmung auch von den angrenzenden Fähigkeiten der Reflexion und Diagnose abzugrenzen, weil diese per se im Vorwege geplant werden können.

Innerhalb einer Situation laufen die drei Dimensionen der professionellen Unterrichtswahrnehmung hierarchisch ab. Auf eine Unterrichtsstunde bezogen bedingt eine Dimension allerdings die nächste, weil die gewählten Entscheidungen mögliche zukünftige Situationen und somit die Wahrnehmung dieser beeinflussen, sodass sich ein spiralförmiger Prozess ergibt.

Die lernrelevanten, meist mehrdimensionalen Situationen werden für die Operationalisierung als *kritische Situationen* bezeichnet und müssen folgende Charakteristika aufweisen: (a) Sie finden im Physikunterricht statt, (b) sie liefern Hinweise auf das Lernen der Lernenden, (c) sie betreffen Konzepte, Begriffe und Methoden der Physik, (d) sie berühren physikdidaktisch relevante Aspekte (hier Schülervorstellungen und Erkenntnisgewinnung), (e) die Geschehnisse in den Situationen sind für den Physikunterricht und das Physiklernen zentral und (f) die kritischen Situationen erfordern eine Entscheidung der Lehrkraft. Mit dieser Einschränkung des Konstrukts auf eine physikspezifische professionelle Unterrichtswahrnehmung Lehramtsstudierender von Physikunterricht mit speziellem Fokus auf Schülervorstellungen und Erkenntnisgewinnung wurde das Instrument – bestehend aus drei Videos begleitet von teils geschlossenen und offenen Fragen – entwickelt.

Basis der Entwicklung eines validen Instruments waren Drehbücher für die Videos, die als Stimuli für lautes Denken von (angehenden) Lehrkräften über ihre Wahrnehmung gedient haben, wodurch die professionelle Unterrichtswahrnehmung gemessen werden sollte. Diese wurden in der Entwicklungsstudie I mit Orientierung an der Operationalisierung der zu messenden Fähigkeit entwickelt und beschreiben daher Situationen, in denen Schülervorstellungen und Erkenntnisgewinnung lernrelevant werden. Die Drehbücher wurden durch $N = 6$ Personen mit einer physikdidaktischen Expertise inhaltlich validiert. Das bedeutet, dass in den Drehbüchern Situationen enthalten sind, die der professionellen Unterrichtswahrnehmung einer Lehrkraft bedürfen, weil sie die oben genannten Charakteristika erfüllen (a-f).

Zusätzlich sicherten $N = 5$ praktizierende Physiklehrkräfte die curriculare und ökologische Validität der Drehbücher. Hieraus folgt, dass die Interaktionen, Methoden und Sozialformen der beschriebenen Unterrichtsstunden ausreichend realistisch sind und für die Tätigkeit einer Lehrkraft nicht irrelevant.

Es resultierten nach der Auswahl und dem Verfilmen dreier Drehbücher Videos mit 38 intendiert kritischen Situationen, in denen die Aspekte Schülervorstellungen und Erkenntnisgewinnung nicht im gleichen Verhältnis auftreten. Allerdings ist der Schwerpunkt auf

Schülervorstellungen, der durch die Videos erzeugt wird, angemessen, da diese in fast allen Formen von Physikunterricht vorkommen, während Erkenntnisgewinnungsprozesse nicht in allen Stunden relevant werden. Damit ist durch die Videos ein valides Messverfahren ermöglicht, sodass aus den Ergebnissen der Messung gültige Interpretationen abgeleitet werden können.

In der Entwicklungsstudie II wurden mit diesen Videos als Stimuli insgesamt $N = 22$ Interviews mit Anteilen des lauten Denkens mit Studierenden und Lehrenden geführt. Ziel der Interviews war es, herauszuarbeiten, was die Testpersonen wahrnehmen und wie sie die Situationen interpretieren sowie möglicherweise Handlungspläne generieren. Die qualitative Inhaltsanalyse der Interviews ergab, dass es den Testpersonen ohne Übung schwer fällt, das Instrument in gewünschter Weise zu bedienen. Daher wurde ein Training der Testsituation mit Hilfe einer Übung für den Umgang mit den Videos eingebunden. Aus den Aussagen der Testpersonen kann zusätzlich abgeleitet werden, dass es vor allem den Studierenden schwer fällt, Handlungspläne vorzuschlagen, und praktizierende Lehrende ihre Aufmerksamkeit meist auf Aspekte des Classroom Managements richten.

Neben diesen explorativen Erkenntnissen über die physikspezifische professionelle Unterrichtswahrnehmung, diente das verbale Datenmaterial der Teilstudie als Grundlage für die Entwicklung von geschlossenen Fragen für das quantitative Messinstrument. Die Analyse resultierte in 38 geschlossenen Fragen im Ordered Multiple-Choice-Design mit fünf Antwortkategorien. Eine Antwortmöglichkeit beschreibt den Kern der kritischen Situation (erhält 4 Punkte), eine weitere fokussiert andere relevante didaktische Aspekte wie Sprache oder Gender im Physikunterricht (erhält 3 Punkte), zwei Antworten beschreiben relevante Aspekte des Classroom Managements (erhält 2 Punkte) und eine letzte Option handelt von kontrafaktischen Inhalten der kritischen Situation (erhält 1 Punkt). Da diese Antwortmöglichkeiten von dem Interviewmaterial abgeleitet wurden, spiegeln diese mögliche reale Aspekte des Unterrichts wider, die zukünftige Testpersonen wahrnehmen könnten. Unabhängig von der Wahl der Antwortmöglichkeiten durch die Testpersonen, misst das Instrument somit die Wahrnehmung der Testpersonen und nicht nur die Wahrnehmung eines vorher festgelegten Aspekts. Die Konstruktion der Antwortmöglichkeiten trägt daher zur Inhaltsvalidität des Instruments bei. Würde die Stichprobe noch mehr Testpersonen umfassen, könnte diese Aussage mit einer höheren Sicherheit getätigt werden. Nichtsdestotrotz bilden diese geschlossenen Fragen eine Grundlage für valide Interpretationen, weil die Inhalte für die physikspezifische professionelle Unterrichtswahrnehmung nicht irrelevant sind.

Da das Fassen eines Handlungsplans (Decision Making) den Studierenden schwer fiel und offenbar anderer Dispositionen bedarf als die ersten beiden Dimensionen der professionellen Unterrichtswahrnehmung, wurde diese Dimension für das PUPiL-Instrument gesondert in einem offenen Frageformat⁵⁵ direkt nach Beantwortung der geschlossenen Frage zur Interpretation der erkannten kritischen Situation erhoben.

Bezüglich der Konstruktvalidität des Instruments ist anzumerken, dass die Testperso-

⁵⁵Die Testpersonen mussten keine Antwort angeben, um mit dem Test fortzufahren.

nen der Stichprobe häufig in ein Bewerten der in den Videos gezeigten Lehrkraft und ihrer Handlungen verfielen. Dies entspricht dem advokatorischen Ansatz, über welchen ebenfalls Aussagen über die professionelle Unterrichtswahrnehmung der Testpersonen abgeleitet werden könnten. In der Erhebung mit dem PUPiL-Instrument sollte der advokatorische Ansatz aber nicht zur Klärung der professionellen Unterrichtswahrnehmung genutzt werden, weil letzterer die direkte Wahrnehmung von Lernprozessen einschränkt und im eigenen Unterricht nur in Form von Reflexion der eigenen Handlungen vorkommt.

In der Hauptstudie standen zur Analyse der Validität des Instruments drei Videos mit 38 Items, die jeweils aus einer kritischen Situation und einer geschlossenen Frage mit fünf Antwortmöglichkeiten bestehen, sowie ein zusätzliches Übungsvideo bereit. Die $N = 160$ Testpersonen der Hauptstudie haben jeweils zwei der drei Videos zur Bearbeitung vorgelegt bekommen.

Neben der physikspezifischen professionellen Unterrichtswahrnehmung wurden für die Validierung außerdem Daten bezüglich studiertem Fach, belegtem Semester, studiertem Studiengang und physikdidaktischem Wissen, sowie Selbsteinschätzungen zur Lehrerfahrung, zum Wissen über Physik und Didaktik auf universitärem Niveau erhoben, um die Hypothesen H1 und H2 untersuchen zu können. Des Weiteren wurden Daten zu vorgelegten Videos und besuchter Hochschule erhoben, um mögliche Effekte durch die gezeigten Videos und der Hochschule feststellen zu können.

Die Daten des PUPiL-Instruments wurden mit einer IRT-Skalierung analysiert. Gemäß dem Ordered Multiple-Choice-Design wurde ein fünfstufiges Partial-Credit-Modell an die Daten angelegt. Die Ergebnisse zeigen, dass 151 Testpersonen aufgrund ihrer erbrachten Leistung berücksichtigt werden konnten und 19 der 38 Items mit einem dreistufigen Partial-Credit-Modell konform sind (EAP-Reliabilität 0.698), in dem nur die Kategorien 0 – nicht wahrgenommen, 1 – nicht aufgabengerecht wahrgenommen und 2 – aufgabengerecht wahrgenommen berücksichtigt wurden. Durch die Reduktion auf drei Stufen verliert das Instrument an Informationen über die professionelle Unterrichtswahrnehmung, besteht dafür aber aus 19 zum Modell passenden Items.

Die Personenfähigkeiten und Kategorienschwierigkeiten zeigen, dass der Test für die Stichprobe eher schwer ist, wobei aber nicht auszumachen war, dass einer der beiden Aspekte (Schülervorstellungen und Erkenntnisgewinnung) schwieriger ist. Auch eine Verschiebung des Fokus der professionellen Unterrichtswahrnehmung von physikspezifischen Aspekten auf Classroom Management bei der Analyse der erhobenen Daten scheint keinen erheblichen Einfluss auf die Fähigkeit der Testpersonen zu haben. Es scheint demnach nicht nur *einen* ausgeprägten Fokus dieser situationsspezifischen Fähigkeit zu geben. Vielmehr ist dieser Befund ein Indiz dafür, dass die professionelle Unterrichtswahrnehmung breit gefächert ist und somit verschiedene Fokusse gesetzt werden können und durch die Testpersonen auch gelegt wurde. Die Schwierigkeit scheint eher durch verschiedene Themen als durch verschiedene Aspekte oder Fokusse ausdifferenziert zu werden. Fachspezifisch wahrzunehmen ist anscheinend nicht schwieriger oder leichter als Classroom Management fokussiert wahrzunehmen. Möglicherweise sind dies zwei Ausprägungen der selben Fähigkeit.

Diese Interpretation wird durch die Analyse⁵⁶ der Zusammenhänge mit anderen Variablen gestützt. Die physikspezifische professionelle Unterrichtswahrnehmung korreliert schwach mit dem physikdidaktischen Wissen⁵⁷ ($\tau = 0.093, p = 0.110$), mit dem selbst eingeschätzten Wissen auf universitärem Niveau zu den Themen in den Videos ($\tau = 0.153, p = 0.007$) und mit dem Wissen über die didaktischen Themen in den Videos ($\tau = 0.145, p = 0.011$). Das bedeutet, dass diejenigen mit hohem Wissen tendenziell besser wahrnehmen. Besteht weniger Wissen zu den behandelten Themen, zeigt sich das auch in der physikspezifischen professionellen Unterrichtswahrnehmung der Testpersonen. Die Zusammenhänge sind allerdings zu schwach, um von einer konvergenten Validität zu sprechen.

Neben diesen Korrelationen konnten die Analysen des Tests eine signifikante Abhängigkeit vom Fach und Studiengang nachweisen, was die Hypothese H1 bestätigt. Es ergeben sich signifikante Unterschiede der professionellen Unterrichtswahrnehmung bezüglich der Fächer Physik und Geographie und bezüglich der Studiengänge Lehramt für Gymnasien und Berufsschulen. Die vermutete Abhängigkeit vom Fortschritt im Studium, die in Hypothese H2 formuliert wurde, konnte allerdings nicht bestätigt werden. Das bedeutet, dass für das Wahrnehmen relevante Inhalte und Fähigkeiten über den Studienverlauf keine systematische Progression erfahren, Lerngelegenheiten sich eher unsystematisch über das gesamte Studium verteilen und professionelle Unterrichtswahrnehmung während der universitären Lehramtsausbildung nur in geringem Maße gefördert wird oder dass das Instrument eine Progression nicht anzeigt.

Des Weiteren hängt die physikspezifische professionelle Unterrichtswahrnehmung nicht von der Hochschule ab, an der die Testpersonen studiert haben und getestet wurden. Die Auswahl an Videos, die die Testpersonen gesehen haben und auf deren Grundlage demnach die Fähigkeit bewertet wurde, hat ebenfalls keinen signifikanten Einfluss.

Die Dimensionen Perception und Interpretation wurden getrennt vom Decision Making aber mit dem gleichen Instrument erhoben. Es besteht ein schwacher fast signifikanter Zusammenhang zwischen dem Decision Making und dem physikdidaktische Wissen ($\tau = 0.104, p = 0.076$) – ähnlich wie der Zusammenhang der physikspezifischen professionellen Unterrichtswahrnehmung und dem physikdidaktischen Wissen.

Diese Befunde belegen insgesamt mit einigen zu beachtenden Limitationen (siehe nächster Abschnitt), dass das PUPiL-Instrument eine professionelle Unterrichtswahrnehmung valide messen kann.

Die Limitationen der Studie und damit auch der resultierenden Aussagen werden im Folgenden diskutiert.

⁵⁶Für weitere Analysen wurde festgehalten, dass die Daten zur physikspezifischen professionellen Unterrichtswahrnehmung nicht normalverteilt sind.

⁵⁷Das physikdidaktische Wissen wurde anhand zweier Skalen (Experimentieren und fachdidaktische Konzepte) des Wissenstest aus Paderborn (Gramzow, 2015) gemessen.

8.2 Limitationen der Studie

In den Unterabschnitten zu Interpretation und Diskussion der einzelnen Teilstudien wurden bereits Grenzen der durch die Analyse getroffenen Aussagen aufgezeigt⁵⁸. Von diesen Überlegungen ausgehend werden hier weitere Aspekte der Studie beleuchtet, die Grenzen für die Gültigkeit der Interpretation der vorliegenden Ergebnisse und somit für die Nutzung des Instruments darstellen. Damit werden explizit Punkte hervorgehoben, die die Gültigkeit der Interpretationen (Validität) mindern könnten.

Einige limitierende Aspekte dieser Arbeit, die sich vor allem aufgrund von Entscheidungen bezüglich der Drehbuchkonzeption ergeben (vgl. Kapitel 5), beziehen sich auf das finale Instrument an sich. Es können *Limitationen bezüglich der Gültigkeit der Interpretationen* und *bezüglich der möglichen Leistung der Testpersonen* unterschieden werden. Letzteres bedeutet, dass die Testpersonen aufgrund einiger Entscheidungen in der vorliegenden Arbeit nicht die Möglichkeit bekommen haben, ihre Fähigkeit bestmöglichst zu zeigen.

1. Limitationen bezüglich der Gültigkeit der Interpretationen:

- Für das Instrument wurde der physikdidaktische Fokus auf die Aspekte Schülervorstellungen und Erkenntnisgewinnung begrenzt. Dieser Fokus ist künstlich und erlaubt valide Aussagen über die professionelle Unterrichtswahrnehmung nur für genau diese Aspekte im Physikunterricht. Das bedeutet, dass zur Wahrnehmung aller anderen physikspezifischen Aspekte wie Kommunikation, Bewerten und Sprache im Physikunterricht sowohl aus den Ergebnissen der Hauptstudie als auch aus zukünftigen Studien zwar Hypothesen aufgestellt, aber keine validen Schlüsse gezogen werden können. Allerdings konnten die Daten die Frage, ob es überhaupt einen physikspezifischen Anteil der professionellen Unterrichtswahrnehmung gibt, nicht sicher beantworten, was an einem Stichprobeneffekt liegen könnte. Das Instrument kann demnach eingesetzt werden, um generell zu untersuchen, wie die professionelle Unterrichtswahrnehmung von Studierenden ausgeprägt ist und sich gegebenenfalls durch eine Intervention verändert. Verschiedene Fokusse und Aspekte sollte nur mit Bedacht untersucht werden.
- Die validierten Drehbücher für das Instrument behandeln nur drei der vielen verschiedenen Themen im Physikunterricht. Diese Themen sind außerdem nicht gleichermaßen zentral für das Physiklernen in deutschen Schulen. Die Elektrostatik und Energieumwandlung sind explizit durch die nationalen Bildungsstandards abgedeckt, die Brown'sche Molekularbewegung nur als Spezialfall der Wärmelehre (vgl. Ergebnisse zur curricularen Validierung in Kapitel 5 zur Entwicklungsstudie I und Beschluss der Kultusministerkonferenz, 2004). Neben diesen Themen gibt es weitere wichtige Themen im Physikunterricht, wie

⁵⁸Vgl. Abschnitte 5.2, 6.1.2.2, 6.1.3.2, 6.2.2.2, 6.2.3.2, 7.2.3, 7.3.2, 7.4.2.

beispielsweise Elektrodynamik, Magnetismus, Mechanik, Dynamik und viele mehr. Da diese aber nicht Gegenstand des gezeigten Unterrichts sind, können zur professionellen Unterrichtswahrnehmung dieser Themen auch keine gültigen Aussagen abgeleitet werden. Außerdem kann die Frage, ob die professionelle Unterrichtswahrnehmung themenspezifisch ist, nicht valide beantwortet werden, weil nicht alle möglichen Themen dargeboten wurden. Dies geht einher mit einer weiteren Limitation:

- Der gezeigte Unterricht finden in der Mittelstufe in der 8. und 9. Klasse statt. Das Unterrichten anderer Klassenstufen kann sich beispielsweise durch die vorgeschriebenen Themen, aber auch durch andere sich im Laufe des Alters der Lernenden verändernden Merkmale von Unterricht unterscheiden. Somit ist es auch möglich, dass die professionelle Unterrichtswahrnehmung in der 8. Klasse anders ist als die in der 10. Klasse, weil beispielsweise andere Lernhindernisse existieren und andere Voraussetzungen seitens der Lernenden gegeben sind oder die Lehrkraft mehr Erfahrungen in der einen als in der anderen Klassenstufe hat. Eine Lehrkraft, die mehr Erfahrung im Unterrichten einer bestimmten Klassenstufe und somit vermutlich auch mit bestimmten Themen hat, nimmt in anderen Klassenstufen und Unterrichtsthemen anders – wenn nicht sogar schlechter – wahr.

2. Limitationen bezüglich der möglichen Leistung der Testpersonen:

- In den verwendeten gestellten Videos spielen die Lernenden eine Rolle und durchleben keinen eigenen Lernprozess. Dieser Umstand kann dazu geführt haben, dass die Testpersonen die Notwendigkeit einer lernförderlichen Wahrnehmung nicht erkannt haben und ihre Performanz deshalb so ausgefallen ist, wie es in den Erhebungen mit Studierenden zu beobachten war. Insbesondere das Planen anschließender Handlungen kann dadurch erschwert werden, dass keine echten Lernprozesse unterstützt werden würden.

Weitere Limitationen ergeben sich aus der Art der Erhebung der professionellen Unterrichtswahrnehmung in der Entwicklungsstudie II und der Hauptstudie:

1. Limitationen bezüglich der Gültigkeit der Interpretationen:

- Die Stichprobe in der Entwicklungsstudie II ist verhältnismäßig klein ($N = 22$). Die Interpretationen der Analysen des Interviewmaterials können daher nicht ohne Weiteres verallgemeinert werden. Die Distraktoren und somit die Antworten der Testpersonen in der Hauptstudie sind mit Vorsicht zu interpretieren. Des Weiteren studierten die Testpersonen in der Entwicklungsstudie II alle an der Universität Hamburg. Über die professionelle Unterrichtswahrnehmung und insbesondere deren Teilfähigkeit der Interpretation können folglich keine validen Aussagen für andere Teile Deutschlands abgeleitet werden. Da das Instrument in drei verschiedenen Standorten eingesetzt wurde (Hochschu-

le 1, Hochschule 2, Hochschule 3), schwächt dieser Umstand die Validität des Instruments.

- Auch die Stichprobe in der Hauptstudie ($N = 160$) ist für die verwendeten Analysemethoden relativ klein. Die Ergebnisse des gerechneten Partial-Credit-Modells können daher vorläufig nur als Tendenzen der wahren Fähigkeit der professionellen Unterrichtswahrnehmung interpretiert werden. Dass beispielsweise keine verschiedenen Fokusse (fachspezifisch oder generisch) der professionellen Unterrichtswahrnehmung nachgewiesen werden konnten, bedeutet nicht, dass es diese verschiedenen Fokusse und somit gegebenenfalls verschiedene Fähigkeiten nicht gibt. Es können diesbezüglich auf Grundlage der vorliegenden Analyse aber noch keine validen Aussagen getroffen werden.
- Um die professionelle Unterrichtswahrnehmung explorativ zu untersuchen, wurden weitere Variablen wie das fachdidaktische Wissen, das studierte Fach und der gewählte Studiengang erhoben. Die analysierten Zusammenhänge sind nur schwach. Das bedeutet zum einen, dass das PUPiL-Instrument tatsächlich mehr beziehungsweise etwas anderes misst als das physikdidaktische Wissen. Zugleich lassen diese Ergebnisse aber nicht erkennen, wovon die professionelle Unterrichtswahrnehmung *stark* abhängt. Denkbar wären neben den erhobenen kognitiven Dispositionen auch affektiv-motivationale Dispositionen wie Selbstwirksamkeitserwartungen, Erwartungen an und Erfahrungen mit Unterricht, wie es in dem theoretischen Modell von Blömeke, Gustafsson & Shavelson (2015) vorgeschlagen ist. Bei der Beurteilung des Zusammenhangs zwischen der professionellen Unterrichtswahrnehmung und dem physikdidaktischen Wissen muss allerdings berücksichtigt werden, dass die Analyse letzterer eine geringere Passung der Stichprobe mit dem Test indiziert. Möglicherweise würde sich ein stärkerer oder kein signifikanter Zusammenhang ergeben, wenn der physikdidaktische Wissenstest belastbarere Aussagen über das Wissen der Testpersonen liefern würde⁵⁹.
- Die Interpretationen der Ergebnisse des PUPiL-Tests wurden zwar inhaltlich und entlang des Konstrukts validiert, es fehlt aber ein Außenkriterium, anhand dessen weitere Validierungen vorgenommen werden können. Es wurde zwar der Zusammenhang vom physikdidaktischen Wissen und der professionellen Unterrichtswahrnehmung erfasst, allerdings liefern diese Ergebnisse noch keine Aussagen darüber, wie die Testpersonen in einer realen Unterrichtssituation reagieren, welche Wissensbestände sie aktivieren und worauf sie ihre Aufmerksamkeit richten würden. Für eine Beurteilung, ob und inwiefern der PUPiL-Test professionelle Unterrichtswahrnehmung valide misst, müsste der Zusammenhang mit einem solchen Außenkriterium untersucht werden. Welchen Einfluss die professionelle Unterrichtswahrnehmung auf das Lernen und

⁵⁹Die ist keineswegs eine Kritik an dem fachdidaktischen Wissenstest aus Profile-P. Die gewählten Skalen und die verhältnismäßig kleine Stichprobe führen aber zu nur wenig belastbaren Aussagen, beispielsweise aufgrund von geringer EAP-Reliabilität.

den Lernerfolg der Lernenden hat, kann aus dieser Studie nicht abgeleitet werden.

2. Limitationen bezüglich der möglichen Leistung der Testpersonen:

- Die Untersuchung einer physikspezifischen professionellen Unterrichtswahrnehmung erfolgte in jeder Teilstudie der vorliegenden Arbeit in einer Laborsituation. Das bedeutet, dass die Testpersonen nicht in ihrer eigentlichen Tätigkeit untersucht wurden, sondern einer Situation ausgesetzt waren, die eine ähnliche Handlung evozieren sollte. Zwar kommt das gewählte Vorgehen im PUPiL-Instrument näher an die interessierende Tätigkeit der professionellen Unterrichtswahrnehmung der Lehrkräfte, als es bisher anderen Studien gelungen ist (vgl. Kapitel 3 zum Stand der empirischen Forschung), dennoch führen einige Umstände der Laborsituation dazu, dass die Testpersonen möglicherweise anders handeln, als sie es im eigenen Unterricht täten. Dazu gehört unter anderem, dass nur wenige Informationen über das Ziel der Unterrichtsstunde, Voraussetzungen und Besonderheiten über die Klasse und einzelne Lernende bekannt sind. Die Testpersonen haben den Unterricht, in dem sie wahrnehmen sollen, nicht selber geplant. Da die professionelle Unterrichtswahrnehmung von Dispositionen abhängt (Blömeke, Gustafsson & Shavelson, 2015), beeinflussen nicht nur das fachliche und fachdidaktische Wissen die Performanz der Lehrkräfte, sondern auch das Wissen über die Klasse und die Lernenden sowie über das Ziel der Unterrichtsstunde. Daneben haben auch Erwartungen an den Unterricht und Erfahrungen mit dem Lehren des Themas und Unterrichten der Lernenden Einfluss darauf, was von den Lehrkräften bemerkt werden kann. Da es im PUPiL-Instrument nur basale Informationen diesbezüglich gab, weicht die gezeigte professionelle Unterrichtswahrnehmung in dem Test möglicherweise von der Fähigkeit ab, im eigenen Unterricht lernförderlich wahrzunehmen. Im realen Unterricht entwickeln Lehrkräfte durch ihre Erfahrung Erwartungen an das Verhalten der Lernenden in ihren Klassen. Diese Erwartungen steuern dann möglicherweise die professionelle Unterrichtswahrnehmung.
- Die Möglichkeit für die Testpersonen, ihre Fähigkeit zu zeigen, wurde außerdem durch das Format mit geschlossenen Items eingeschränkt. Zum einen wird die Aufmerksamkeit der Testpersonen durch die vorgegebenen Antwortmöglichkeiten post-hoc gelenkt. Zum anderen werden weitere Interpretationen der Situation nicht zugelassen, beziehungsweise durch das Partial-Credit-Modell möglicherweise schlechter bewertet. Die Interpretationen (als Teilprozess der professionellen Unterrichtswahrnehmung) in Form der Distraktoren sind zum Teil auf einer eher beschreibenden Ebene. Nur in wenigen Fällen konnten Distraktoren mehr als nur eine Beschreibung der Situation liefern, ohne dabei zu attraktiv zu werden. Auf welcher Ebene die Interpretationen der Testpersonen in einem realen Unterricht stattfinden – beschreibend, erklärend, bewertend (Seidel & Stürmer, 2014) – kann aus den Daten der vorliegenden Studie nicht abgeleitet werden.

- Die Ergebnisse des PUPiL-Tests weisen darauf hin, dass das Lösen der Items also das physikspezifische professionelle Wahrnehmen im Unterricht für die Testpersonen eher schwierig ist. Ein Grund dafür kann sein, dass es schwerer ist in (gestellten) Videos professionell wahrzunehmen, als im realen Unterricht. Ein anderer Grund könnte der hohe kognitive Aufwand sein, den die Testpersonen leisten mussten. Neben dem Test zur professionellen Unterrichtswahrnehmung mussten die Testpersonen im Vorwege Fragen zu demographischen Daten beantworten. Möglicherweise war die wahrgenommene Aufgabe für die Studierenden demnach – anders als im Unterricht – die Fragen der Erhebung zu beantworten und weniger lernförderlich wahrzunehmen. Sie konnten sich gegebenenfalls nicht vollständig auf das professionelle Wahrnehmen in den Videos einstellen. Im Nachgang wurden sie außerdem bezüglich ihres professionellen Wissens befragt. Auch das kann ein Grund für die geringe Passung der Stichprobe mit dem physikdidaktischen Wissenstest sein.

8.2.1 Folgerungen für die Validität der Interpretationen

Einige der oben aufgeführten Limitationen schränken den Bereich ein, in dem valide Aussagen über die professionelle Unterrichtswahrnehmung abgeleitet werden können. Dazu gehört, dass das Instrument Fokusse bezüglich des Unterrichtsaspekts (Schülvorstellungen und Erkenntnisgewinnung versus Classroom Management), des behandelten Themas (Brown'sche Molekularbewegung, Elektrostatik und Energieumwandlung) und der Klassenstufe (8. und 9. Klasse) vorgibt. Zudem gelten die Aussagen zunächst auch nur in Bezug auf gestellte Videos.

Egal welche Fokusse gesetzt wurden, führen andere Limitationen zur Verminderung der Validität der präsentierten Interpretationen: Zum einen sind in allen empirischen Teilen die Stichproben relativ klein gewesen, sodass eine Verallgemeinerung der Aussagen nur bedingt möglich ist. Zum anderen wurden die Testpersonen in eine Laborsituation gebracht, die zwar möglichst vergleichbar mit der eigentlich zu messenden Fähigkeit war, aber die trotzdem zu anderen Handlungen als die realen Handlungen führen muss. Zusätzlich wurden zugunsten einer ökonomischen Auswertung geschlossene Items eingesetzt, in welchen die Testpersonen ihre Interpretation erkannter Situationen markieren sollten. Offene Antworten hätten gegebenenfalls zu anderen Interpretationen geführt. Schließlich sind die Daten der Hauptstudie in einem kognitiv aufwendigen Verfahren erhoben wurden, in dem die Testpersonen möglicherweise anders agiert haben, als sie es in einer anders oder weniger kognitiv aufwendigen Situation getan hätten. Aussagen zu einer externen Validität können an dieser Stelle nicht abgeleitet werden, weil kein Außenkriterium zur Validierung untersucht wurde.

8.3 Diskussion

Im vorherigen Kapitel wurde zusammengefasst, welche Erkenntnisse durch die vorliegende Arbeit bezüglich der professionellen Unterrichtswahrnehmung und der Entwicklung eines Instruments zum Messen dieser Fähigkeit gewonnen werden konnten. Hier wird berichtet, was immer noch unsicher oder unbekannt bleibt, und wie diese Forschungsdesiderate möglicherweise angegangen werden können.

Meiner Einschätzung nach gibt es zwei Fragen, die durch die vorliegenden Studien nicht (im vollen Umfang) beantwortet wurden, deren Beantwortung allerdings notwendig ist, um Unterrichtsqualität durch die professionelle Unterrichtswahrnehmung von Lehrkräften zu verbessern:

1. Womit hängt die professionelle Unterrichtswahrnehmung (stark) zusammen?
2. Wie hängen die professionelle Unterrichtswahrnehmung und die realen Lernprozesse und -erfolge der Lernenden zusammen?

Im folgenden Abschnitt werden Implikationen für die (physik-)didaktische Forschung, die aus der vorliegenden Studie resultieren, vorgestellt und im Zuge dessen Vorschläge bereitet, wie die oben stehenden Fragen beantwortet werden können.

Abschließend werden Folgerungen für die Aus- und Fortbildung von Physiklehrkräften aus den Ergebnissen der Studie abgeleitet.

8.3.1 Implikationen für (physik-)didaktische Forschung

Teile der Ergebnisse dieser Studie unterstützen einige der empirischen Befunde, die in Kapitel 3 zum Stand der empirischen Forschung vorgestellt wurden:

- *Studien zum Experten-Novizen-Vergleich:* Es wurden nur in der Entwicklungsstudie II insgesamt neun Lehrkräfte im Vorbereitungsdienst oder mit weiterer Unterrichtserfahrung befragt. Bei der Analyse der Interviews hat sich aber gezeigt, dass die praktizierenden Lehrkräfte auf Aktivitäten der Lernenden fokussieren und häufiger Handlungspläne nennen. Diese Befunde unterstützen die Ergebnisse der vorgestellten Studien zum Experten-Novizen-Vergleich (Carter et al., 1988; Sabers et al., 1991; Seidel & Prenzel, 2007; Oser et al., 2010). Ob die Interpretationen von angehenden Lehrkräften wie bei Carter et al. (1988), Seidel & Prenzel (2007), Oser et al. (2010) auch eher oberflächlich sind, kann aus der vorliegenden Arbeit nicht abgeleitet werden, da die Interpretationsebene vorgegeben war.
- *Studien zur Entwicklung:* Da das Ziel der vorliegenden Arbeit war, ein valides Instrument zu entwickeln, wurde zunächst keine Intervention zur Förderung der professionellen Unterrichtswahrnehmung durchgeführt. Daher kann auch keine Aussage über die mögliche Veränderung der professionellen Unterrichtswahrnehmung auf Grundlage von gezielten Interventionen berichtet werden.

Versteht man das Studium als eine Art Intervention, kann nicht bestätigt werden,

dass sich die Intervention als förderlich für die professionelle Unterrichtswahrnehmung erweist. Allerdings liefern die Ergebnisse der vorliegenden Studie Hinweise darauf, dass die Fähigkeit, Unterricht professionell wahrzunehmen, themenspezifisch sein könnte.

Bezüglich der Förderung und Entwicklung der professionellen Unterrichtswahrnehmung von Lehramtsstudierenden gaben Kramer et al. (2017) zu bedenken, dass komplexe Unterrichtsvideos möglicherweise überfordernd und daher nicht lernförderlich auf die Studierenden wirken könnten. Die Ergebnisse der Erhebung mit dem PUPiL-Instrument machen deutlich, dass für eine Erhebungssituation – möglicherweise anders als für eine Fördersituation – die Verminderung der Komplexität von Unterricht (beispielsweise in Form von parallel ablaufenden Lernprozessen und Störungen) für die Videos zu einem weniger authentischen Umgang führen könnte. Je nach Ziel einer Studie sind demnach gegebenenfalls verschiedene Level an Komplexität angezeigt.

- *Studien zum Zusammenhang zu anderen Variablen:* Die Befunde der vorliegenden Studie unterstützen die Ergebnisse von Stürmer (2011) zum Zusammenhang mit der Anzahl belegter Seminare und von Stürmer et al. (2015) zur Anzahl der Lerngelegenheiten durch die nachgewiesene Abhängigkeit der physikspezifischen professionellen Unterrichtswahrnehmung vom Fach und vom Studiengang. In den verschiedenen Fächern und Studiengängen werden verschieden viele Seminare mit (physik-)didaktischem Fokus angeboten, sodass der gezeigte Medianunterschied möglicherweise auf die Lerngelegenheiten zurückgeführt werden kann.

Im Rahmen der PUPiL-Erhebung wurden keine affektiv-motivationalen Dispositionen erfasst, sodass über den Zusammenhang zur professionellen Unterrichtswahrnehmung nichts gefolgert werden kann. Allerdings weisen die Befunde der vorliegenden Studie auch darauf hin, dass die erhobenen kognitiven Dispositionen nur einen geringen Teil der professionellen Unterrichtswahrnehmung erklären können. Dies liefert weitere Hinweise darauf, dass wie bei Meschede et al. (2017) geschehen, affektiv-motivationale Dispositionen zusätzlich erhoben werden müssten, um zu untersuchen wie ihr Einfluss auf die situative Fähigkeit der professionellen Unterrichtswahrnehmung ist.

Das Ergebnis, dass die physikspezifische und Classroom Management fokussierte professionelle Unterrichtswahrnehmung mittel miteinander korrelieren unterstützt die Ergebnisse von Kaiser et al. (2017) zum Zusammenhang der sogenannten mathematikspezifischen und pädagogischen professionellen Unterrichtswahrnehmung. Auch den schwachen Zusammenhang vom fachdidaktischen Wissen und der professionellen Unterrichtswahrnehmung bei Kaiser et al. (2017) können die Ergebnisse des PUPiL-Tests bestätigen. Offenbar wird das fachdidaktische Wissen noch nicht immer hinsichtlich der professionellen Unterrichtswahrnehmung aktiviert, sodass eine Förderung des fachdidaktischen Wissens (entweder bereits in der Universität oder erst im Vorbereitungsdienst) die Lehramtsstudierenden darin unterstützen müsste, ihr Wissen handlungsgeleitet anzuwenden.

Die Befunde zu nicht signifikanten Zusammenhängen und Kausalitäten bezüglich

der Erfahrung in Praktika und des Fachsemesters unterstützen die (unerwarteten) Ergebnisse von Stürmer (2011) und Gold et al. (2013), die ebenfalls nachweisen konnten. Stürmer (2011) stellte fest, dass es keinen signifikanten Zusammenhang zwischen den Vorerfahrungen durch Praktika und der professionellen Unterrichtswahrnehmung gibt. Gold et al. (2013) wiederum konnten keinen signifikanten Zusammenhang zwischen dem Fachsemester und der professionellen Unterrichtswahrnehmung feststellen.

Inwiefern die einzelnen Dimensionen der professionellen Unterrichtswahrnehmung ein eindimensionales Konstrukt aufspannen, wie es bei Seidel & Stürmer (2014) vorliegt, kann aus den Analysen nicht geschlussfolgert werden. Allerdings weisen die Ergebnisse der Entwicklungsstudie II und der Hauptstudie darauf hin, dass die Fähigkeit Decision Making in der praktischen Umsetzung der professionellen Unterrichtswahrnehmung getrennt von Perception und Interpretation betrachtet werden sollte.

Steffensky et al. (2015) ziehen statt einer Domänenspezifität sogar eine Aspektspezifität der professionellen Unterrichtswahrnehmung in Betracht. Mit Hilfe der vorliegenden Studie kann dies nicht sicher bestätigt werden. Vielmehr zeigte sich, dass die Themen der Aspekte (beispielsweise Schülervorstellungen (Aspekt) zur Energieumwandlung (Thema)) mehr Einfluss auf das Gelingen der professionellen Unterrichtswahrnehmung haben als die Aspekte selbst. Um diese Unsicherheit zu erhellen, wäre ein Design denkbar, in dem zu einem Thema gleich viele kritische Situationen zu mehreren Aspekten (wie Schülervorstellungen und Erkenntnisgewinnung) angeboten und gleichzeitig zu den Aspekten verschiedene Themen behandelt werden – sofern dies möglich ist. Vor allem die generischen Unterrichtsaspekte wie Classroom Management werden weniger vom fachlichen Thema der Unterrichtsstunde beeinflusst als Aspekte wie Schülervorstellungen.

- *Studien zur Instrumententwicklung:* In Kapitel 3 zum Stand der empirischen Forschung wurden Studien vorgestellt, in denen Messinstrumente zur Erfassung der professionellen Unterrichtswahrnehmung bezüglich der Mathematik und naturwissenschaftlichen Fächern entwickelt wurden. Dort wurde ebenfalls abgeleitet, welche Schwächen die Instrumente in Bezug zur Validität aufweisen. Auf dieser Grundlage wurde das PUPiL-Instrument entwickelt. Dieses erweitert die bestehenden Instrumente um eine Studie, in der die professionellen Unterrichtswahrnehmung von Lehramtsstudierenden in zwei aufeinanderfolgenden Prozessen des Erkennens samt Interpretierens und des Handlungsplan Fassens mit großer Handlungsnähe erhoben wird. Die fokussierten Unterrichtsaspekte sind Schülervorstellungen und Erkenntnisgewinnung der Lernenden. Die Testpersonen sehen gestellte Unterrichtsvideos und demonstrieren das Erkennen und Interpretieren von kritischen Situationen durch Stoppen der Videos, wenn sie etwas Lernrelevantes in Bezug auf Schülervorstellungen und Erkenntnisgewinnung wahrnehmen, und das sofort anschließende Beantworten einer geschlossenen Frage zu ihrer Wahrnehmung im Ordered Multiple-Choice-Design. Erst im Anschluss ist ihnen die Möglichkeit gegeben, einen Hand-

lungsplan für die erkannte und interpretierte Situation schriftlich zu notieren. Das PUPiL-Instrument ist somit ein Instrumente, dessen Erhebung auch auf gestellten Unterrichtsvideos basiert, die Erhebungssituation ist aber handlungsnäher, weil die Testpersonen ihre Aufmerksamkeit selber steuern, aus verschiedenen Fokussen wählen können und anschließend eigenständig einen Handlungsplan entwickeln dürfen. Anders als bei den meisten anderen (vorgestellten) Instrumenten steht im Fokus der Erhebung die professionelle Unterrichtswahrnehmung von fachdidaktischen Aspekten (Schülervorstellungen und Erkenntnisgewinnung) und nicht nur generische Aspekte (wie Lernbegleitung oder Lernatmosphäre). Im Gegensatz zu den bisherigen Instrumenten baut das PUPiL-Instrument außerdem auf nur drei Videos, deren Länge aber zwischen fünf und acht Minuten liegt. Damit sind die Videos länger, geben aber einen tieferen Einblick in den Kontext der Unterrichtsstunde, als es ein dreiminütiger Clip vermag. Wie in Abschnitt 4.4 zur Operationalisierung erläutert, macht gerade die Kontextualisierung die professionelle Unterrichtswahrnehmung zu einer anders gearteten Fähigkeit als beispielsweise das Planen oder Reflektieren von Unterricht.

Für die Unterstützung der anderen in Kapitel 3 vorgestellten Befunde (wie beispielsweise zu der Entwicklung der professionellen Unterrichtswahrnehmung von Lehrkräften oder zu Zusammenhängen mit Interesse (Stürmer, 2011) und Beliefs (Meschede et al., 2017)) liefern die Daten des PUPiL-Instruments keine oder keine eindeutigen Hinweise.

Zu Beginn dieses Kapitels wurden Limitationen der Studie vorgestellt. Zukünftige fachdidaktische Forschung kann dabei helfen, einige dieser Limitationen zu eliminieren. Was müsste also unternommen werden, um die Befunde dieser Studie weiter verallgemeinern zu können und somit die professionelle Unterrichtswahrnehmung noch valider zu messen, beziehungsweise zu validieren Aussagen zu kommen?

Sowohl in der Entwicklungsstudie II als auch in der Hauptstudie wurden die Testpersonen einer Laborsituation ausgesetzt, an der nach jetzigem Erkenntnisstand einige Änderungen vorgenommen werden könnten, um das bestehende Instrument so zu verändern, dass sich die Erhebungssituation der Situation im realen Unterricht weiter nähert:

- Um genauere Daten bezüglich des Erkennens und des Interpretierens zu erlangen, könnte der Blickverlauf der Testpersonen beispielsweise durch die Methode des Eye-Trackings aufgezeichnet werden, ähnlich wie es Stürmer et al. (2017) mit mobiler Eye-Tracking Technologie in Mikroteaching-Kontextes bereits untersucht haben. Anschließend könnten sich die Testpersonen ihren Blickverlauf selber ansehen und in einem post-hoc Interview zusätzlich zum bisherigen Verfahren ihren Blickverlauf und somit die Interpretationen der Situationen erläutern und reflektieren. Dieses Verfahren verlangt allerdings ein hohes Bewusstsein der Testpersonen über die Lenkung ihrer Aufmerksamkeit.
- Die Distraktoren wurden auf Grundlage der Antworten von Testpersonen entwickelt, die laut zu den kritischen Situationen in den Videos gedacht haben. Daher sind einige Interpretationen der Situationen in den Distraktoren tiefer interpretiert und

einige eher oberflächlich dargestellt. In einem Gespräch mit der wissenschaftlichen Mitarbeiterin aus der Didaktik der deutschen Sprache und Literatur Yannah Wiechmann entstand die Idee, die Distraktoren zu den Interpretationen auf zwei Ebenen zu formulieren: (1) Halten die Testpersonen das Video an einer kritischen Situation an, erhalten sie erst eine Auswahl an Distraktoren, die auf einer eher beschreibenden Ebene formuliert sind. (2) Anschließend könnte ihnen zum ausgewählten Distraktor eine weitere Auswahl an tiefer interpretierenden Aussagen vorgelegt werden, aus denen sie die für sie am besten passende Aussage wählen müssen. Auf diese Weise kann neben dem Fokus zusätzlich die Tiefe der Interpretationen erhoben werden.

- Ein weiteres Problem bezüglich der Validität des Instruments ist, dass die Testpersonen weder die Lernenden noch die genaue Planung des Unterrichts kennen. In einem zukünftigen Verfahren wäre denkbar, dass die Testpersonen die Planung und deren Begründung im Vorfeld erhalten und sich dadurch über die Intention des gestellten Unterrichts bewusst werden und somit ihre Aufmerksamkeit entsprechend lenken können. Die Testpersonen könnten auch selber Unterricht zum Thema der gestellten Videos planen und sich somit mögliche Lernschwierigkeiten bewusst machen. Dieses Vorgehen scheint zwar aufwändiger, ist aber sinnvoll, um dem starken Kontextbezug der professionellen Unterrichtswahrnehmung gerecht zu werden. Die Aufmerksamkeit der Testpersonen wäre dann vermutlich schon durch die Voraufgaben auf fachliches Lernen gerichtet.
- Die Validität des Instruments wird auch durch die Anzahl der kritischen Situationen und Aufteilung dieser auf Themen und Aspekte verringert. Denkbar wäre daher, mehr Videos zu weiteren Themen zu entwickeln. In einer weiteren Entwicklungsstudie könnte dann aus einer größeren Auswahl von kritischen Situationen gewählt und somit eine repräsentative Auswahl zusammengestellt werden.
- Eine größere Veränderung könnte darin bestehen, dass Videos, die realen Unterricht zeigen, die Videos des PUPiL-Instruments ersetzen. Entsprechend müssten die Distraktoren angepasst und die Validität neu bestimmt werden. Um passende kritische Situationen im realen Unterricht zu finden, ist möglicherweise das Beobachten und Filmen vieler Unterrichtsstunden notwendig. Dieser Aufwand wäre aber gerechtfertigt, wenn dadurch die Validität des Instruments steigen würde. Dies wäre im Fall der curricularen und ökologischen Validität gegeben. Die Inhaltsvalidität und Konstruktvalidität müssten allerdings neu beurteilt werden.
- Um die Konstruktvalidität einer Erhebung mit dem PUPiL-Instrument genauer bestimmen zu können, müsste ein besserer Schätzer für das physikdidaktische Wissen eingesetzt werden. Die berechneten Zusammenhänge basieren auf einer Messung des physikdidaktischen Wissens, welche keine guten Kennwerte hat. Es könnte zukünftig entweder der ganze Test von Gramzow (2015) verwendet werden oder es müsste ein zu den Themen im PUPiL-Instrument passender Test neu entwickelt werden.

Neben den Veränderungen am PUPiL-Instrument lassen sich aus den Ergebnissen der vorliegenden Studie weitere Implikationen ableiten, was in der physikdidaktischen Forschung

unternommen werden müsste, um die professionelle Unterrichtswahrnehmung und ihren Effekt auf Lernen besser verstehen und erklären zu können:

- Sowohl die Entwicklungsstudie II als auch die Hauptstudie haben gezeigt, dass immer noch recht wenig über den Zusammenhang der ersten beiden Dimensionen der professionellen Unterrichtswahrnehmung (Perception, Interpretation) mit der dritten Dimension (Decision Making) bekannt ist. Es stellt sich die Frage, was die Handlung und somit den Handlungsplan einer Lehrkraft „antreibt“. In der Studie von Wider et al. (2019) wurde beispielsweise erhoben, welche Erwartungen die Lehrenden an die Unterrichtsstunde haben, die sie für die Erhebung der professionellen Unterrichtswahrnehmung sehen werden. Ein ähnliches Vorgehen bezüglich aller möglichen Antriebsfaktoren (beispielsweise Wissen, Erwartungen, Haltungen, Einstellungen oder Erfahrungen) für das Handeln im Unterricht könnte Aufschluss darüber geben, mit welcher Zielsetzung Lehrkräfte (vermutlich eher unbewusst) ihre Aufmerksamkeit lenken.
- Es hat sich aber auch gezeigt, dass allein das fachdidaktische Wissen und Aspekte wie Fach und Studiengang einen nur geringen Zusammenhang mit der professionellen Unterrichtswahrnehmung haben. Um zum einen das Kompetenzmodell von Blömeke, Gustafsson & Shavelson (2015) und zum anderen das PUPiL-Instrument zu validieren, sollte der Zusammenhang sowohl der kognitiven als auch der affektiv-motivationalen Aspekte der Dispositionen (wie bei Meschede et al. (2017), Honskusová & Rusek (2011) und Wider et al. (2019) bereits angedacht beziehungsweise geschehen) mit der professionellen Unterrichtswahrnehmung untersucht werden.
- Wenn die Größen bekannt sind, die mit der professionellen Unterrichtswahrnehmung zusammenhängen, könnte in einer Interventionsstudie mit einem Prä-Post-Follow-Up-Design ebenfalls untersucht werden, wie die jeweilige Intervention auf die professionellen Unterrichtswahrnehmung wirkt. Eine denkbare Stichprobe könnte aus Lehramtsstudierenden und praktizierenden Lehrkräften bestehen. Auf diese Weise würde ein weiterer Pseudo-Längsschnitt durchgeführt werden. Für die praktizierenden Lehrkräfte wäre des Weiteren denkbar, dass die Lernprozesse beziehungsweise Lernerfolge der Lernenden in der Klasse vor und nach der Intervention anhand eines Tests erhoben werden, um ein Maß für die externe Validität zu generieren. Dafür müsste die Intervention aber eher langfristig angelegt sein, weil sich ansonsten die Änderungen in der Leistung der Lernenden nur schwer zeigen lassen würden.
- Für die Förderung der professionellen Unterrichtswahrnehmung ist es außerdem wissenswert, ob diese situative Fähigkeit domänen-, aspekt- oder themenspezifisch ist, weil mit diesem Wissen adaptive Lerngelegenheiten für Aus- und Fortbildung geschaffen werden könnten. Studien, die diese verschiedenen Dimensionen systematisch ändern, könnten Aufschluss über die Spezifität der professionellen Unterrichtswahrnehmung geben. Abhängig von den Ergebnissen einer solchen Studie könnte die Fähigkeit von Lehramtsstudierenden dementsprechend gefördert werden.

Mit den Erfahrungen aus der vorliegenden Studie lässt sich für zukünftige Studien ableiten, dass es sinnvoll ist, die Inhaltsvalidität zu priorisieren, um eine Fähigkeit zu messen, die der Fähigkeit im Unterricht möglichst gut entspricht. Das bedeutet, dass für Studien zu Entwicklung eines Instruments als erstes eine klare Definition des zu messenden Konstrukts entwickelt werden sollte. Jeder Schritt einer Studie zur Entwicklung eines Instruments (oder auch zur Messung einer Fähigkeit) muss dann mit dieser Definition abgestimmt werden. Die Passung der Elemente für das Instrument zu dem Konstrukt kann beispielsweise unter Hinzuziehung weiterer Personen mit Expertise im Bereich der zu messenden Fähigkeit gesichert werden. Die Ergebnisse dürfen außerdem nur in Hinblick auf die durch die Konstruktdefinition klar definierte Fähigkeit interpretiert werden. Verallgemeinerungen, die über diese Definition hinausgehen, sollte nur vorsichtig formuliert und möglicherweise eher als Ausgangshypothesen für weitere Untersuchungen verstanden werden.

Damit die sorgfältigen Überlegungen zur Erhebung der professionellen Unterrichtswahrnehmung auch im gewünschten Verhalten der Testpersonen resultieren, wird des Weiteren angeraten, den Umgang mit dem Instrument zunächst üben zu lassen, wie es in der vorliegenden Studie praktiziert wurde. Um dadurch aber keinen Trainingseffekt zu messen, sollte die Übung andere als die in der Studie untersuchten Inhalte betreffen.

8.3.2 Implikationen für die Aus- und Fortbildung von Physiklehrkräften

Neben den Implikationen bezüglich zukünftiger Forschung lassen sich aus den Ergebnissen der vorliegenden Studie ebenfalls Konsequenzen für die Aus- und Fortbildung von Physiklehramtsstudierenden ableiten. Da die professionelle Unterrichtswahrnehmung eine Fähigkeit ist, die unter anderen den Lehrkräften eine adaptive Lernunterstützung ermöglicht, sollte die Fähigkeit, Lerngelegenheiten im Unterricht professionell wahrzunehmen, bereits in der universitären Lehrkräfteausbildung gefördert werden. Wie diese Förderung aussehen kann, wird im Folgenden berichtet:

- Die professionelle Unterrichtswahrnehmung ist offenbar eine Voraussetzung für die Lernbegleitung und -unterstützung und daher sollten Lehrkräfte darin geschult werden, ihre Aufmerksamkeit lernförderlich auf kritische Situationen im Unterricht zu richten.

Während der universitären Lehramtsausbildung sollte dafür zunächst die Dispositionen mit Bezug zur professionellen Unterrichtswahrnehmung gefördert werden (das fachliche, fachdidaktische und pädagogische Wissen, sowie affektiv-motivationale Anteile). Daneben sollte aber auch die Fähigkeit, Unterricht professionell wahrzunehmen, gefördert werden. Wie in Kapitel 3 zum Stand der empirischen Forschung berichtet, kann letzteres beispielsweise durch die Arbeit mit Unterrichtsvideos in Seminaren geschehen (Gold et al., 2013; Sunder et al., 2016). Dazu sollte den Studierenden die Möglichkeit geben werden, beispielsweise in Unterrichtsvideos professionell wahrzunehmen und somit die Dimensionen der professionellen Unterrichts-

wahrnehmung sichtbar gemacht werden. Anschließend können alle Seminarteilnehmenden über diesen situativen Prozess reflektieren. Ähnlich wie in den Video Clubs von Sherin & van Es (2008) könnte man mit den Studierenden eher mit einer reflexiven Haltung über die lernrelevanten Vorkommnisse in den – möglichst realen und gegebenenfalls eigenen – Unterrichtsvideos ins Gespräch kommen. Auf diese Weise lernen die Studierenden andere Perspektiven auf das Gesehene kennen und können zusätzlich durch Beobachtungsaufträge gezielt neue Aspekte fokussieren. In schulpraktischen Phasen können die Studierenden ebenfalls – durch gezielte Aufträge der Praktikumsbegleitung – verschiedene Fokusse während des Unterrichtens legen. Das sind mögliche Schritte, um die Aufmerksamkeitslenkung zu trainieren. An dieser Stelle wird der Zusammenhang zwischen der professionellen Unterrichtswahrnehmung der Reflexion wieder deutlich: Zum einen ist die Wahrnehmung im Unterricht Voraussetzung für die Reflexion (vgl. Rosaen, Lundeborg, Cooper, Fritzen & Terpstra, 2008, beispielsweise) und zum anderen kann die Reflexion dabei helfen, die eigene professionelle Unterrichtswahrnehmung weiterzuentwickeln. Der Zusammenhang dieser beiden Fähigkeiten sollte auch empirisch untersucht werden, um mögliche Synergieeffekte nutzen zu können. Desweiteren müsste den Studierenden die Notwendigkeit von kognitiven aber auch affektiv-motivationalen Dispositionen transparent gemacht werden, um ihre Bereitschaft zu fördern. Dies könnte beispielsweise anhand von Beispielen geschehen, in denen eine mangelnde (Bereitschaft zur) professionelle(n) Unterrichtswahrnehmung zu nicht lernförderlichen Verhalten der Lehrkraft geführt hat. Dabei gestaltet es sich allerdings schwierig, den Studierenden Kriterien zu bieten, anhand derer sie beurteilen, wann ihre professionelle Unterrichtswahrnehmung mit welchem Fokus notwendig ist, ohne dabei allzu normativ zu werden. Aufgrund der teilweise geringen Ausprägung der professionellen Unterrichtswahrnehmung bei Lehramtsstudierenden ist es allerdings empfehlenswert, Arbeiten mit den Unterrichtsvideos enger begleitet zu halten, als es im PUPiL-Instrument der Fall war.

- Damit die erkannten kritischen Situationen auch in einer Lerngelegenheit für die Lernenden resultieren, müssen die Lehramtsstudierenden auch lernen, das Erkannte möglichst evidenzbasiert zu interpretieren und entsprechende Handlungspläne zu fassen beziehungsweise umzusetzen. Das kann durch die wissensgesteuerte Reflexion während der Video Club ähnlichen Seminare oder nach hospitierten Unterrichtsstunden in Praxisphasen geübt werden. Wenn die Studierenden nicht lernen, wie man mit kritischen Situationen umgehen kann, nützt das Erkennen der Situation den Lernenden eher wenig, weil die Lernenden unter Umständen nicht die benötigte Förderung erhalten. Voraussetzung dafür ist selbstverständlich, dass die Studierenden die entsprechenden Wissensbestände bereits erlernt haben.
- Ähnliches gilt auch für das Fassen eines Handlungsplans. Die Lehramtsstudierenden sollten zwar nicht rezeptartig lernen, in welchem Fall sie wie reagieren müssen, um Lernen zu fördern. Dennoch bietet die universitäre Lehrkräfteausbildung die Gelegenheit, das Repertoire an möglichen, lernförderlichen Handlungen zu erweitern

und in schulpraktischen Phasen zu lernen, diese situativ und vom Kontext abhängig einzusetzen.

9 Schlusswort

Die professionelle Unterrichtswahrnehmung ist eine grundlegende Fähigkeit, um Lernprozesse adaptiv zu unterstützen. In der vorliegenden Arbeit ist es gelungen, diese Fähigkeit handlungsnäher und somit valider zu messen als es zuvor geschehen ist. Das Resultat und die Erkenntnisse dieser Arbeit tragen damit zum Wissensstand bezüglich der Entwicklung valider Instrumente zur Erfassung der latenten Fähigkeiten der professionellen Unterrichtswahrnehmung in der (physik-)didaktischen Forschung bei. Folgendes Vorgehen hat laut den Erfahrungen dieser Studie dazu geführt, ein Instrument zu erhalten, durch welches valide Aussagen über Fähigkeiten wie die professionelle Unterrichtswahrnehmung abgeleitet werden können:

1. *Klare Definition und Operationalisierung der zu messenden Fähigkeit:* Am Anfang der Entwicklung des Instruments zur Erfassung der professionellen Unterrichtswahrnehmung stand eine klare Definition dieser Fähigkeit und eine Eingrenzung auf den spezifischen Teil der Fähigkeit, der gemessen werden soll. Diese Definition ist notwendig, weil alle folgenden Schritte sowie die Interpretation der erhobenen Daten auf der Konstruktdefinition basieren. Für diese Arbeit wurde die professionelle Unterrichtswahrnehmung nur in Hinblick auf die Wahrnehmung von physikspezifischen Aspekten wie Schülervorstellungen und Erkenntnisgewinnung konzeptualisiert. Außerdem wurden die Situationen, in der diese (spezifisch eingegrenzte) Fähigkeit zum Tragen kommt, und der Prozess an sich klar operationalisiert. Für die physikspezifische Unterrichtswahrnehmung sind dies Situationen im Physikunterricht, die relevant für das Fachlernen der Klasse sind.
2. *Sicherung der Inhaltsvalidität:* Alle Elemente des Instruments wurden auf Grundlage der Konstruktdefinition entwickelt. Durch Expertenrating wurde die inhaltliche, curriculare und ökologische Validität der einzelnen Elemente für das Instrument gesichert.
3. *Erhebungssituation handlungsnah gestalten:* Das Vorgehen in der Erhebungssituation wurde nah an die Ausführung der Fähigkeit, Unterricht professionell wahrzunehmen, angelehnt. Das bedeutet, dass die Testpersonen selbst mit der Unterrichtssequenz interagieren können und in situ lernförderliche Handlungspläne aufstellen müssen, statt erst im Nachhinein über die vorgegebenen Situationen zu urteilen. Damit ist sichergestellt worden, dass genau die Fähigkeit gemessen wird, über die Informationen gewonnen werden sollen. Zu dieser Handlungsnähe gehört aber nicht nur die Nähe zur Tätigkeit, sondern auch dass die Testpersonen Informationen erhalten, zu denen sie in einer realen Unterrichtssituation auch Zugang hätten.

Für eine ökonomische Umsetzung und Auswertung der Erhebung mussten Entscheidungen getroffen werden, die zu einer verminderten Handlungsnahe führten:

- Im Instrument werden gestellte Unterrichtsvideos statt Videos realen Unterrichts verwendet, was dazu führt, dass die Testpersonen nur wenig über die Unterrichtsplanung, -zielsetzung und die Lernenden in den Videos erfahren. Da die professionelle Unterrichtswahrnehmung mit Erwartungen, Erfahrungen und Wissen zusammenhängt, beeinträchtigt dies die Validität der Messung.
 - Die geschlossenen Multiple-Choice-Items für die Hauptstudie wurden aufbauend auf den Verbaldaten von Testpersonen in Vorstudien entwickelt. Diese Formulierungen führen möglicherweise zu einer extrinsischen Lenkung der Aufmerksamkeit.
 - Das Instrument setzt einen Fokus für die Wahrnehmung durch die Aufgabenstellung, womit wiederum eine Aufmerksamkeitslenkung durch das Instrument selbst initiiert wird.
4. *Angemessene Interpretationen:* Die Daten, die durch die Erhebung generiert wurden, wurden nur unter Rückbezug auf die klare Konstruktdefinition und die entsprechenden Limitationen interpretiert. Das hat zur Folge, dass nur Aussagen über die physikspezifische professionelle Unterrichtswahrnehmung in Bezug auf Schülervorstellungen und Erkenntnisgewinnung der Lernenden abgeleitet werden konnten. Die Generalisierung dieser Aussagen ist auf Basis des Instruments nicht möglich. Valide Schlussfolgerungen können nur über die definierte Teilfähigkeit gezogen werden.

Bisherige Instrumente zur Erfassung der professionellen Unterrichtswahrnehmung basieren meist darauf, dass Testpersonen Unterrichtssequenzen im Sinne des advokatorischen Ansatzes mithilfe von Likertskalen einschätzen. Die Testpersonen konzentrieren sich dabei in einer reflexiven Haltung auf die Handlungen der Lehrkräfte statt auf die Lernprozesse der Lernenden, wie sie es im eigenen Unterricht täten. Das Instrument der vorliegenden Arbeit bietet den Testpersonen hingegen die Möglichkeit, ihre Aufmerksamkeit auf die Lernprozesse der Klasse zu richten, ohne dabei eine reflexive Haltung dem Gesehenen gegenüber einzunehmen und misst somit eine Fähigkeit, die näher an der professionellen Unterrichtswahrnehmung ist, als die vorherigen Instrumente.

In der Zukunft könnten weitere Optimierungen vorgenommen werden, um die professionelle Unterrichtswahrnehmung noch valider zu messen. Statt geschlossener Fragen könnten ausschließlich offene Fragen gestellt werden. Weil Lehrkräfte ihren Unterricht planen und daher im Vorwege Erwartungen an die Unterrichtsstunde haben, was ihre Aufmerksamkeitslenkung beeinflusst, sollten sich Testpersonen ebenfalls ihre Erwartungen vor dem Sehen der Unterrichtsvideos bewusst machen und möglicherweise eine detaillierte Planung der wahrzunehmenden Unterrichtsstunde erhalten.

Final kann die Forschungsfrage dieser Dissertation folgendermaßen beantwortet werden: Die physikspezifische professionelle Unterrichtswahrnehmung (von Schülervorstellungen und Erkenntnisgewinnung) kann mithilfe von Videos und einem teilweise geschlossenen

Fragebogen durch das entwickelte Instrument valide gemessen werden. Das bedeutet aber nicht, dass durch Veränderungen des bestehenden Instruments nicht auch eine höhere Validität der ableitbaren Aussagen über die professionelle Unterrichtswahrnehmung erreicht werden kann.

Literatur

- Baldinger, E. M. (2017). „Maybe It’s a Status Problem.“: Development of Mathematics Teacher Noticing for Equity. In E. O. Schack, M. H. Fisher & J. A. Wilhelm (Hrsg.), *Teacher Noticing* (S. 231–249). Research in Mathematics Education. Cham: Springer International Publishing.
- Barmeier, M. (Hrsg.). (2013). *Prisma Physik* (Ausg. A, differenzierende Ausg., 1. Aufl., Dr. 1). Stuttgart: Klett.
- Bartels, H., Kulgemeyer, C. & Schecker, H. (2016). Entwicklung eines computerbasierten Testinstruments für Erklärfähigkeit. *PhyDid B - Didaktik der Physik*, (Beiträge zur DPG-Frühjahrstagung).
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, 35(5), 463–482.
- Beschluss der Kultusministerkonferenz. (2004). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. Luchterhand Verlag.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond Dichotomies. *Zeitschrift für Psychologie*, 223(1), 3–13.
- Blömeke, S. & Kaiser, G. (2017). Understanding the Development of Teachers’ Professional Competencies as Personally, Situationally and Socially Determined. In D. J. Clandini & J. Hsu (Hrsg.), *International Handbook of Research on Teacher Education*. (S. 783–802). Sage Publishers: Thousand Oaks.
- Blömeke, S., Kaiser, G. & Clarke, D. (2015). Preface for the Special Issue on “Video-Based Research on Teacher Expertise”. *International Journal of Science and Mathematics Education*, 13(2), 257–266.
- Bögeholz, S., Hößle, C., Höttecke, D. & Menthe, J. (2018). Bewertungskompetenz. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (S. 261–281). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (2nd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., vollständig überarbeitete und erweiterte Auflage). Springer-Lehrbuch. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Brennan, R. L. & Prediger, D. J. (1981). Coefficient Kappa: Some Uses, Misuses, and Alternatives. *Educational and Psychological Measurement*, 41, 687–699.

- Briggs, D., Alonzo, A., Schwab, C. & Wilson, M. (2006). Diagnostic Assessment With Ordered Multiple-Choice Items. *Educational Assessment*, 11(1), 33–63.
- Bruckmaier, G., Krauss, S., Blum, W. & Leiss, D. (2016). Measuring mathematics teachers' professional competence by using video clips (COACTIV video). *ZDM*, 48(1-2), 111–124.
- Carter, K., Cushing, K., Sabers, D., Stein, P. & Berliner, D. (1988). Expert-Novice Differences in Perceiving and Processing Visual Classroom Information. *Journal of Teacher Education*, 39(3), 25–31.
- Cauet, E. (2016). *Testen wir relevantes Wissen? Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterricht*. Studien zum Physik- und Chemielernen. Berlin: Logos Berlin.
- Chan, E. K. H. (2014). Standards and Guidelines for Validation Practices: Development and Evaluation of Measurement Instruments. In B. D. Zumbo & E. K. Chan (Hrsg.), *Validity and Validation in Social, Behavioral, and Health Sciences* (Bd. 54, S. 9–24). Social Indicators Research Series. Cham: Springer International Publishing.
- Christmann, U. (1999). Wahrnehmung, Kognition und Metakognition. Köln. Zugriff unter <http://www.uni-koeln.de/dfg-spp-lesesoz/Literatur/begriffe.pdf>
- diSessa, A. A. (1988). Knowledge in Pieces. In G. Forman & P. B. Pufall (Hrsg.), *Constructivism in the Computer Age* (S. 49–70). Jean Piaget Symposia Series. Hoboken: Taylor and Francis.
- Düchs, G. & Ingold, G.-L. (2017). Physik hält Kurs: Statistiken zum Physikstudium an den Universitäten in Deutschland 2017. *Physik Journal*, 16(8/9), 28–33.
- Duit, R. (1986). Der Energiebegriff im Physikunterricht. *Habilitationsschrift. Universität Kiel*.
- Duit, R., Hopf, M., Schecker, H. & Wilhelm, T. (Hrsg.). (2018). *Schülervorstellungen und Physikunterricht: Ein Lehrbuch für Studium, Referendariat und Unterrichtspraxis*. Berlin: Springer Spektrum.
- Dunekacke, S., Jenßen, L., Eilerts, K. & Blömeke, S. (2016). Epistemological beliefs of prospective preschool teachers and their relation to knowledge, perception, and planning abilities in the field of mathematics: a process model. *ZDM*, 48(1-2), 125–137.
- Dyer, E. B. (2013). Investigating the Relationship Between Teacher Professional Vision and Classroom Practices: A Case of Misalignment. In M. Martinez & A. Castro Superfine (Hrsg.), *Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Chicago, IL.
- Dyer, E. B. & Sherin, M. G. (2016). Instructional reasoning about interpretations of student thinking that supports responsive teaching in secondary mathematics. *ZDM*, 48(1-2), 69–82.
- Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1), 32–64.

- Erickson, F. (2011). On noticing teacher noticing. In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Hrsg.), *Mathematics teacher noticing* (S. 17–34). Studies in mathematical thinking and learning. New York: Routledge.
- Feser, M. S. (eingereicht). *Physiklehrkräfte korrigieren Schülertexte: Eine Explorationsstudie zur fachlich-konzeptuellen und sprachlichen Leistungsfeststellung und -beurteilung im Physikunterricht* (Dissertation, Universität Hamburg, Hamburg).
- Floro, B. & Bostic, J. D. (2017). A Case Study of Middle School Teachers' Noticing During Modeling with Mathematics Tasks. In E. O. Schack, M. H. Fisher & J. A. Wilhelm (Hrsg.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (S. 73–89). Research in Mathematics Education. Cham: Springer International Publishing.
- Gebhard, U., Höttecke, D. & Rehm, M. (Hrsg.). (2017). *Pädagogik der Naturwissenschaften: Ein Studienbuch*. Lehrbuch. Wiesbaden: Springer VS.
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41(6), 867–888.
- Gold, B., Förster, S. & Holodynski, M. (2013). Evaluation eines videobasierten Trainingsseminars zur Förderung der professionellen Wahrnehmung von Klassenführung im Grundschulunterricht. *Zeitschrift für Pädagogische Psychologie*, 27(3), 141–155.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist, New Series*, 96(3), 606–633.
- Gramzow, Y. (2015). *Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik: Modellierung und Testkonstruktion*. Studien zum Physik- und Chemielernen. Berlin: Logos.
- Gropengießer, H., Harms, U. & Kattmann, U. (Hrsg.). (2013). *Fachdidaktik Biologie* (9. Auflage). Seelze: Aulis Verlag.
- Hadenfeldt, J. C., Repenning, B. & Neumann, K. (2014). Die kognitive Validität von Ordered Multiple Choice Aufgaben zur Erfassung des Verständnisses von Materie. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 57–68.
- Hartig, J. & Frey, A. (2013). Sind Modelle der Item-Response-Theorie (IRT) das „Mittel der Wahl“ für die Modellierung von Kompetenzen? In D. Leutner, E. Klieme & Fleischer, J. Kuper, H. (Hrsg.), *Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen: aktuelle Diskurse im DFG-Schwerpunktprogramm*. 18. Sonderheft der Zeitschrift für Erziehungswissenschaft. Wiesbaden: VS.
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 143–171). Springer-Lehrbuch. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Hattie, J. (2014). *Lernen sichtbar machen: Überarbeitete deutschsprachige Ausgabe von „Visible Learning“ besorgt von Beywl, W. und Zierer K.* (2., korrigierte Aufl.). Baltmannsweiler: Schneider-Verl. Hohengehren.
- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (5. Auflage). Unterricht verbessern - Schule entwickeln. Seelze-Velber: Klett Kallmeyer.

- Hertel, I. V. & Großmann, S. (Hrsg.). (2016). *Physik in der Schule: Hauptteil mit Anlage Basiskonzepte*.
- Honskusová, L. & Rusek, M. (2011). Evaluation of the Effectiveness of a Pre-Service Chemistry Teachers' Professional Vision Development Course.
- Hopf, M. & Wilhelm, T. (2018). Conceptual Change - Entwicklung physikalischer Vorstellungen. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 23–38). Berlin: Springer Spektrum.
- Hoth, J., Döhrmann, M., Kaiser, G., Busse, A., König, J. & Blömeke, S. (2016). Diagnostic competence of primary school mathematics teachers during classroom situations. *ZDM*, 48(1-2), 41–53.
- Höttecke, D., Buth, K., Koenen, J., Masanek, N., Reichwein, W., Scholten, N., ... Wöhlke, C. (2018). Vernetzung von Fach und Fachdidaktik in der Hamburger Lehrerbildung. In I. Glowinski, J. Gillen, A. Borowski, S. Schanze & J. v. Meien (Hrsg.), *Kohärenz in der universitären Lehrerbildung* (S. 29–52). Potsdam: Universitätsverlag Potsdam.
- IBM Corp. (Released 2017). IBM SPSS Statistics for Windows. Version 26.0. Armonk, NY: IBM Corp.
- Irtel, H. (1996). *Entscheidungs- und testtheoretische Grundlagen der psychologischen Diagnostik*. Frankfurt am Main: Lang.
- Jacobs, V. R., Lamb, L. & Philipp, R. (2010). Professional Noticing of Children's Mathematical Thinking. *Journal for Research in Mathematics Education*, 41(2), 169–202.
- Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M. & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers—cognitive versus situated approaches. *Educational Studies in Mathematics*, 94(2), 161–182.
- Keppens, K., Consuegra, E., Goossens, M., Maeyer, S. d. & Vanderlinde, R. (2019). Measuring pre-service teachers' professional vision of inclusive classrooms: A video-based comparative judgement instrument. *Teaching and Teacher Education*, 78, 1–14.
- Kersting, N. B. (2008). Using Video Clips of Mathematics Classroom Instruction as Item Prompts to Measure Teachers' Knowledge of Teaching Mathematics. *Educational and Psychological Measurement*, 68(5), 845–861.
- Kersting, N. B., Givvin, K. B., Sotelo, F. L. & Stigler, J. W. (2010). Teachers' Analyses of Classroom Video Predict Student Learning of Mathematics: Further Explorations of a Novel Measure of Teacher Knowledge. *Journal of Teacher Education*, 61(1-2), 172–181.
- Kircher, E. (2015). Warum Physikunterricht? In E. Kircher (Hrsg.), *Physikdidaktik* (S. 15–73). Springer-Lehrbuch. Berlin: Springer Spektrum.
- Konrad, K. (2010). Lautes Denken. In G. Mey & K. Mruck (Hrsg.), *Handbuch qualitative Forschung in der Psychologie* (S. 476–490). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kramer, C., König, J., Kaiser, G., Ligtvoet, R. & Blömeke, S. (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. *Zeitschrift für Erziehungswissenschaft*, 20(S1), 137–164.
- Krauss, S., Lindl, A. & Schilcher, A. (Hrsg.). (2017). *FALKO: fachspezifische Lehrerkompetenzen: Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik*. Münster: Waxmann.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M. & Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, 29(3-4), 233–258.
- Krippendorff, K. (2011). Computing Krippendorff's Alpha-Reliability. Zugriff unter http://repository.upenn.edu/asc_papers/43
- Krupa, E. E., Huey, M., Kesseig, K., Casey, S. & Monson, D. (2017). Investigating Secondary Preservice Teacher Noticing of Students' Mathematical Thinking. In E. O. Schack, M. H. Fisher & J. A. Wilhelm (Hrsg.), *Teacher Noticing* (S. 49–72). Research in Mathematics Education. Cham: Springer International Publishing.
- Kulgemeyer, C. & Schecker, H. (2009). Physikalische Darstellungsformen: Ein Beitrag zur Klärung von Kommunikationskompetenz. *Der mathematische und naturwissenschaftliche Unterricht*, 62(6), 331–336.
- Kunter, M., Baumert, J. & Blum, W. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. StandardWissen Lehramt. Paderborn: Schöningh.
- Lamme, V. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 7(1), 12–18.
- Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174.
- Lazarevic, C. (2016). *Professionelle Wahrnehmung und Analyse von Unterricht durch Mathematiklehrkräfte: Eine fallrekonstruktive Studie*. Springer-Verlag.
- Lee, M. Y. & Choy, B. H. (2017). Mathematical Teacher Noticing: The Key to Learning from Lesson Study. In E. O. Schack, M. H. Fisher & J. A. Wilhelm (Hrsg.), *Teacher Noticing* (S. 121–140). Research in Mathematics Education. Cham: Springer International Publishing.
- Lefstein, A. & Snell, J. (2011). Professional vision and the politics of teacher learning. *Teaching and Teacher Education*, 27(3), 505–514.
- Lin, J., Chu, K.-L. & Meng, Y. (2010). Distractor rationale taxonomy: Diagnostic assessment of reading with ordered multiple-choice items. *American Educational Research Association, Denver, CO*.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an: Empirische Evidenz für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung im Lehrerberuf. Ausbildung und Beruf* (S. 47–70). 51. Beiheft. Weinheim: Beltz.

- Lipowsky, F. (2007). Was wissen wir über guten Unterricht? Im Fokus: die fachliche Lernentwicklung. *Friedrich Jahresheft*, 26–29.
- Males, L. M. (2017). Using Video of Peer Teaching to Examine Grades 6-12 Preservice Teachers' Noticing. In E. O. Schack, M. H. Fisher & J. A. Wilhelm (Hrsg.), *Teacher Noticing* (S. 91–109). Research in Mathematics Education. Cham: Springer International Publishing.
- Martin, J.-P. (1988). Schüler in komplexen Lernumwelten: Vorschlag eines kognitionspsychologisch fundierten Curriculums für den Fremdsprachenunterricht. *Praxis des neu sprachlichen Unterrichts*, (3), 294–302.
- Mason, J. (2011). Noticing: Roots and branches. In *Mathematics teacher noticing* (S. 65–80). Routledge.
- Mausfeld, R. (2006). Wahrnehmung: Geschichte und Ansätze. *Handbuch der Allgemeinen Psychologie-Kognition*, 97–107.
- Mausfeld, R. (2010). Wahrnehmungspsychologie. In A. Schütz (Hrsg.). H. Selg (Hrsg.). M. Brand (Hrsg.). S. Lautenbacher (Hrsg.), *Psychologie. Eine Einführung in ihre Grundlagen und Anwendungsfelder*. Stuttgart: Kohlhammer.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarb. Aufl.). Beltz Pädagogik. Weinheim: Beltz.
- McClelland, D. C. (1973). Testing for competence rather than for intelligence. *American Psychologist*, 28(1), 1–14.
- Merzyn, G. (2015). Guter Physikunterricht: Die Sicht von Schülern, Lehrern und Wissenschaftlern. *Didaktik der Physik*.
- Meschede, N., Fiebranz, A., Möller, K. & Steffensky, M. (2017). Teachers' professional vision, pedagogical content knowledge and beliefs: On its relation and differences between pre-service and in-service teachers. *Teaching and Teacher Education*, 66, 158–170.
- Meyer, H. (2003). Zehn Merkmale guten Unterrichts: Empirische Befunde und didaktische Ratschläge. *Pädagogik*, 10, 36–43.
- Neisser, U. (1976). *Cognition and reality: Principles and implications*. San Francisco: Freeman.
- Neumann, K., Kauertz, A., Lau, A., Notarp, H. & Fischer, H. E. (2007). Die Modellierung physikalischer Kompetenz und ihrer Entwicklung. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 101–121.
- Neuweg, G. (2014). Das Wissen der Wissensvermittler: Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrerwissen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 583–614).
- Niebert, K. (2010). *Den Klimawandel verstehen: Eine didaktische Rekonstruktion der globalen Erwärmung* (1. Aufl.). Beiträge zur didaktischen Rekonstruktion. Hannover: Technische Informationsbibliothek u. Universitätsbibliothek & Didaktisches Zentrum, Carl von Ossietzky Univ.
- Nienhaus, H. (2018). *Physik für das Lehramt*. De Gruyter.
- Oser, F., Heinzer, S. & Salzmann, P. (2010). Die Messung der Qualität von professionellen Kompetenzprofilen von Lehrpersonen mit Hilfe der Einschätzung von Filmvignetten:

- Chancen und Grenzen des advokatorischen Ansatzes. *Unterrichtswissenschaft*, 38(H. 1).
- Pankow, L., Kaiser, G., Busse, A., König, J., Blömeke, S., Hoth, J. & Döhrmann, M. (2016). Early Career Teachers' ability to focus on typical students errors in relation to the complexity of a mathematical topic. *ZDM*, 48(1-2), 55–67.
- Peetz, M. K. (2019). *Evaluation von Schülervorstellungen mithilfe von Animationen: Der Löseprozess von Zucker und Salz in Wasser* (Dissertationsschrift, Carl-von-Ossietzky Universität Oldenburg, Oldenburg).
- Petersen, P. (1946). *Der kleine Jena-Plan*. J. Beltz.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen* (1. Aufl.). Konzepte der Humanwissenschaften. Stuttgart: Klett.
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science education*, 66, 2001–2227.
- Reichhart, B. (2018). *Lehrerprofessionalität im Bereich der politischen Bildung: Eine Studie zu motivationalen Orientierungen und Überzeugungen im Sachunterricht*. Empirische Forschung in den gesellschaftswissenschaftlichen Fachdidaktiken. Wiesbaden: Springer Fachmedien Wiesbaden.
- Renkl, A. (2008). Lernen und Lehren im Kontext der Schule. In A. Renkl (Hrsg.), *Lehrbuch Pädagogische Psychologie* (S. 109–153). Bern: Huber.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften: Zugl.: Paderborn, Univ., Diss., 2009*. Studien zum Physik- und Chemielernen. Berlin: Logos Verlag.
- Robitzsch, A. (2019). Test Analysis Modules. Zugriff unter <https://cran.r-project.org/web/packages/TAM/TAM.pdf>
- Rosaen, C. L., Lundeberg, M., Cooper, M., Fritzen, A. & Terpstra, M. (2008). Noticing Noticing: How Does Investigation of Video Records Change How Teachers Reflect on Their Experiences? *Journal of Teacher Education*, 59(4), 347–360.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion* (2., vollst. überarb. und erw. Aufl.). Psychologie Lehrbuch. Bern: Huber.
- Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K. & Wickler, N. I. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, 48(2), 117–148.
- RStudio Team. (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc. Boston, MA. Zugriff unter <http://www.rstudio.com/>
- Sabers, D., Cushing, K. & Berliner, D. (1991). Differences Among Teachers in a Task Characterized by Simultaneity, Multidimensional, and Immediacy. *American Educational Research Journal*, 28(1), 63–88.
- Santagata, R. & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM*, 43(1), 133–145.
- Santagata, R. & Yeh, C. (2016). The role of perception, interpretation, and decision making in the development of beginning teachers' competence. *ZDM*, 48(1-2), 153–165.

- Santagata, R., Zannoni, C. & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: an empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education*, 10(2), 123–140.
- Schack, E. O., Fisher, M. H. & Wilhelm, J. A. (Hrsg.). (2017). *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks*. Research in Mathematics Education. Cham: Springer International Publishing.
- Schäfer, S. & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. *Journal for Educational Research Online*, 7(2), 34–58.
- Schecker, H. & Höttecke, D. (2007). Aufgaben zum Kompetenzbereich Bewerten. *Naturwissenschaften im Unterricht - Physik*, 18(97), 29–36.
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 45–67.
- Scheiner, T. (2016). Teacher noticing: Enlightening or blinding? *ZDM*, 48(1-2), 227–238.
- Schlesinger, L., Jentsch, A., Kaiser, G., König, J. & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM*, 50(3), 475–490.
- Schmiemann, P. & Lücken, M. (2014). Validität–Misst mein Test, was er soll? In *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 107–118). Springer.
- Schoenfeld, A. H. (2011). Noticing matters. A lot. Now what? In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Hrsg.), *Mathematics teacher noticing* (S. 223–238). Studies in Mathematical Thinking and Learning. New York: Routledge.
- Scholten, N., Höttecke, D. & Sprenger, S. (2018). Conceptualizing Geography Teachers' Subject-Specific Noticing During Instruction. *European Journal of Geography*, 9(3), 80–97.
- Scholten, N., Höttecke, D. & Sprenger, S. (2019). How do geography teachers notice critical incidents during instruction? *International Research in Geographical and Environmental Education*, 3(2), 1–15.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. [New York]: Basic Books.
- Seidel, T. (2009). Klassenführung. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 135–148). Springer-Lehrbuch. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Seidel, T., Blomberg, G. & Stürmer, K. (2010). Observer - Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE. *Zeitschrift für Pädagogik*, 56(Beiheft), 296–306.
- Seidel, T. & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 201–216). Zeitschrift für Erziehungswissenschaft Sonderheft. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden.

- Seidel, T., Prenzel, M., Rimmele, R., Dalehefte, I. M., Herweg, C., Kobarg, M. & Schwindt, K. (2006). Blicke auf den Physikunterricht. Ergebnisse der IPN Videostudie. *Zeitschrift für Pädagogik*, 52(6), 799–821.
- Seidel, T. & Stürmer, K. (2014). Modeling and Measuring the Structure of Professional Vision in Preservice Teachers. *American Educational Research Journal*, 51(4), 739–771.
- Sherin, M. G. (2001). Developing a professional vision of classroom events: Teaching elementary school mathematics. In *Beyond classical pedagogy* (S. 75–93). Erlbaum.
- Sherin, M. G., Jacobs, V. R. & Philipp, R. A. (Hrsg.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. Studies in Mathematical Thinking and Learning. New York: Routledge.
- Sherin, M. G., Russ, R. S., Sherin, B. & Colestock, A. A. (2008). Professional Vision in Action: An Exploratory Study. *Issues in Teacher Education*, Fall, 27–46.
- Sherin, M. G. & van Es, E. A. (2008). Effects of Video Club Participation on Teachers' Professional Vision. *Journal of Teacher Education*, 60(1), 20–37.
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 4–14.
- Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1–22.
- Stahnke, R., Schueler, S. & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM*, 48(1-2), 1–27.
- Star, J. R. & Strickland, S. K. (2008). Learning to observe: using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, 11(2), 107–125.
- Stark, R. (2003). Conceptual Change: kognitiv oder situiert? *Zeitschrift für Pädagogische Psychologie*, 17(2), 133–144.
- Steffensky, M., Gold, B., Holodynski, M. & Möller, K. (2015). Professional Vision of Classroom Management and Learning Support in Science Classrooms—Does Professional Vision Differ Across General and Content-Specific Classroom Interactions? *International Journal of Science and Mathematics Education*, 13(2), 351–368.
- Stöcker, K. (1975). *Neuzeitliche Unterrichtsgestaltung*. Ehrenwirth.
- Strobl, C. (2010). *Das Rasch-Modell: Eine verständliche Einführung für Studium und Praxis* (1. Aufl.). Sozialwissenschaftliche Forschungsmethoden. München: Hampp.
- Stüber, A. T. (2016). *Das Partial Credit Modell* (Bachelorarbeit, Ludwig-Maximilians-Universität, München).
- Stürmer, K. (2011). *Voraussetzungen für die Entwicklung professioneller Unterrichtswahrnehmung im Rahmen universitärer Lehrerbildung* (Dissertation, Technische Universität München, München).
- Stürmer, K., Könings, K. D. & Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: effect of courses in teaching and learning. *The British journal of educational psychology*, 83(Pt 3), 467–483.

- Stürmer, K., Könings, K. D. & Seidel, T. (2015). Factors Within University-Based Teacher Education Relating to Preservice Teachers' Professional Vision. *Vocations and Learning*, 8(1), 35–54.
- Stürmer, K. & Seidel, T. (2015). Assessing Professional Vision in Teacher Candidates. *Zeitschrift für Psychologie*, 223(1), 54–63.
- Stürmer, K., Seidel, T., Müller, K., Häusler, J. & Cortina, K. S. (2017). What is in the eye of preservice teachers while instructing? An eye-tracking study about attention processes in different teaching situations. *Zeitschrift für Erziehungswissenschaft*, 20(S1), 75–92.
- Sunder, C., Todorova, M. & Möller, K. (2016). Kann die professionelle Unterrichtswahrnehmung von Sachunterrichtsstudierenden trainiert werden? – Konzeption und Erprobung einer Intervention mit Videos aus dem naturwissenschaftlichen Grundschulunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 1–12.
- Szogs, M., Kobl, C., Volmer, M. & Korneck, F. (2019). Bedeutsamkeit von Reflexion und Reflexivität in der Professionalisierung von Lehrkräften sowie ihre Beziehung zu anderen Prozessen und Konstrukten. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 317–320). Regensburg: Universität Regensburg.
- Tenorth, H.-E. & Tippelt, R. (2007). Wahrnehmung. In H.-E. Tenorth & R. Tippelt (Hrsg.), *Beltz Lexikon Pädagogik* (S. 761). Weinheim: Beltz Verlagsgruppe.
- Terhart, E. (2011). Lehrerberuf und Professionalität: Gewandeltes Begriffsverständnis - neue Herausforderungen. *Zeitschrift für Pädagogik*, 57, 202–224.
- Thode, D. & Schecker, H. (2017). Bewertung im Fokus: Unterrichtskonzeption für die gymnasiale Oberstufe. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (S. 688–691). Regensburg : Universität Regensburg.
- Treagust, D. F., Chittleborough, G. & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education*, 24(4), 357–368.
- Treisch, F. (2018). *Die Entwicklung der Professionellen Unterrichtswahrnehmung im Lehr-Lern-Labor Seminar*. Logos Verlag Berlin.
- van Es, E. A. & Sherin, M. G. (2002). Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *Journal of Technology and Teacher Education*, 10(4), 571–596.
- van Es, E. A. & Sherin, M. G. (2006). How Different Video Club Designs Support Teachers in Learning to Notice. *Journal of Computing in Teacher Education*, 22(4), 125–135.
- Vogelsang, C. (2014). *Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften: Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz: Zugl.: Paderborn, Univ., Diss., 2014*. Studien zum Physik- und Chemielernen. Berlin: Logos Berlin.
- von Aufschnaiter, C., Cappell, J., Dübbelde, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., ... Wolgast, A. (2015). Diagnostische Kompetenz: Theoretische Überle-

- gungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, 61(5), 738–758.
- von Aufschnaiter, C., Fraij, A. & Kost, D. (2019). Reflexion und Reflexivität in der Lehrerbildung. *Herausforderung Lehrer_innenbildung - Zeitschrift zur Konzeption, Gestaltung und Diskussion*, 2(1), 144–159.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4(1), 45–69.
- Wagenschein, M. (1982). *Verstehen lehren: Genetisch - Sokratisch - Exemplarisch* (7. durchges. Aufl. Weinheim). Basel: Beltz.
- Wahl, D. (1991). *Handeln unter Druck: Der weite Weg vom Wissen zum Handeln bei Lehrern, Hochschullehrern und Erwachsenenbildern: Zugl.: Tübingen, Univ., Habil.-Schr., 1990*. Weinheim: Dt. Studien-Verl.
- Wahl, D. (2001). Nachhaltige Wege vom Wissen zum Handeln. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 19(2), 157–174.
- Weger, D. (2019). Professional Vision - State of the art zum Konstrukt der professionellen Unterrichtswahrnehmung in der Lehrer- (innen)bildung. *Fremdsprachen Lehren und Lehren*, 48(1), 14–31.
- Weinert, F. (2001a). Concept of competence. A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Seattle, Washington, USA: Hogrefe & Huber.
- Weinert, F. (2001b). *Leistungsmessungen in Schulen* (F. E. Weinert, Hrsg.). Beltz-Pädagogik. Weinheim: Beltz.
- Wells, K. J. (2017). Noticing Students' Conversations and Gestures During Group Problem-Solving in Mathematics. In E. O. Schack, M. H. Fisher & J. A. Wilhelm (Hrsg.), *Teacher Noticing* (S. 183–204). Research in Mathematics Education. Cham: Springer International Publishing.
- Wider, V., Schwichow, M. & Mikelskis-Seifert, S. (2019). Diagnostizieren von Modellvorstellungen im Physikunterricht. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 830–833). Regensburg: Universität Regensburg.
- Widodo, A. & Duit, R. (2004). Konstruktivistische Sichtweisen vom Lehren und Lernen und die Praxis des Physikunterrichts. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 233–255.
- Wiedmann, M. (2015). *Teacher's Professional Vision of Student Interaction in Collaborative Learning Settings: Development and Evaluation of an Assessment Tool* (Inaugural-Dissertation, Albert-Ludwigs-Universität, Freiburg i.Br.).
- Wolters, M. (2015). *Wie kompetent sind (angehende) Lehrkräfte in der professionellen Wahrnehmung kognitiv anregender Situationen im naturwissenschaftlichen Grundschulunterricht?* (Inaugural-Dissertation, Westfälische Wilhelms-Universität, Münster).

A Entwicklungsstudie I – Arbeitsblätter der Drehbücher

A.1 Drehbuch 2

Magnetfeld zweier Hufeisenmagneten

- 1.) Überlegt euch zu zweit eine Hypothese, wie das Magnetfeld von zwei benachbarten Hufeisenmagneten aussieht (siehe Abb. 1).



Abb. 1 Zwei benachbarte Hufeisenmagneten

Hypothese:

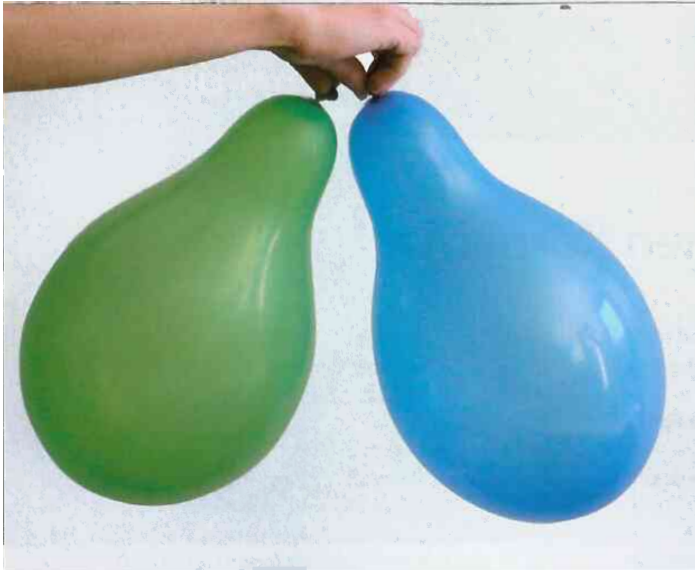
- 2.) Testet eure Hypothese anschließend mit folgendem Experiment.

- Materialien:
 - 2 Hufeisenmagneten
 - 1 Blatt Papier
 - Eisenspäne
- Durchführung:
 - Bringt zwei Hufeisenmagneten wie in Abb.1 zusammen. Legt ein Blatt Papier auf die beiden Magneten und streut VORSICHTIG Eisenspäne auf das Blatt Papier.
- Beobachtung:

- Ergebnisskizzen:

A.2 Drehbuch 3

Seite 1



1 Zwei gleich geladene Luftballons stoßen einander ab.



2 Ein positiv und ein negativ geladener Luftballon ziehen einander an.

⊙ Elektrische Kräfte

Anziehung und Abstoßung

Elektrisch geladene Körper können einander anziehen oder abstoßen. Das hängt davon ab, wie die Körper geladen sind.

Wenn du zwei Luftballons kräftig aneinander reibst und wieder trennst, ist danach der eine Luftballon positiv und der andere negativ geladen. Du beobachtest, dass die unterschiedlich geladenen Luftballons einander anziehen (▷ B2).

Eine andere Feststellung wirst du machen, wenn du zwei gleich geladene Luftballons einander näherst. Wenn beide Luftballons positiv oder beide negativ geladen sind, dann stoßen sie einander ab (▷ B1).

Geladene Körper ziehen ungeladene an

Elektrische Kräfte wirken auch zwischen elektrisch geladenen und neutralen Körpern. Deshalb zieht beispielsweise ein elektrisch geladener Luftballon ungeladene Papierschnipsel an.

Gleich geladene Körper stoßen einander ab. Unterschiedlich geladene Körper ziehen einander an.

AUFGABEN

- 1 ○ Nenne die Kraftwirkungen zwischen elektrisch geladenen Körpern.
- 2 ● Eine geriebene Folie „klebt“ ohne weitere Hilfsmittel an einer Wand. Probiere es aus und erkläre.
- 3 ● Anna behauptet: „Mit einem Kunststoffstab kann ich die Richtung eines Wasserstrahls ändern.“ Prüfe, ob Annas Aussage stimmt und begründe.

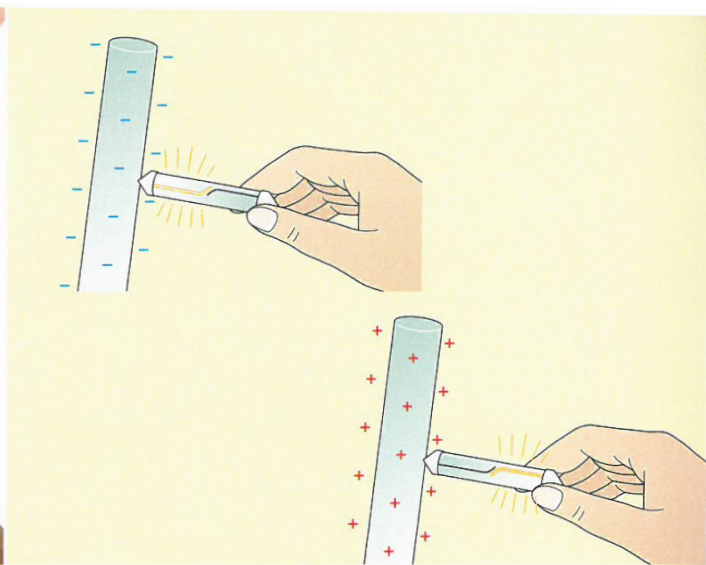
VERSUCHE

- 1 Schneide zwei breite Streifen Folie von einem Tiefkühlbeutel ab. Reibe die Folien kräftig aneinander und trenne sie danach voneinander. Beobachte, wie sich die Folien verhalten.
- 2 Lege einen Streifen Folie auf den Tisch und reibe ihn kräftig. Greife den Streifen in der Mitte und ziehe ihn vom Tisch ab. Beobachte das Verhalten der Folie.

Quelle:

254 Barmeier, M. (Hg.) (2013): Prisma Physik. Ausg. A, differenzierende Ausg., 1. Aufl., Dr. 1. Stuttgart: Klett, S. 254.

Seite 2



Haare werden von einem Luftballon angezogen. 2 Nachweis von elektrischen Ladungen

Elektrisch geladene Körper

Seltene Beobachtungen

Paula reibt einen Luftballon kräftig an ihren Haaren. Danach stehen ihr plötzlich „die Haare zu Berge“ (▷ B1). Vielleicht hat sie dabei auch ein Knistern gehört oder sogar kleine Funken gesehen. Hast du auch schon einmal eine ähnliche Erfahrung gemacht?

Laden und Entladen

Vor dem Reiben sind Luftballon und Haare nicht elektrisch geladen. Man sagt: Sie sind elektrisch neutral.

Wenn du den Luftballon an deinen Haaren reibst und ihn dann etwas abhebst, sind beide Körper **elektrisch geladen**. Der Luftballon und die Haare sind jedoch unterschiedlich geladen: Ein Körper ist **elektrisch positiv (+)**, der andere ist **elektrisch negativ (-)** geladen. Wenn sich die unterschiedlich geladenen Körper wieder berühren, dann entladen sie sich.

Die Glimmlampe

Mit der Glimmlampe kannst du **elektrische Ladungen** nachweisen. Wenn du den geladenen Luftballon mit der Glimmlampe berührst, blitzt die Glimmlampe auf. Dabei

wird der Luftballon an dieser Stelle entladen – sie ist wieder elektrisch neutral. Du kannst mit einer Glimmlampe auch die Ladungsart feststellen: Wenn ein Körper negativ geladen ist, dann leuchtet die Glimmlampe auf der Seite auf, die den Körper berührt (▷ B2). Wenn ein Körper positiv geladen ist, dann leuchtet die Glimmlampe auf der anderen Seite auf.

Körper können elektrisch geladen werden, indem man sie aneinander reibt und wieder voneinander trennt. Dann ist ein Körper elektrisch positiv (+) geladen, der andere Körper ist elektrisch negativ (-) geladen.

AUFGABEN

- 1 ○ Beschreibe, wie Körper elektrisch geladen werden können.
- 2 ● Erkläre, wie man elektrisch geladene Körper nachweisen kann.
- 3 ● Wenn du einen Pullover ausziehst, stehen dir manchmal „die Haare zu Berge“. Erkläre, wie es dazu kommt.

Quelle:

252 Barmeier, M. (Hg.) (2013): Prisma Physik. Ausg. A, differenzierende Ausg., 1. Aufl., Dr. 1. Stuttgart: Klett, S. 252.

Woher kommen die Ladungen?

Das Atom

Alle Körper sind aus kleinsten Teilchen – den Atomen aufgebaut. Ein Atom ist unvorstellbar klein, sein Durchmesser beträgt nur etwa 0,000 000 1 mm.

Jedes Atom hat einen **Atomkern** und eine **Atomhülle** (► B 1). Im Atomkern befinden sich **Protonen** und **Neutronen**. Die Protonen sind positiv geladen. Die Neutronen sind elektrisch neutral.

Um den Atomkern bewegen sich negativ geladene Elektronen. Sie bilden die Atomhülle.

Neutrale Atome

Wenn es gleich viele Protonen wie Elektronen im Atom gibt, dann gleichen sich positive und negative Ladungen aus. Das Atom ist elektrisch neutral.

Die Protonen im Atomkern können nicht verschoben werden. Aus der Atomhülle können jedoch einzelne Elektronen entfernt werden. Umgekehrt ist es auch möglich, der Hülle Elektronen hinzuzufügen. In diesen Fällen ist das Atom nicht mehr elektrisch neutral.

Was passiert beim Aufladen?

Wenn du beispielsweise einen Kunststoffstab mit einem Tuch reibst, kommen beide

Körper in sehr engen Kontakt. Dabei gehen einige Elektronen vom Tuch auf den Stab über (► B 2).

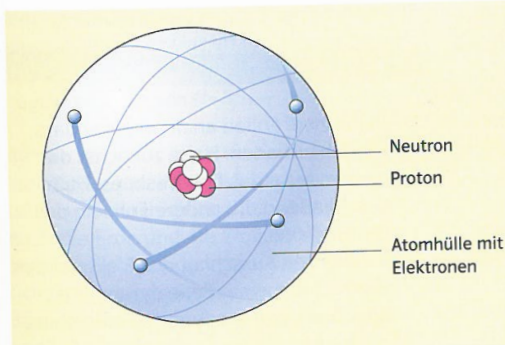
Das Tuch ist positiv geladen, weil es Elektronen abgegeben hat. Im Gegensatz dazu ist der Kunststoffstab negativ geladen, denn er hat Elektronen aufgenommen.

Ein Atom besteht aus Neutronen (elektrisch neutral), aus Protonen (positiv geladen) und aus Elektronen (negativ geladen).

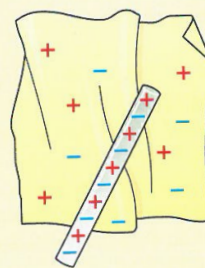
Beim Aufladen eines Körpers gehen Elektronen von einem Körper auf den anderen Körper über.

AUFGABEN

- 1 ○ Beschreibe den Aufbau eines Atoms.
- 2 ● Mara meint: „Alle Körper besitzen elektrische Ladungen.“ Beurteile ihre Aussage.
- 3 ● An eine positiv geladene Kugel wird eine negativ geladene Kugel gehalten. Erkläre, was passiert, wenn beide Kugeln sich berühren. Fertige eine Skizze dazu an.



1 Vereinfachtes Kern-Hülle-Modell



vor dem Reiben



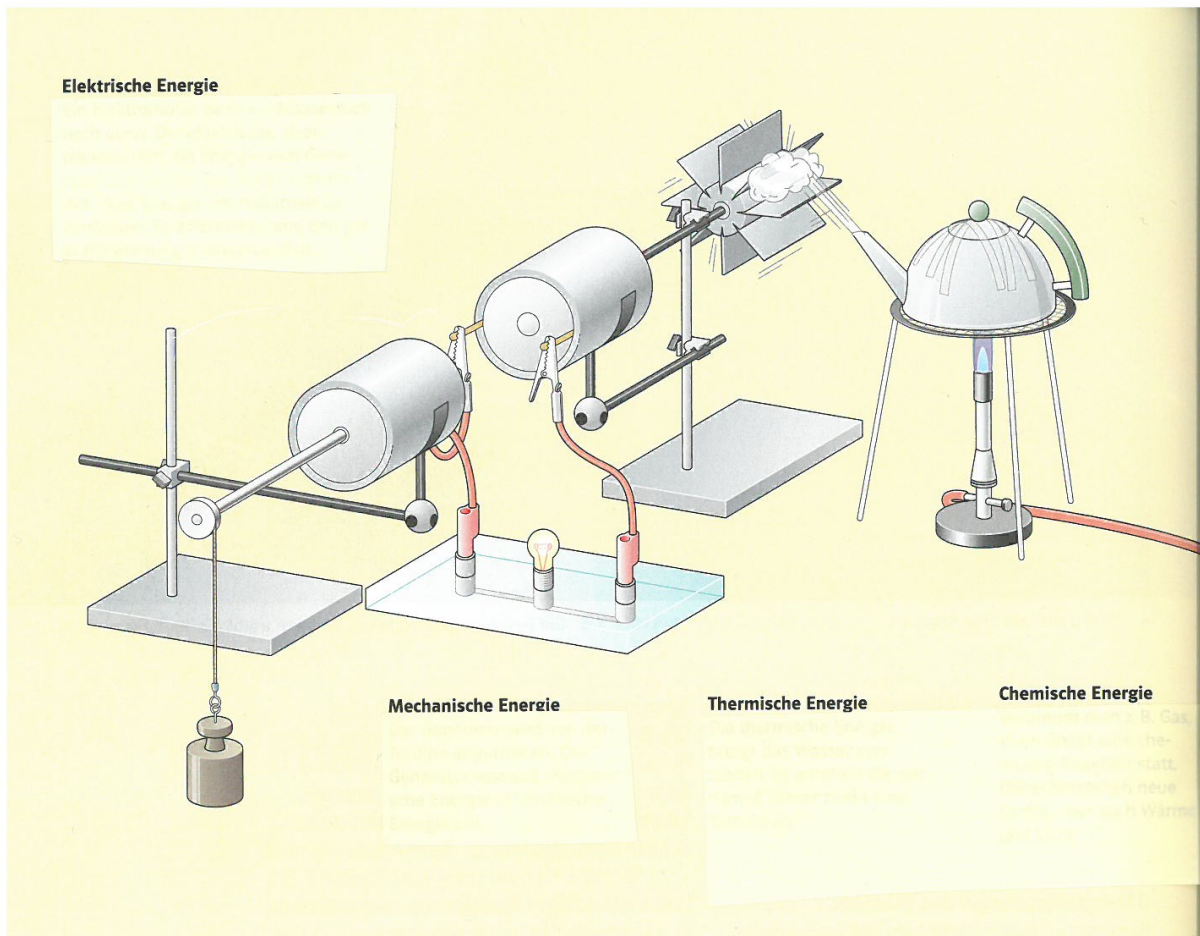
nach dem Reiben

2 Der Übergang von Elektronen führt zur Aufladung.

Quelle:

Barmeier, M. (Hg.) (2013): Prisma Physik. Ausg. A, differenzierende Ausg., 1. Aufl., Dr. 1. Stuttgart: Klett, S. 255.

A.3 Drehbuch 4



1 Verschiedene Energieformen

Energieformen – Umwandlungen

Verschiedene Energieformen

Energie kommt in verschiedenen Formen vor. Im Bild 1 kannst du einige erkennen: Chemische Energie, thermische Energie, mechanische Energie und elektrische Energie. Im Bild 3 sind es Höhenenergie und Bewegungsenergie.

Mit Schwung nach oben

Inlineskaten und Skateboardfahren auf der Half-Pipe ist bei Jugendlichen sehr beliebt (> B 3). Immer wieder wird auf der Bahn hin und her geskatet und die unterschiedlichsten Figuren werden dabei ausprobiert.

Was hat das mit Energie zu tun?

Betrachte einen Skater auf seiner Bahn von links nach rechts. Er kommt beim ersten Lauf fast auf dieselbe Höhe. Dabei sind zwei Energieformen beteiligt: Höhenenergie und Bewegungsenergie. Dies nennt man eine **Energieumwandlung**.

Energieerhaltung

Auf der Half-Pipe wandeln sich Höhenenergie und Bewegungsenergie ineinander um. Auch andere Energieformen können sich ineinander umwandeln (> B 2). Bei allen Energieumwandlungsprozessen wird ein

170 Quelle:

Barmeier, M. (Hg.) (2013): Prisma Physik. Ausg. A, differenzierende Ausg., 1. Aufl., Dr. 1. Stuttgart: Klett, S. 170.



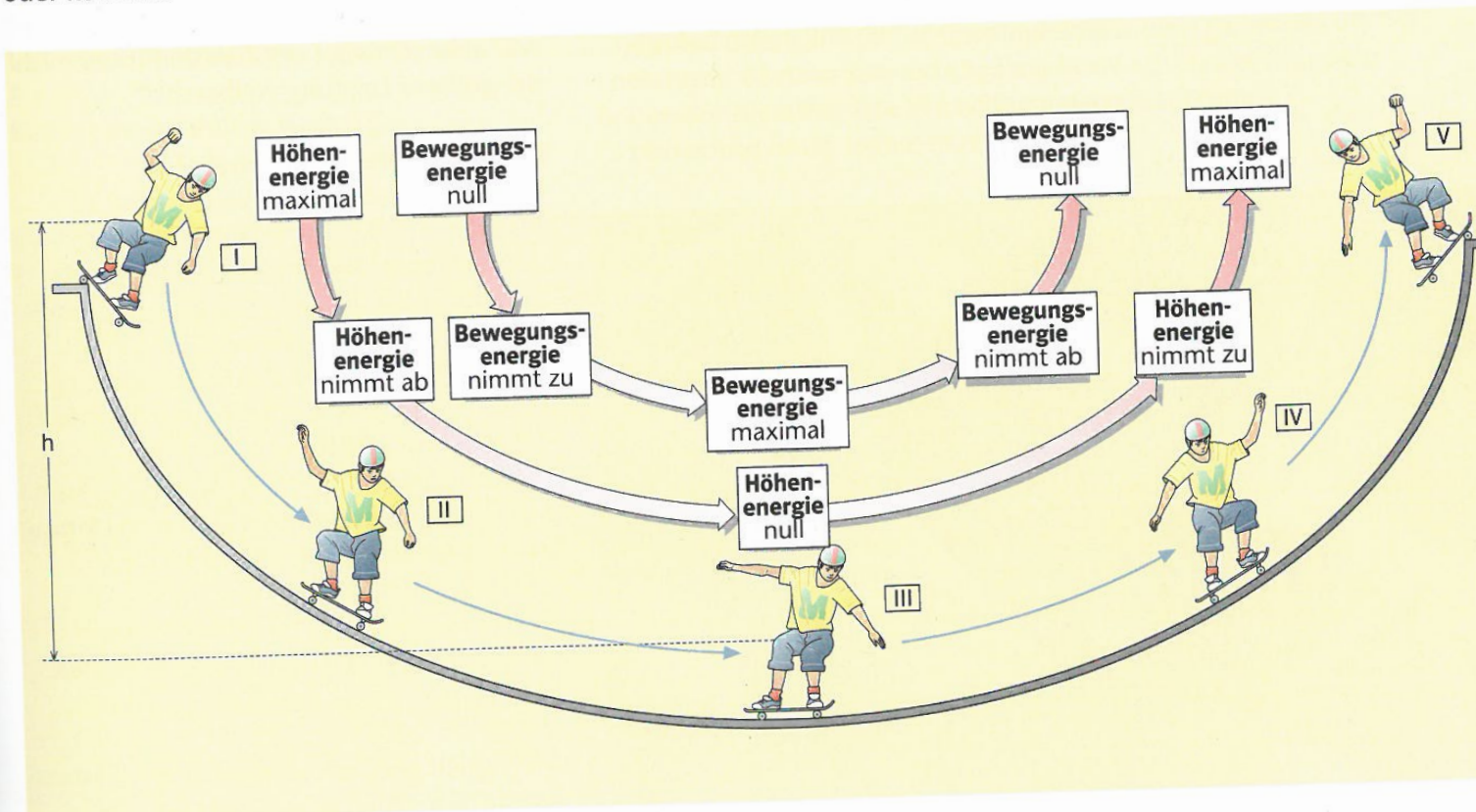
2 Eine Autobatterie wandelt chemische Energie in elektrische Energie um.

Teil der Energie in Wärme umgewandelt. Es wird also keine Energie „verbraucht“. Es gilt der **Energieerhaltungssatz**. Es wurde allerdings Energie in Formen (hier Wärme) umgewandelt, die wir nicht nutzen wollen oder können.

Es gibt verschiedene Energieformen. Energie kann von einer Form in eine andere umgewandelt werden. Energieumwandlung ist immer mit Verrichtung von Arbeit verbunden.

AUFGABEN

- 1 ○ Zähle verschiedene Energieformen auf.
- 2 ○ Beschreibe die Energieumwandlungen auf einer Half-Pipe.
- 3 ● Bild 1 zeigt das Modell eines Wärmekraftwerks. Erstelle die Energieumwandlungskette eines Wärmekraftwerks.
- 4 ● Beim Inlineskaten auf der Half-Pipe wandeln sich Höhenenergie und Bewegungsenergie ineinander um. Würde sich der Skater am oberen Rand nicht immer leicht abstoßen, würde er nicht mehr ganz den oberen Rand erreichen. Erkläre diesen Sachverhalt.
- 5 ● Vergleiche und bewerte die verschiedenen Energieformen miteinander.
- 6 ● Recherchiere, wie Erdöl, Erdgas und Kohle entstanden sind.



3 Energieumwandlung beim Skaten

Quelle:

Barmeier, M. (Hg.) (2013): Prisma Physik. Ausg. A, differenzierende Ausg., 1. Aufl., Dr. 1. Stuttgart: Klett, S. 171.

[illegible]

Fragen zur ökologischen Validität:

1. In dieser Unterrichtssequenz ist **die Interaktion zwischen den Schüler_innen und der Lehrkraft** authentisch für den Physikunterricht.

- ☐ Stimme zu.
- ☐ Stimme nicht zu.

Bitte begründen Sie Ihre Antwort:

A full-page sheet of white graph paper featuring a uniform grid of thin, light gray horizontal and vertical lines. The grid consists of small squares covering the entire area of the page.

2. In dieser Unterrichtssequenz ist **die Interaktion zwischen den Schüler_innen** authentisch für den Physikunterricht.

- ☐ Stimme zu.
- ☐ Stimme nicht zu.

Bitte begründen Sie Ihre Antwort:

A full page of blank graph paper with a uniform grid of small squares. The grid consists of 20 columns and 20 rows, creating a total of 400 small squares. The lines are thin and gray, set against a white background. There are no margins or additional markings on the page.

3. In dieser Unterrichtssequenz ist **die Sozialform** authentisch für den Physikunterricht.

- ☐ Stimme zu.
- ☐ Stimme nicht zu.

Bitte begründen Sie Ihre Antwort:

[illegible]

4. In dieser Unterrichtssequenz sind **die Arbeitsweisen der Schüler_innen** authentisch für den Physikunterricht.

- ☐ Stimme zu.
- ☐ Stimme nicht zu.

Bitte begründen Sie Ihre Antwort:

[illegible]

5. In dieser Unterrichtssequenz sind **die Unterrichtsmethoden** authentisch für den Physikunterricht.

- ☐ Stimme zu.
- ☐ Stimme nicht zu.

Bitte begründen Sie Ihre Antwort:

[illegible]

6. Diese Unterrichtssequenz könnte im Physikunterricht einer allgemeinbildenden Schule so tatsächlich stattfinden.

- ☐ Stimme zu.
- ☐ Stimme nicht zu.

Bitte begründen Sie Ihre Antwort:

[illegible]

C Entwicklungsstudie II – Interviewleitfaden

Leitfaden Interviews in der Pilotierung

Mai 2017

Informationen über die Interviewdurchführung:

Anonymer Code (3 Vorname Mutter, Körpergröße, Zahl)	
Versuchsdauer	
Datum	
Ort	
Studiensemester: Studiengang: Fächer: Muttersprache: Noten Physik: Noten Erzwiss.:	
Besondere Vorkommnisse	

Materialien

Sie benötigen folgende Materialien:

- Leitfaden
 - Einführung
 - Codeerstellung
 - Information über Think aloud
 - Übung zum Think aloud
 - Aufgabe
 - Hinweise für Interviewenden
 - Fragen zum Feedback
 - Zusatzmaterial
- Diktiergerät
- Videos (auf einem Rechner)

Einführung

Zuerst möchte ich mich bei Ihnen für Ihre Bereitschaft bedanken, bei dieser Untersuchung mitzumachen. Ihre Angaben und Daten aus der Untersuchung werden selbstverständlich absolut anonym gehalten. Hierzu generieren wir zunächst Ihren anonymen Code:

- Notieren Sie hierzu:
 - Die ersten drei Buchstaben des Vornamens der Mutter
 - Die Größe des Probanden in cm
 - Eine Beliebige Zahl zwischen 0 und 9

_ _ _ / _ _ _ / _

- Tragen Sie diesen Code auf der zweiten Seite dieses Leitfadens ein.
- Starten Sie die Tonbandaufzeichnung. Sagen Sie zu Beginn den anonymen Code und das Datum.
- Information zu Think aloud:

Lesen Sie dem Probanden folgende Instruktion vor:

*Ziel der Studie ist es, möglichst gut zu verstehen, wie (angehende) Lehrkräfte Unterrichtssituationen wahrnehmen. Es geht **nicht** darum, Ihre Wahrnehmung zu kontrol-*

lieren. Vielmehr geht es darum, mehr über die Begründungen für die Wahrnehmung herauszufinden. Die Studie ist daher so angelegt, dass Sie sich möglichst wie eine Lehrkraft in einer Unterrichtsstunde fühlen. Versuchen Sie die Rolle der Lehrkraft einzunehmen. Wir verwenden die sogenannte Think Aloud Methode. Das bedeutet: Während Sie das Video sehen, sollen Sie *all* das laut äußern, was Sie gerade denken oder machen. **Stellen Sie sich am besten vor, dass Sie alleine im Raum sind und mit sich selbst sprechen.** Das ist nicht ganz einfach, weil man sehr viel schneller denkt als man verbalisieren kann. Deshalb werde ich Ihnen bevor wir mit der eigentlichen Untersuchung beginnen eine Übungsaufgabe geben. Haben Sie bis hierher Fragen?

- Übung zu Think aloud:
Legen Sie die Aufgabenstellung verdeckt vor die Lehrkraft, nehmen Sie Ihren Platz hinter dem Probanden ein. Lesen Sie dem Probanden die Aufgabe vor:

Verbesserung technischer Geräte:

Erfinden Sie **fünf** Verbesserungen für Ihr Handy.

Auf dem Zettel steht noch einmal die Aufgabenstellung. Sie können nun mit der Aufgabe beginnen.

- Hinweise nach der Übung:
 - Es soll alles laut gesagt werden, was man denkt und macht.
 - Es sollte gut hörbar und deutlich gesprochen werden.
 - Lieber zu viel als zu wenig
 - Es sollte jede noch so kleine Tätigkeit kommentiert werden.
 - Stellen Sie sich vor, ich bin gar nicht anwesend und Sie führen ein Selbstgespräch.

Durchführung Untersuchung:

Video 1

- Einleitung:
 - Legen Sie die Aufgaben und Zusatzmaterialien verdeckt vor den Probanden, setzen Sie sich wieder außer Sichtweite und führen Sie in die Untersuchung ein:
 - *Bitte sehen Sie sich gleich das erste Video. Denken Sie daran die Perspektive der Lehrkraft einzunehmen; Immer wenn Schüler_innen zu sehen sind, sieht die Lehrkraft die Schüler_innen auch. Beantworten Sie folgende Aufgaben, indem Sie die ganze Zeit lang laut denken.*
 - 1. *Pausieren Sie das Video, wenn Sie etwas in Bezug auf:*
 - Vorstellungen **über** Fachinhalte der Schüler_innen wahrnehmen
 - Erkenntnisgewinnung der Schüler_innen wahrnehmen
 - Vorstellungen **über** Erkenntnisgewinnung der Schüler_innen wahrnehmen
 - 2. *Begründen Sie, warum Ihnen diese Situation aufgefallen ist.*
 - 3. *Stellen Sie sich vor, Sie stehen neben den agierenden Schüler_innen. Wie würden Sie handeln?*
 - *Sie haben ein Blatt mit Zusatzmaterial vor sich liegen, hier finden Sie Informationen zu Schülervorstellungen und Erkenntnisgewinnung. Gibt es Fragen zu den obenstehenden Aufgaben?*
 - *Beginnen Sie **jetzt** mit der Bearbeitung der Aufgabe und geben Sie ein Zeichen, wenn Sie meinen die Bearbeitung beendet zu haben.*
- Durchführung
- Abschluss:
 - *Bitte nennen Sie zwei Situationen, die für Sie herausstechend waren.*
 - *Vielen Dank, wir fahren jetzt mit dem zweiten Video fort.*

Video 3

- Einleitung:
 - Legen Sie die Aufgaben und Zusatzmaterialien verdeckt vor den Probanden, setzen Sie sich wieder außer Sichtweite und führen Sie in die Untersuchung ein:
 - *Bitte sehen Sie sich gleich das erste Video. Denken Sie daran die Perspektive der Lehrkraft einzunehmen; Immer wenn Schüler_innen zu sehen sind, sieht die Lehrkraft die Schüler_innen auch. Beantworten Sie folgende Aufgaben, indem Sie die ganze Zeit lang laut denken.*
 - 1. *Pausieren Sie das Video, wenn Sie etwas in Bezug auf:*
 - Vorstellungen **über** Fachinhalte der Schüler_innen wahrnehmen
 - Erkenntnisgewinnung der Schüler_innen wahrnehmen
 - Vorstellungen **über** Erkenntnisgewinnung der Schüler_innen wahrnehmen
 - 2. *Begründen Sie, warum Ihnen diese Situation aufgefallen ist.*
 - 3. *Stellen Sie sich vor, Sie stehen neben den agierenden Schüler_innen. Wie würden Sie handeln?*
 - *Sie haben ein Blatt mit Zusatzmaterial vor sich liegen, hier finden Sie Informationen zu Schülervorstellungen und Erkenntnisgewinnung. Sie finden außerdem das Aufgabenblatt der Schülerinnen. Gibt es Fragen zu den obenstehenden Aufgaben?*
 - *Beginnen Sie **jetzt** mit der Bearbeitung der Aufgabe und geben Sie ein Zeichen, wenn Sie meinen die Bearbeitung beendet zu haben.*
- Durchführung
 - Abschluss:
 - *Bitte nennen Sie zwei Situationen, die für Sie herausstechend waren.*
 - *Vielen Dank für die Teilnahme an der Untersuchung, ich überreiche Ihnen nun Ihren Gutschein.*

Video 4

- Einleitung:
 - Legen Sie die Aufgaben und Zusatzmaterialien verdeckt vor den Probanden, setzen Sie sich wieder außer Sichtweite und führen Sie in die Untersuchung ein:
 - *Bitte sehen Sie sich gleich das erste Video. Denken Sie daran die Perspektive der Lehrkraft einzunehmen; Immer wenn Schüler_innen zu sehen sind, sieht die Lehrkraft die Schüler_innen auch. Beantworten Sie folgende Aufgaben, indem Sie die ganze Zeit lang laut denken.*
- 4. *Pausieren Sie das Video, wenn Sie etwas in Bezug auf:*
 - *Vorstellungen **über** Fachinhalte der Schüler_innen wahrnehmen*
 - *Erkenntnisgewinnung der Schüler_innen wahrnehmen*
 - *Vorstellungen **über** Erkenntnisgewinnung der Schüler_innen wahrnehmen*
- 5. *Begründen Sie, warum Ihnen diese Situation aufgefallen ist.*
- 6. *Stellen Sie sich vor, Sie stehen neben den agierenden Schüler_innen. Wie würden Sie handeln?*
 - *Sie haben ein Blatt mit Zusatzmaterial vor sich liegen, hier finden Sie Informationen zu Schülervorstellungen und Erkenntnisgewinnung. Sie finden außerdem das Aufgabenblatt der Schülerinnen. Gibt es Fragen zu den obenstehenden Aufgaben?*
 - *Beginnen Sie **jetzt** mit der Bearbeitung der Aufgabe und geben Sie ein Zeichen, wenn Sie meinen die Bearbeitung beendet zu haben.*
- Durchführung
- Abschluss:
 - *Bitte nennen Sie zwei Situationen, die für Sie herausstechend waren.*

Vielen Dank für die Teilnahme an der Untersuchung, ich überreiche Ihnen nun Ihren Gutschein.

Zusatzmaterial Video 1, 3, 4:

1) Definition Schülervorstellungen:

Schülervorstellungen werden als mentale Konstrukte verstanden, die auf bisherige – nicht zwingend schulischen – (Lern-)Erfahrungen zurückgehen. Sie sind dem Bewusstsein des Lernenden zugänglich und beziehen sich auf spezifische Themenfelder des Faches, wie zum Beispiel die Mechanik, den Sehvorgang oder chemische Reaktionen. (nach Krüger 2016)

2) Was heißt „Erkenntnisgewinnung“?

Die Schülerinnen und Schüler

- beschreiben Phänomene und führen sie auf bekannte physikalische Zusammenhänge zurück,
- wählen Daten und Informationen aus verschiedenen Quellen zur Bearbeitung von Aufgaben und Problemen aus, prüfen sie auf Relevanz und ordnen sie,
- verwenden Analogien und Modellvorstellungen zur Wissensgenerierung,
- wenden einfache Formen der Mathematisierung an, nehmen einfache Idealisierungen vor,
- stellen an einfachen Beispielen Hypothesen auf,
- führen einfache Experimente nach Anleitung durch und werten sie aus,
- planen einfache Experimente, führen sie durch und dokumentieren die Ergebnisse,
- werten gewonnene Daten aus, ggf. auch durch einfache Mathematisierungen,
- beurteilen die Gültigkeit empirischer Ergebnisse und deren Verallgemeinerung

D Entwicklungsstudie II – Distraktoren

Video 1, 3 und 4

Distraktoren für Video 1

Situation	Nr.	Distraktor
1	1	Die Schülerin äußert die Vorstellung, dass der Tee schmilzt.
	2	Die Schülerin stellt erfolgreich eine erste testbare Hypothese auf.
	3	Die Schülerin und der Schüler hören einander aufmerksam zu.
	4	Es meldet sich fast die ganze Klasse, alle sind aktiviert und motiviert.
	5	Es tragen zwei Schülerinnen Kopftuch, eventuell muss man hier einen Schwerpunkt auf sprachliche Aspekte legen.
2	1	Die Schülerin versteht noch nicht, was eine Hypothese ist.
	2	Die Schülerin und der Schüler sind sich einig, dass sich Teeteilchen im heißen Wasser schneller bewegen.
	3	Die beiden Schülerinnen hinten rechts arbeiten nicht.
	4	Die Schülerin und der Schüler tauschen sich in der Partnerarbeit über Vorstellungen zum Phänomen aus.
	5	Die Schülerin und der Schüler sprechen kaum fachsprachlich über das Phänomen.
3	1	Die Schülerin stellt sich vor, dass sich Tee besser löst, wenn sich heißes Wasser stark bewegt.
	2	Die Schülerin nutzt ein Modell, um das Phänomen zu verstehen.
	3	Der Schüler könnten sich aus dem Schulbuch informieren. Schulbücher müssen hier vom Tisch.
	4	Die meisten Schülerinnen und Schüler im Klassenraum arbeiten in dieser Unterrichtsphase gewissenhaft mit.
	5	Die Schülerin beschreibt und interpretiert das Phänomen nicht fachsprachlich genug.
4	1	Die Schülerin stellt sich vor, dass Tee sich auch im kalten Wasser löst.
	2	Die Schülerin meint, dass man bei der Hypothesenbildung notiert, was im Experiment passiert ist.
	3	Die Schülerinnen im Hintergrund spielen auf ihrem College Block ein Spiel.
	4	Die Schülerin und der Schüler unterstützen sich beim Lernen.
	5	Es sind viele Schülerin und Schüler mit Migrationshintergrund zu sehen, das kann sprachliche Probleme bedeuten.
5	1	Der Schüler stellt sich vor, dass der Tee im warmen Wasser besser schmilzt.
	2	Der Schüler stellt eine eigene Hypothese auf.
	3	Die Schülerin und der Schüler kooperieren nicht genügend miteinander.

	4	Die Schülerin und der Schüler arbeiten in der Partnerarbeit nicht an einem gemeinsamen Ergebnis.
	5	Der Schüler kann seine Gedanken sofort verschriftlichen.
6	1	Die Schülerin und der Schüler bleiben bei ihrer problematischen Vorstellung.
	2	Die Schülerinnen und Schüler hätten zur Durchführung des Experimentes weitere Materialien wie Stoppuhren und Thermometer benötigt.
	3	Die Schülerin und der Schüler hätten für die fünf Minuten Wartezeit eine Aufgabe erhalten müssen.
	4	Die Schülerin und der Schüler organisieren ihren Arbeitsplatz zweckgemäß, so dass sie zügig arbeiten können.
	5	Zwischen der Schülerin und dem Schüler gibt es eine klare Rollenverteilung.
7	1	Der Schüler wiederholt seine Annahme, dass Tee schmilzt.
	2	Der Schüler hat eine Idee das Experiment zu verbessern / Der Schüler optimiert das Experiment, indem er einen Löffel verwenden möchte.
	3	Viele Schülergruppen arbeiten an ihrem Experiment und sind aktiviert.
	4	Den Schülerinnen und Schülern fehlen die Materialien zum Umrühren.
	5	Die Schülerin und der Schüler können über das Vorgehen beim Experimentieren diskutieren.
8	1	Die Schülerin stellt sich vor, dass sich im warmen Wasser mehr Farbstoffe gelöst haben.
	2	Die Schülerin äußert Beobachtungen zu ihrem Experiment und wertet diese aus.
	3	Die Schülerin und der Schüler sind am Ende der Arbeitsphase immer noch diszipliniert.
	4	Der Schüler im Hintergrund hat seinen Kopf auf dem Tisch liegen und arbeitet nicht.
	5	Der Schüler nutzt den Vergleich mit Schokolade, um sich sprachlich zu entlasten.
9	1	Der Schüler nimmt nach dem Experiment doch an, dass der Tee sich löst.
	2	Die Schülerin und der Schüler haben ihre falschen Hypothesen bestätigt und verallgemeinert, aber nicht adäquat begründet.
	3	Die Schülerinnen im Hintergrund beschäftigen sich nicht mit ihrem Experiment.

4	Der Schüler arbeitet besonders zügig, indem er einen Text formuliert und zugleich aufschreibt.
5	Die Schülerin diskutiert nicht mehr mit ihrem Sitznachbarn.

Distraktoren für Video 3

Situation	Nr.	Distraktor
1	1	Der Schüler erklärt Knistern mit Reibungselektrizität.
	2	Der Schüler beschreibt eine Vermutung über das anstehende Phänomen mit Hilfe eines Vergleichs.
	3	Es melden sich nur wenige Lernende, die Aktivität ist nicht sehr hoch.
	4	Die Schülerinnen in der letzten Reihe sind nicht bei der Sache, sie sehen gelangweilt aus.
	5	Der Schüler redet frei und selbstbewusst.
2	1	Der Schüler bezeichnet das Phänomen als Kleben.
	2	Der Schüler stellt einen Bezug zu einer Alltagserfahrung her.
	3	Die Schülerinnen beteiligen sich nicht am Unterricht.
	4	Ein Schüler kommentiert unaufgefordert die Äußerung eines anderen Schülers.
	5	Die männlichen Schüler dominieren den Unterricht.
3	1	Die Schülerin stellt sich vor, dass der Luftballon durch Reiben klebrig geworden ist.
	2	Die Schülerin gibt eine Vermutung ohne Erklärung.
	3	Einer Schülerin wurde sofort nach der Frage des Lehrers das Wort erteilt, ohne dass sie sich gemeldet hat.
	4	Einige Lernende sind nicht aktiviert, obwohl das Experiment eindrucksvoll war.
	5	Die Schülerin drückt ihre Vermutung in Alltagssprache aus.
4	1	Der Schüler meint, dass der Ballon die Haare wegen eines elektrischen Feldes anzieht.
	2	Der Schüler benutzt Informationen aus anderen Medien zur Erklärung.
	3	Im Hintergrund schläft ein Schüler.
	4	Der Schüler unterbricht ungefragt die Aussage einer Mitschülerin.
	5	Der Schüler verwendet Fachbegriffe wie Ladung.
5	1	Der Schüler benennt Knistern als Ursache für das Blitzen.
	2	Der Schüler beschreibt seine Beobachtungen zum Versuch.
	3	Eine Schülerin ist frustriert, weil ihr nicht das Wort erteilt wurde.
	4	Ein Schüler lässt seinen Blick in die Ferne schweifen.
	5	Der Schüler drückt sich sprachlich nicht angemessen genug aus.
6	1	Der Schüler unterstellt dem Blitz, er wolle sich bewegen.
	2	Der Schüler ergänzt eine weitere Beobachtung.
	3	Weitere Lernende werden durch Rückfragen motiviert.
	4	Der Großteil der Lernenden ist passiv.
	5	Der Schüler drückt mit seiner Wortwahl eine Unsicherheit aus.

7	1	Der Schüler erklärt Blitze mit Temperaturunterschieden.
	2	Der Schüler vergleicht das Experiment mit der Entstehung von Gewitter.
	3	Ein Schüler hat keinen Partner und schaut orientierungslos durch die Klasse.
	4	Fast alle Lernende diskutieren angeregt in Partnerarbeit.
	5	Der Schüler signalisiert mit seinem sprachlichen Ausdruck Unsicherheit.
8	1	Die Schülerin stellt sich vor, dass man von elektronischer Diebstalsicherungen im Kaufhaus einen elektrischen Schlag bekommt.
	2	Die Schülerin bringt ein Beispiel aus ihrem Alltag in die Diskussion ein, um das Phänomen aus einer anderen Perspektive zu erläutern.
	3	Der Schüler im Hintergrund richtet seine Aufmerksamkeit nur auf das Gespräch vor ihm und spricht nicht mit seiner Sitznachbarin.
	4	Der Schüler in der zweiten Reihe sitzt in einem Rollstuhl, seine Möglichkeiten in allen Unterrichtssituationen teilzunehmen sind eingeschränkt.
	5	Die Schülerin kann durch den Gebrauch von Alltagssprache ihre Erfahrungen mit dem Phänomen berichten.
9	1	Der Schüler stellt sich vor, dass Blitze etwas mit der Kälte zu tun haben.
	2	Die Schülerin und der Schüler planen zur Überprüfung ihrer Hypothese einen eigenen Versuch.
	3	Die Schülerin und der Schüler unterhalten sich gerne über das Thema.
	4	Die Schülerin und der Schüler verlassen unaufgefordert ihren Arbeitsplatz.
	5	Der Schüler hat Schwierigkeiten seine Gedanken in der Unterrichtssprache zu formulieren.
10	1	Der Schüler sagt, dass die Kugel ein Modell für die Ladung ist.
	2	Der Schüler zieht Erkenntnisse über Ladungen aus einer Fernsehsendung.
	3	Dieses Schülerpaar ist sehr leistungsheterogen.
	4	Die Schülerinnen im Hintergrund sind nicht bei der Sache.
	5	Ein Schüler mit Migrationshintergrund hat fachsprachliche Defizite.
11	1	Der Schüler äußert seine Vorstellung, dass man mit kälteren Objekten kleinere Blitze erzeugt.
	2	Der Schüler nutzt sein Erfahrungswissen, um das Phänomen besser zu verstehen.
	3	Im Hintergrund ist ein Schüler unaufmerksam.

	4	Nur Lernende in der ersten Reihe melden sich.
	5	Der Schüler hat Schwierigkeiten mit der Verwendung fachlicher Begriffe.
12	1	Der Schüler meint, dass Ladungen bei Blitzen aufeinanderstoßen und das Phänomen nichts mit Kälte zu tun hat.
	2	Die Lernenden im Hintergrund haben dem sprechenden Schüler nicht zugehört.
	3	Der Schüler beschreibt seine Beobachtungen des Phänomens.
	4	Der Schüler hantiert mit einer Schere, während er spricht.
	5	Der Schüler geht mit seiner Antwort inhaltlich auf die Aussage seines Mitschülers ein.
13	1	Die Schülerin stellt sich vor, dass die Folienstreifen immer verschieden geladen sind.
	2	Die Schülerin und der Schüler wissen noch nicht warum die Streifen verschieden geladen sind, lassen das aber als Fakt in ihre Hypothese einfließen.
	3	Die Schülerin und der Schüler praktizieren eine sinnvolle Arbeitsaufteilung.
	4	Der Schüler ohne Partner hat in dieser Phase wieder einen Nachteil.
	5	Die Schulbuchaufgabe bereitet der Schülerin Schwierigkeiten beim Lesen.
14	1	Die Schüler stellen sich vor, dass die Folienstreifen immer unterschiedliche Ladungen tragen.
	2	Die Schüler entwickeln aus eigenem Interesse eine neue Hypothese und planen dazu ein Experiment.
	3	Die Schüler bereichern sich in der Partnerarbeit gegenseitig.
	4	Die Schüler leihen sich Versuchsmaterial von anderen Schülern, das ist in Ordnung, wenn diese sie nicht mehr benötigen.
	5	Der Schüler drückt mit seiner Sprechweise große Sicherheit mit dem Thema der Ladung aus.
15	1	Die Schüler glauben, dass es nicht wichtig ist, welcher Streifen mit welchem kombiniert wird.
	2	Die Schüler führen ihr eigenständig geplantes Experiment durch, machen dabei aber einen Fehler.
	3	Im Hintergrund sitzen Schülerinnen, die nicht arbeiten.
	4	Durch die kreative Aktivität der beiden Schüler, wurden auch andere Lernende motiviert.
	5	Die Schüler verwenden nach dieser Arbeitsphase zum Teil fachsprachliche Aspekte.

Distraktoren für Video 4

Situation	Nr.	Distraktor
1	1	Der Schüler erklärt sein Verständnis des Begriffs Energieumwandlung mit der Umwandlung von Energie.
	2	Der Schüler stellt eine auf Alltagserfahrungen basierende Hypothese auf.
	3	Alle Lernenden folgen dem Unterrichtsgespräch.
	4	Die Lernenden haben ausreichend Zeit sich Gedanken zu machen und sich zu melden.
	5	Der Schüler reichert seine Alltagssprache mit Fachsprache an.
2	1	Der Schüler nennt ein Beispiel, in dem sich ein fachlich angemessenes Verständnis von Energieumwandlung anbahnt.
	2	Die Lernenden erklären physikalische Zusammenhänge, indem sie bereits gewonnene Erkenntnis damit in Verbindung bringen.
	3	Ein Schüler ergänzt die Aussage eines anderen Schülers unaufgefordert.
	4	Alle Lernenden folgen dem Unterrichtsgespräch.
	5	Der Schüler formuliert sprachlich einfache Sätze.
3	1	Die Schülerin nimmt an, dass Energie überall im Alltag verbraucht wird.
	2	Die Schülerin erklärt Energieumwandlung mit Hilfe von Alltagsbeispielen.
	3	Die anderen Lernenden wenden sich der sprechenden Schülerin zu.
	4	Viele Lernende kommen zu Wort.
	5	Die Schülerin geht nicht auf die Aussage ihres Mitschülers ein.
4	1	Die Schülerin stellt sich vor, dass Energie auf Objekte übertragen werden kann.
	2	Die Schülerin schätzt die Relevanz der Informationen des Textes ein.
	3	Es meldet sich nur eine Schülerin nach der Einzelarbeitsphase.
	4	Die Lernenden hatten wenig Zeit ihre Gedanken zur Beantwortung der Frage zu sortieren.
	5	Die Schülerin formuliert frei und in vollständigen Sätzen.
5	1	Der Schüler hat die Vorstellung, dass Energie umgewandelt wird.
	2	Der Schüler bewertet die Gültigkeit der Informationen aus dem Text.
	3	Der Schüler äußert sich ungefragt.
	4	Die Lernenden hören zu.
	5	Der Schüler verwendet Energie-Begriffe, die er nicht aus dem Text kennt.

6	1	Die Schülerin fasst zusammen, dass Energie nach der Umwandlung manchmal weg ist.
	2	Die Schülerin ordnet die Informationen aus dem Text.
	3	Die Schülerinnen aus der ersten Reihe sind weiterhin auf den Text konzentriert.
	4	Die Lernenden hatten nicht genug Zeit ihre Gedanken zu sortieren.
	5	Die Schülerin kann sich sprachlich problemlos ausdrücken.
7	1	Der Schüler macht die Erfahrung, dass sein bisheriges Konzept nicht ausreicht.
	2	Der Schüler stellt eine Analogie her.
	3	Der Schüler darf seine Aussage nicht als Erster tätigen.
	4	Der Schüler im Hintergrund passt nicht auf.
	5	Der Schüler drückt sich adäquat aus.
8	1	Die Schülerin links fragt sich, ob die Energie am Ende der Kette verloren ist.
	2	Die Schülerin stellt die Hypothese auf, dass die Energie nach dem Anschalten des Fernsehers weg ist.
	3	Die Schülerinnen sollten ihre Ideen auch schriftlich festhalten.
	4	Die Schülerin und der Schüler im Hintergrund arbeiten nicht zusammen.
	5	Die Schülerinnen diskutieren angeregt.
9	1	Die Schülerin rechts stellt fachlich richtig klar, dass Energie nicht verbraucht, sondern umgewandelt wird.
	2	Die Schülerin rechts widerlegt die Hypothese ihrer Sitznachbarin.
	3	Die Schülerin und der Schüler im Hintergrund arbeiten nicht zusammen.
	4	Die Schülerinnen sollten ihre Ideen auch schriftlich festhalten.
	5	Die Schülerinnen diskutieren über den Sachverhalt.
11	1	Der Schüler stellt sich vor, dass Energie in Form von Kalorien verbraucht wird.
	2	Der Schüler verwendet Modellvorstellungen zur Wissensgenerierung.
	3	Die Schülerin und der Schüler nutzen ihr Arbeitsblatt für die Bearbeitung der Aufgabe.
	4	Die Schülerin und der Schüler notieren ihre Überlegungen zur Aufgabe nicht.
	5	Die Schülerin und der Schüler tauschen sich über ihr Verständnis aus.
12	1	Die Schülerin argumentiert mit dem Kraft-Begriff.
	2	Die Schülerin nimmt zur Verdeutlichung eine Idealisierung vor.
	3	Die Schülerin und der Schüler nutzen ihr Arbeitsblatt für die Bearbeitung der Aufgabe.

	4	Die Schülerin und der Schüler notieren ihre Überlegungen zur Aufgabe nicht.
	5	Die Schülerin und der Schüler tauschen sich über ihr Verständnis aus.
	1	Der Schüler kann Verbrauch und Umwandlung nicht konzeptionell voneinander trennen.
	2	Der Schüler formuliert auf Basis des Phänomens mit dem Skater eine Hypothese für das Phänomen des Flummis.
	3	Der Schüler Hardy freut sich, dass ihm das Wort erteilt wurde.
13	4	Der Schüler im Hintergrund guckt in der Klasse herum.
	5	Der Schüler verwendet Fachbegriffe, wie Höhenenergie und Bewegungsenergie.
	1	Der Schüler meint, dass Reibungskraft Wärme ist.
	2	Der Schüler verallgemeinert seine Erkenntnisse bezüglich des Flummis.
	3	Die Lernatmosphäre ist ruhig.
14	4	Die Lernenden präsentieren ihre Ergebnisse ausschließlich mündlich.
	5	Der Schüler verwendet seine Gestik zur Veranschaulichung seiner Ausführungen.

E Hauptstudie

E.1 Fragen zu demographischen Daten 1

Zu Ihrer Person

Geschlecht: ☐ Männlich ☐ Weiblich ☐

Alter: Jahre

Hochschulsemester:

Fachsemester: (Aktuelles Studium)

Zu Ihrem aktuellem Studium

Welchen Abschluss streben Sie in Ihrem **aktuellen Studium** an?

- ☐ Lehramt an Gymnasien ☐ Lehramt NUR an Sek I
☐ Lehramt NUR an Prima ☐ Lehramt an Prima UND Sek I
☐ Lehramt an Berufsschulen ☐ Lehramt für Sonderpädagogik
☐ B.Sc. Physik ☐ M.Sc. Physik
☐

Zu Ihre Fächern

Welche Fächer studieren Sie?

1.
2.
3.
4.

Zu Ihren bisherigen Abschlüssen

Welchen Abschluss haben Sie bereits vor dem aktuellen Studium abgeschlossen?

- ☐ Abitur: Durchschnittsnote ☐ Fachabitur:
Durchschnittsnote
☐ abgeschlossene Berufsausbildung:
☐ vorheriges Studium mit den Fächern:

E.2 Fragen zu demographischen Daten 2

Zu Ihren Lehr-Erfahrungen: Nachhilfe

Haben Sie Erfahrungen mit dem Unterrichten von Schüler_innen im Rahmen von **Nachhilfeunterricht**?

- ☐ Nein
☒ Ja

Dauer insgesamt:

- ☐ < 1 Jahr ☐ 1 - 5 Jahre ☐ > 5 Jahre

Mittlere Häufigkeit:

- ☐ 1-2 mal wöchentlich ☐ 3-4 mal wöchentlich
☐ 5-6 mal wöchentlich ☐ 7-8 mal wöchentlich

Zu Ihren Lehr-Erfahrungen: Unterrichtspraktikum mit Schul- oder Fachbezug

Haben Sie Erfahrungen mit dem Unterrichten von Schüler_innen im Rahmen von **Unterrichtspraktika** mit Schul- oder Fachbezug?

- ☐ Nein
☒ Ja

Anzahl der **Praktika**:

- ☐ 1 ☐ 2 ☐ 3 ☐ > 3

Zu Ihren Lehr-Erfahrungen: Lehrauftrag

Haben Sie Erfahrungen mit dem Unterrichten von Schüler_innen im Rahmen von **Lehraufträgen**?

- ☐ Nein
☒ Ja

Daure insgesamt:

- ☐ 1 Jahr ☐ 1 - 5 Jahre ☐ > 5 Jahre

Mittlere Häufigkeit:

- ☐ 1 - 10 Stunden wöchentlich ☐ 11 - 20 Stunden wöchentlich
☐ 21 - 30 Stunden wöchentlich

F Lister der digitalen Anlagen der vorliegenden Arbeit

Da in der vorliegenden Arbeit viel mit Excel-Tabellen und R- sowie SPSS-Skripten gearbeitet wurde, findet sich auf der beiliegenden DVD ein digitaler Anhang, der aus folgenden Elementen besteht:

Digitale Version der Dissertation

- Dissertationsschrift (Dateiformat: PDF).

Dateien und Dokument der Entwicklungsstudie I

- Die Drehbücher zu den Video 1, 3 und 4 (Dateiformat: PDF).
- Dateien zur Validierung der Drehbücher (Dateiformat: MX18; PDF).
- Die Videos 1, 3 und 4 (Dateiformat: MP4-Datei).

Dateien und Dokument der Entwicklungsstudie II

- Die Audiodateien der 22 Interviews (Dateiformat: MP3-Datei).
- Die Transkripte der 22 Interviews (Dateiformat: PDF).
- Dateien zur Kodierung der Transkripte (Dateiformat: MX18).

Dateien und Dokument der Hauptstudie

- Die Rohdaten der Erhebung (Dateiformat: CSV und PDF).
- Die Dateien zur Analyse der Daten:
 - Dateien zur Kodierung der freien Antworten (Dateiformat: CSV, DOCS, EXCEL).
 - Tabellen zu den verwendeten Daten (Dateiformat: CSV; Dateiname: Alle Variablen, Daten_PI, Date_PI151, Daten_PI151_19, Daten_PI_CM_29, Daten_fdw, Daten_fdw_151, Daten sonstige).

- Tabellen zu den verschiedenen Modellen und deren Kennwerten (Dateiformat: CSV oder EXCEL).
- Die Skripte und Workspaces der Analyse mit R-Studio (Dateiformat: R-Datei, R Workspace).
- Die Dateien der Analyse mit SPSS (Dateiformat: SPSS Statistics Data Document, SPSS Statistics Output Document).
- Die Plots der Analyse der Daten (Dateiformat: PNG).

G Lister der aus der Dissertation hervorgegangenen Publikationen

2020

- Wöhlke, C. & Höttecke, D. (2020). Professionelle Unterrichtswahrnehmung und fachdidaktisches Wissen. In C. Maurer (Hrsg.), Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Wien. o.V. (o.N.), o.S.

2019

- Wöhlke, C. & Höttecke, D. (2019). Development and Validation of an Instrument for Identifying Pre-Service Physics Teachers' Professional Vision. Paper presented at the ESERA conference 26st-30th August 2019 in Bologna.
- Wöhlke, C. & Höttecke, D. (2019). Ist Noticing valide messbar? Erste Befunde eines Videovignetentests. In: C. Maurer (Hrsg.), Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Kiel 2018. (S. 257). Universität Regensburg

2018

- Wöhlke, C. & Höttecke, D. (2018). Erfassung von Noticing von Physiklehrkräften – Instrumentenentwicklung. In: C. Maurer (Hrsg.), Qualitätsvoller Chemie- und Physikunterricht – normative und empirische Dimensionen. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Regensburg 2017. (S. 67). Universität Regensburg.

2017

- Wöhlke, C. & Höttecke, D. (2017). Development of an instrument for identifying pre-service physics teachers' noticing. Paper presented at the ESERA conference 21st-25th August 2017 in Dublin, keynote.conference-services.net/resources/444/5233/pdf/ESERA2017_0516_paper.pdf (14.09.2017).

Abbildungsverzeichnis

2.1	Modell der Professionellen Unterrichtswahrnehmung	13
2.2	Professionelle Kompetenzen von Lehrkräften	17
2.3	Das PID-Modell	18
2.4	Konzepte des Lehrerwissen nach Neuweg (2014)	20
2.5	Überarbeitetes PID-Modell	22
2.6	Situationen im Unterrichtsverlauf	23
2.7	Eigene Konzeptualisierung der professionellen Unterrichtswahrnehmung . .	26
4.1	Die drei Teilstudien des Forschungsdesigns	58
4.2	Beispiel für die Online-Oberfläche für das PUPiL-Instrument	59
4.3	Angebots-Nutzungs-Modell	62
4.4	Operationalisierter Prozess der professionellen Unterrichtswahrnehmung . .	66
4.5	Operationalisierung des Konzepts der professionellen Unterrichtswahrnehmung	67
5.1	Energieumwandlungskette	80
6.1	Ausprägungsstufen der professionellen Unterrichtswahrnehmung	107
6.2	Ergebnisse Prüfung der Passung und Schwierigkeit des Tests – Entwicklungsstudie II.1	109
6.3	Verteilung der Aussagen auf die Kategorien nach Testpersonen	118
6.4	Verteilung der Aussagen auf die Kategorien	120
6.5	Ergebnisse Prüfung der Passung und Schwierigkeit des Tests Entwicklungsstudie II.2	124
6.6	Verteilung der Aussagen auf die Kategorien	134
6.7	Verteilung der Aussagen auf die Kategorien nach Testpersonen	135
6.8	Vergleich der beiden Teilstudien bezüglich Verteilung der Aussagen auf die Kategorien	137
7.1	Interface bei der Übung zum Umgang mit den Videos I	146
7.2	Interface bei der Übung zum Umgang mit den Videos II	148
7.3	Interface zu einem angehaltenen Video 1 während einer kritischen Situation.	148
7.4	CCC bei (Strobl, 2010)	156
7.5	Categorical Characteristic Curce für Item 4 aus Video 3	159
7.6	Wrightmap	165

7.7	Normalwahrscheinlichkeitsplot für die Personenfähigkeit physikspezifische professionelle Unterrichtswahrnehmung	167
7.8	Wrightmap	169
7.9	Normalwahrscheinlichkeitsplot für die Personenfähigkeit Classroom Management fokussierte professionelle Unterrichtswahrnehmung	170
7.10	Boxplots zur Varianzanalyse von der physikspezifischen professionelle Unterrichtswahrnehmung mit Hilfe des Fachs	172
7.11	Boxplots zur Varianzanalyse von der physikspezifischen professionelle Unterrichtswahrnehmung mit Hilfe des Studiengangs	173
7.12	Streudiagramm zum physikalischen Wissen an der Hochschule (Y-Achse) und zur physikspezifischen professionellen Unterrichtswahrnehmung (X-Achse) mit einer Ausgleichsgeraden	173
7.13	Streudiagramm zum physikdidaktischen Wissen an der Hochschule (Y-Achse) und zur physikspezifischen professionellen Unterrichtswahrnehmung (X-Achse) mit einer Ausgleichsgeraden	174
7.14	Wrightmap	185
7.15	Streudiagramm physikdidaktisches Wissen und physikspezifische professionelle Unterrichtswahrnehmung	185
7.16	Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe des Fachs	187
7.17	Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe des Studienfortschritts	188
7.18	Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe des Studiengangs	188
7.19	Boxplots zur Varianzanalyse vom physikdidaktischen Wissen mit Hilfe der Hochschule	189
7.20	Verteilung der Antworten zum Decision Making	193

Tabellenverzeichnis

2.1	Angrenzende Konzepte zur professionellen Unterrichtswahrnehmung	32
3.1	Überblick bisheriger Instrumente	44
3.2	Überblick bisheriger Instrumente-Fortsetzung	45
5.1	Schritte der Entwicklungsstudie I	70
5.2	Kodierschema Teil 1	86
5.3	Kodierschema Teil 2	87
5.4	Kodierschema Teil 3	88
5.5	Kodierschema Teil 4	89
5.6	Kodierschema Teil 5	90
5.7	Ergebnisse der Kodierung von Drehbuch 1	91
5.8	Ergebnisse der Kodierung von Drehbuch 2	92
5.9	Ergebnisse der Kodierung von Drehbuch 3	92
5.10	Ergebnisse der Kodierung von Drehbuch 4	93
5.11	Ergebnisse der Kodierung von Drehbuch 5	93
5.12	Ergebnisse zur inhaltlichen Validität	94
5.13	Ergebnis der Lehrkräftebefragung zur curricularen und ökologische Validität – Drehbuch 1	98
5.14	Ergebnis der Lehrkräftebefragung zur curricularen und ökologische Validität – Drehbuch 3	99
5.15	Ergebnis der Lehrkräftebefragung zur curricularen und ökologische Validität – Drehbuch 4	99
6.1	Stichprobe der Entwicklungsstudie II.1	104
6.2	Kodierschema zur Reaktion auf die Videos – Entwicklungsstudie II.1	110
6.3	Kodierschema zur Distraktorenentwicklung – Teil 1	115
6.4	Kodierschema zur Distraktorenentwicklung – Teil 2	116
6.5	Ergebnisse der Intercoderreliabilität für die Erhebung mit Kappa nach Testpersonen	117
6.6	Ergebnisse der Intercoderreliabilität für die Kategorien mit Kappa	117
6.7	Stichprobe der Entwicklungsstudie II.2	122
6.8	Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II	124
6.9	Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 1 . .	125

6.10	Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 2 . .	126
6.11	Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 3 . .	127
6.12	Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 4 . .	128
6.13	Vergleich der Ergebnisse beider Teile von Entwicklungsstudie II – Teil 5 . .	129
6.14	U-Test für den Vergleich der Verteilung auf die verschiedenen Kategorien .	138
6.15	Antwortmöglichkeiten der Level 1 bis 4 für die Situation 2 aus Video 4 . .	140
7.1	Beschreibung der Stichprobe der Hauptstudie	145
7.2	Kodierschema zum Fassen eines Handlungsplans – Teil 1	152
7.3	Kodierschema zum Fassen eines Handlungsplans – Teil 2	153
7.4	Analysehandlungen und ihre Zwecke sowie Ergebnisse	158
7.5	Charakterisierung der Items	162
7.6	Charakterisierung der Items	164
7.7	Thresholds und Fehler	166
7.8	Thresholds und Fehler	168
7.9	Zusammenfassung zu professionellen Unterrichtswahrnehmung	171
7.10	Itemschwierigkeiten der Skalen zum physikdidaktischen Wissen	184
7.11	Zusammenfassung zum physikdidaktischen Wissen	186
7.12	Zusammenfassung zum Decision Making	195

Bisher erschienene Bände der Reihe „*Studien zum Physik- und Chemielernen*“

ISSN 1614-8967 (vormals *Studien zum Physiklernen* ISSN 1435-5280)

- 1 Helmut Fischler, Jochen Peuckert (Hrsg.): Concept Mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie
ISBN 978-3-89722-256-4 40.50 EUR
- 2 Anja Schoster: Bedeutungsentwicklungsprozesse beim Lösen algorithmischer Physikaufgaben. *Eine Fallstudie zu Lernprozessen von Schülern im Physiknachhilfeunterricht während der Bearbeitung algorithmischer Physikaufgaben*
ISBN 978-3-89722-045-4 40.50 EUR
- 3 Claudia von Aufschnaiter: Bedeutungsentwicklungen, Interaktionen und situatives Erleben beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-89722-143-7 40.50 EUR
- 4 Susanne Haerberlen: Lernprozesse im Unterricht mit Wasserstromkreisen. *Eine Fallstudie in der Sekundarstufe I*
ISBN 978-3-89722-172-7 40.50 EUR
- 5 Kerstin Haller: Über den Zusammenhang von Handlungen und Zielen. *Eine empirische Untersuchung zu Lernprozessen im physikalischen Praktikum*
ISBN 978-3-89722-242-7 40.50 EUR
- 6 Michaela Horstendahl: Motivationale Orientierungen im Physikunterricht
ISBN 978-3-89722-227-4 50.00 EUR
- 7 Stefan Deylitz: Lernergebnisse in der Quanten-Atomphysik. *Evaluation des Bremer Unterrichtskonzepts*
ISBN 978-3-89722-291-5 40.50 EUR
- 8 Lorenz Hucke: Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums
ISBN 978-3-89722-316-5 50.00 EUR
- 9 Heike Theyßen: Ein Physikpraktikum für Studierende der Medizin. *Darstellung der Entwicklung und Evaluation eines adressatenspezifischen Praktikums nach dem Modell der Didaktischen Rekonstruktion*
ISBN 978-3-89722-334-9 40.50 EUR
- 10 Annette Schick: Der Einfluß von Interesse und anderen selbstbezogenen Kognitionen auf Handlungen im Physikunterricht. *Fallstudien zu Interessenhandlungen im Physikunterricht*
ISBN 978-3-89722-380-6 40.50 EUR
- 11 Roland Berger: Moderne bildgebende Verfahren der medizinischen Diagnostik. *Ein Weg zu interessanterem Physikunterricht*
ISBN 978-3-89722-445-2 40.50 EUR

- 12 Johannes Werner: Vom Licht zum Atom. *Ein Unterrichtskonzept zur Quantenphysik unter Nutzung des Zeigermodells*
ISBN 978-3-89722-471-1 40.50 EUR
- 13 Florian Sander: Verbindung von Theorie und Experiment im physikalischen Praktikum. *Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum*
ISBN 978-3-89722-482-7 40.50 EUR
- 14 Jörn Gerdes: Der Begriff der physikalischen Kompetenz. *Zur Validierung eines Konstruktes*
ISBN 978-3-89722-510-7 40.50 EUR
- 15 Malte Meyer-Arndt: Interaktionen im Physikpraktikum zwischen Studierenden und Betreuern. *Feldstudie zu Bedeutungsentwicklungsprozessen im physikalischen Praktikum*
ISBN 978-3-89722-541-1 40.50 EUR
- 16 Dietmar Höttecke: Die Natur der Naturwissenschaften historisch verstehen. *Fachdidaktische und wissenschaftshistorische Untersuchungen*
ISBN 978-3-89722-607-4 40.50 EUR
- 17 Gil Gabriel Mavanga: Entwicklung und Evaluation eines experimentell- und phänomenorientierten Optikcurriculums. *Untersuchung zu Schülervorstellungen in der Sekundarstufe I in Mosambik und Deutschland*
ISBN 978-3-89722-721-7 40.50 EUR
- 18 Meike Ute Zastrow: Interaktive Experimentieranleitungen. *Entwicklung und Evaluation eines Konzeptes zur Vorbereitung auf das Experimentieren mit Messgeräten im Physikalischen Praktikum*
ISBN 978-3-89722-802-3 40.50 EUR
- 19 Gunnar Friege: Wissen und Problemlösen. *Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs*
ISBN 978-3-89722-809-2 40.50 EUR
- 20 Erich Starauschek: Physikunterricht nach dem Karlsruher Physikkurs. *Ergebnisse einer Evaluationsstudie*
ISBN 978-3-89722-823-8 40.50 EUR
- 21 Roland Paatz: Charakteristika analogiebasierten Denkens. *Vergleich von Lernprozessen in Basis- und Zielbereich*
ISBN 978-3-89722-944-0 40.50 EUR
- 22 Silke Mikelskis-Seifert: Die Entwicklung von Metakzepten zur Teilchenvorstellung bei Schülern. *Untersuchung eines Unterrichts über Modelle mithilfe eines Systems multipler Repräsentationsebenen*
ISBN 978-3-8325-0013-9 40.50 EUR
- 23 Brunhild Landwehr: Distanzen von Lehrkräften und Studierenden des Sachunterrichts zur Physik. *Eine qualitativ-empirische Studie zu den Ursachen*
ISBN 978-3-8325-0044-3 40.50 EUR

- 24 Lydia Murmann: Physiklernen zu Licht, Schatten und Sehen. *Eine phänomenografische Untersuchung in der Primarstufe*
ISBN 978-3-8325-0060-3 40.50 EUR
- 25 Thorsten Bell: Strukturprinzipien der Selbstregulation. *Komplexe Systeme, Elementarisierungen und Lernprozessstudien für den Unterricht der Sekundarstufe II*
ISBN 978-3-8325-0134-1 40.50 EUR
- 26 Rainer Müller: Quantenphysik in der Schule
ISBN 978-3-8325-0186-0 40.50 EUR
- 27 Jutta Roth: Bedeutungsentwicklungsprozesse von Physikerinnen und Physikern in den Dimensionen Komplexität, Zeit und Inhalt
ISBN 978-3-8325-0183-9 40.50 EUR
- 28 Andreas Saniter: Spezifika der Verhaltensmuster fortgeschrittener Studierender der Physik
ISBN 978-3-8325-0292-8 40.50 EUR
- 29 Thomas Weber: Kumulatives Lernen im Physikunterricht. *Eine vergleichende Untersuchung in Unterrichtsgängen zur geometrischen Optik*
ISBN 978-3-8325-0316-1 40.50 EUR
- 30 Markus Rehm: Über die Chancen und Grenzen moralischer Erziehung im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-0368-0 40.50 EUR
- 31 Marion Budde: Lernwirkungen in der Quanten-Atom-Physik. *Fallstudien über Resonanzen zwischen Lernangeboten und SchülerInnen-Vorstellungen*
ISBN 978-3-8325-0483-0 40.50 EUR
- 32 Thomas Reyer: Oberflächenmerkmale und Tiefenstrukturen im Unterricht. *Exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe*
ISBN 978-3-8325-0488-5 40.50 EUR
- 33 Christoph Thomas Müller: Subjektive Theorien und handlungsleitende Kognitionen von Lehrern als Determinanten schulischer Lehr-Lern-Prozesse im Physikunterricht
ISBN 978-3-8325-0543-1 40.50 EUR
- 34 Gabriela Jonas-Ahrend: Physiklehrervorstellungen zum Experiment im Physikunterricht
ISBN 978-3-8325-0576-9 40.50 EUR
- 35 Dimitrios Stavrou: Das Zusammenspiel von Zufall und Gesetzmäßigkeiten in der nicht-linearen Dynamik. *Didaktische Analyse und Lernprozesse*
ISBN 978-3-8325-0609-4 40.50 EUR
- 36 Katrin Engeln: Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken
ISBN 978-3-8325-0689-6 40.50 EUR
- 37 Susann Hartmann: Erklärungsvielfalt
ISBN 978-3-8325-0730-5 40.50 EUR

- 38 Knut Neumann: Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker
ISBN 978-3-8325-0762-6 40.50 EUR
- 39 Michael Späth: Kontextbedingungen für Physikunterricht an der Hauptschule. *Möglichkeiten und Ansatzpunkte für einen fachübergreifenden, handlungsorientierten und berufsorientierten Unterricht*
ISBN 978-3-8325-0827-2 40.50 EUR
- 40 Jörg Hirsch: Interesse, Handlungen und situatives Erleben von Schülerinnen und Schülern beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-8325-0875-3 40.50 EUR
- 41 Monika Hüther: Evaluation einer hypermedialen Lernumgebung zum Thema Gasgesetze. *Eine Studie im Rahmen des Physikpraktikums für Studierende der Medizin*
ISBN 978-3-8325-0911-8 40.50 EUR
- 42 Maike Tesch: Das Experiment im Physikunterricht. *Didaktische Konzepte und Ergebnisse einer Videostudie*
ISBN 978-3-8325-0975-0 40.50 EUR
- 43 Nina Nicolai: Skriptgeleitete Eltern-Kind-Interaktion bei Chemiehausaufgaben. *Eine Evaluationsstudie im Themenbereich Säure-Base*
ISBN 978-3-8325-1013-8 40.50 EUR
- 44 Antje Leisner: Entwicklung von Modellkompetenz im Physikunterricht
ISBN 978-3-8325-1020-6 40.50 EUR
- 45 Stefan Rumann: Evaluation einer Interventionsstudie zur Säure-Base-Thematik
ISBN 978-3-8325-1027-5 40.50 EUR
- 46 Thomas Wilhelm: Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung – mit CD-ROM
ISBN 978-3-8325-1046-6 45.50 EUR
- 47 Andrea Maier-Richter: Computerunterstütztes Lernen mit Lösungsbeispielen in der Chemie. *Eine Evaluationsstudie im Themenbereich Löslichkeit*
ISBN 978-3-8325-1046-6 40.50 EUR
- 48 Jochen Peuckert: Stabilität und Ausprägung kognitiver Strukturen zum Atombegriff
ISBN 978-3-8325-1104-3 40.50 EUR
- 49 Maik Walpuski: Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback
ISBN 978-3-8325-1184-5 40.50 EUR
- 50 Helmut Fischler, Christiane S. Reiners (Hrsg.): Die Teilchenstruktur der Materie im Physik- und Chemieunterricht
ISBN 978-3-8325-1225-5 34.90 EUR
- 51 Claudia Eysel: Interdisziplinäres Lehren und Lernen in der Lehrerbildung. *Eine empirische Studie zum Kompetenzerwerb in einer komplexen Lernumgebung*
ISBN 978-3-8325-1238-5 40.50 EUR

- 52 Johannes Günther: Lehrerfortbildung über die Natur der Naturwissenschaften. *Studien über das Wissenschaftsverständnis von Grundschullehrkräften*
ISBN 978-3-8325-1287-3 40.50 EUR
- 53 Christoph Neugebauer: Lernen mit Simulationen und der Einfluss auf das Problemlösen in der Physik
ISBN 978-3-8325-1300-9 40.50 EUR
- 54 Andreas Schnirch: Gendergerechte Interessen- und Motivationsförderung im Kontext naturwissenschaftlicher Grundbildung. *Konzeption, Entwicklung und Evaluation einer multimedial unterstützten Lernumgebung*
ISBN 978-3-8325-1334-4 40.50 EUR
- 55 Hilde Köster: Freies Explorieren und Experimentieren. *Eine Untersuchung zur selbstbestimmten Gewinnung von Erfahrungen mit physikalischen Phänomenen im Sachunterricht*
ISBN 978-3-8325-1348-1 40.50 EUR
- 56 Eva Heran-Dörr: Entwicklung und Evaluation einer Lehrerfortbildung zur Förderung der physikdidaktischen Kompetenz von Sachunterrichtslehrkräften
ISBN 978-3-8325-1377-1 40.50 EUR
- 57 Agnes Szabone Varnai: Unterstützung des Problemlösens in Physik durch den Einsatz von Simulationen und die Vorgabe eines strukturierten Kooperationsformats
ISBN 978-3-8325-1403-7 40.50 EUR
- 58 Johannes Rethfeld: Aufgabenbasierte Lernprozesse in selbstorganisationsoffenem Unterricht der Sekundarstufe I zum Themengebiet ELEKTROSTATIK. *Eine Feldstudie in vier 10. Klassen zu einer kartenbasierten Lernumgebung mit Aufgaben aus der Elektrostatik*
ISBN 978-3-8325-1416-7 40.50 EUR
- 59 Christian Henke: Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. *Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven*
ISBN 978-3-8325-1515-7 40.50 EUR
- 60 Lutz Kasper: Diskursiv-narrative Elemente für den Physikunterricht. *Entwicklung und Evaluation einer multimedialen Lernumgebung zum Erdmagnetismus*
ISBN 978-3-8325-1537-9 40.50 EUR
- 61 Thorid Rabe: Textgestaltung und Aufforderung zu Selbsterklärungen beim Physiklernen mit Multimedia
ISBN 978-3-8325-1539-3 40.50 EUR
- 62 Ina Glemnitz: Vertikale Vernetzung im Chemieunterricht. *Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext*
ISBN 978-3-8325-1628-4 40.50 EUR
- 63 Erik Einhaus: Schülerkompetenzen im Bereich Wärmelehre. *Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*
ISBN 978-3-8325-1630-7 40.50 EUR

- 64 Jasmin Neuroth: Concept Mapping als Lernstrategie. *Eine Interventionsstudie zum Chemielernen aus Texten*
ISBN 978-3-8325-1659-8 40.50 EUR
- 65 Hans Gerd Hegeler-Burkhart: Zur Kommunikation von Hauptschülerinnen und Hauptschülern in einem handlungsorientierten und fächerübergreifenden Unterricht mit physikalischen und technischen Inhalten
ISBN 978-3-8325-1667-3 40.50 EUR
- 66 Karsten Rincke: Sprachentwicklung und Fachlernen im Mechanikunterricht. *Sprache und Kommunikation bei der Einführung in den Kraftbegriff*
ISBN 978-3-8325-1699-4 40.50 EUR
- 67 Nina Strehle: Das Ion im Chemieunterricht. *Alternative Schülervorstellungen und curriculare Konsequenzen*
ISBN 978-3-8325-1710-6 40.50 EUR
- 68 Martin Hopf: Problemorientierte Schülerexperimente
ISBN 978-3-8325-1711-3 40.50 EUR
- 69 Anne Beerenwinkel: Fostering conceptual change in chemistry classes using expository texts
ISBN 978-3-8325-1721-2 40.50 EUR
- 70 Roland Berger: Das Gruppenpuzzle im Physikunterricht der Sekundarstufe II. *Eine empirische Untersuchung auf der Grundlage der Selbstbestimmungstheorie der Motivation*
ISBN 978-3-8325-1732-8 40.50 EUR
- 71 Giuseppe Colicchia: Physikunterricht im Kontext von Medizin und Biologie. *Entwicklung und Erprobung von Unterrichtseinheiten*
ISBN 978-3-8325-1746-5 40.50 EUR
- 72 Sandra Winheller: Geschlechtsspezifische Auswirkungen der Lehrer-Schüler-Interaktion im Chemieanfangsunterricht
ISBN 978-3-8325-1757-1 40.50 EUR
- 73 Isabel Wahser: Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie
ISBN 978-3-8325-1815-8 40.50 EUR
- 74 Claus Brell: Lernmedien und Lernerfolg - reale und virtuelle Materialien im Physikunterricht. *Empirische Untersuchungen in achten Klassen an Gymnasien (Laborstudie) zum Computereinsatz mit Simulation und IBE*
ISBN 978-3-8325-1829-5 40.50 EUR
- 75 Rainer Wackermann: Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer
ISBN 978-3-8325-1882-0 40.50 EUR
- 76 Oliver Tepner: Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-1919-3 40.50 EUR

- 77 Claudia Geyer: Museums- und Science-Center-Besuche im naturwissenschaftlichen Unterricht aus einer motivationalen Perspektive. *Die Sicht von Lehrkräften und Schülerinnen und Schülern*
ISBN 978-3-8325-1922-3 40.50 EUR
- 78 Tobias Leonhard: Professionalisierung in der Lehrerbildung. *Eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*
ISBN 978-3-8325-1924-7 40.50 EUR
- 79 Alexander Kauertz: Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben
ISBN 978-3-8325-1925-4 40.50 EUR
- 80 Regina Hübinger: Schüler auf Weltreise. *Entwicklung und Evaluation von Lehr-/Lernmaterialien zur Förderung experimentell-naturwissenschaftlicher Kompetenzen für die Jahrgangsstufen 5 und 6*
ISBN 978-3-8325-1932-2 40.50 EUR
- 81 Christine Waltner: Physik lernen im Deutschen Museum
ISBN 978-3-8325-1933-9 40.50 EUR
- 82 Torsten Fischer: Handlungsmuster von Physiklehrkräften beim Einsatz neuer Medien. *Fallstudien zur Unterrichtspraxis*
ISBN 978-3-8325-1948-3 42.00 EUR
- 83 Corinna Kieren: Chemiehausaufgaben in der Sekundarstufe I des Gymnasiums. *Fragebogenerhebung zur gegenwärtigen Praxis und Entwicklung eines optimierten Hausaufgabendesigns im Themenbereich Säure-Base*
978-3-8325-1975-9 37.00 EUR
- 84 Marco Thiele: Modelle der Thermohalinen Zirkulation im Unterricht. *Eine empirische Studie zur Förderung des Modellverständnisses*
ISBN 978-3-8325-1982-7 40.50 EUR
- 85 Bernd Zinn: Physik lernen, um Physik zu lehren. *Eine Möglichkeit für interessanteren Physikunterricht*
ISBN 978-3-8325-1995-7 39.50 EUR
- 86 Esther Klaes: Außerschulische Lernorte im naturwissenschaftlichen Unterricht. *Die Perspektive der Lehrkraft*
ISBN 978-3-8325-2006-9 43.00 EUR
- 87 Marita Schmidt: Kompetenzmodellierung und -diagnostik im Themengebiet Energie der Sekundarstufe I. *Entwicklung und Erprobung eines Testinventars*
ISBN 978-3-8325-2024-3 37.00 EUR
- 88 Gudrun Franke-Braun: Aufgaben mit gestuften Lernhilfen. *Ein Aufgabenformat zur Förderung der sachbezogenen Kommunikation und Lernleistung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-2026-7 38.00 EUR
- 89 Silke Klos: Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht. *Der Einfluss eines integrierten Unterrichtskonzepts*
ISBN 978-3-8325-2133-2 37.00 EUR

- 90 Ulrike Elisabeth Burkard: Quantenphysik in der Schule. *Bestandsaufnahme, Perspektiven und Weiterentwicklungsmöglichkeiten durch die Implementation eines Medienservers*
ISBN 978-3-8325-2215-5 43.00 EUR
- 91 Ulrike Gromadecki: Argumente in physikalischen Kontexten. *Welche Geltungsgründe halten Physikanfänger für überzeugend?*
ISBN 978-3-8325-2250-6 41.50 EUR
- 92 Jürgen Bruns: Auf dem Weg zur Förderung naturwissenschaftsspezifischer Vorstellungen von zukünftigen Chemie-Lehrenden
ISBN 978-3-8325-2257-5 43.50 EUR
- 93 Cornelius Marsch: Räumliche Atomvorstellung. *Entwicklung und Erprobung eines Unterrichtskonzeptes mit Hilfe des Computers*
ISBN 978-3-8325-2293-3 82.50 EUR
- 94 Maja Brückmann: Sachstrukturen im Physikunterricht. *Ergebnisse einer Videostudie*
ISBN 978-3-8325-2272-8 39.50 EUR
- 95 Sabine Fechner: Effects of Context-oriented Learning on Student Interest and Achievement in Chemistry Education
ISBN 978-3-8325-2343-5 36.50 EUR
- 96 Clemens Nagel: eLearning im Physikalischen Anfängerpraktikum
ISBN 978-3-8325-2355-8 39.50 EUR
- 97 Josef Riese: Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften
ISBN 978-3-8325-2376-3 39.00 EUR
- 98 Sascha Bernholt: Kompetenzmodellierung in der Chemie. *Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität*
ISBN 978-3-8325-2447-0 40.00 EUR
- 99 Holger Christoph Stawitz: Auswirkung unterschiedlicher Aufgabenprofile auf die Schülerleistung. *Vergleich von Naturwissenschafts- und Problemlöseaufgaben der PISA 2003-Studie*
ISBN 978-3-8325-2451-7 37.50 EUR
- 100 Hans Ernst Fischer, Elke Sumfleth (Hrsg.): nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht
ISBN 978-3-8325-3331-1 40.00 EUR
- 101 Hendrik Härtig: Sachstrukturen von Physikschulbüchern als Grundlage zur Bestimmung der Inhaltsvalidität eines Tests
ISBN 978-3-8325-2512-5 34.00 EUR
- 102 Thomas Grüß-Niehaus: Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht. *Der Effekt von Methoden progressiver und kollaborativer Reflexion*
ISBN 978-3-8325-2537-8 40.50 EUR

- 103 Patrick Bronner: Quantenoptische Experimente als Grundlage eines Curriculums zur Quantenphysik des Photons
ISBN 978-3-8325-2540-8 36.00 EUR
- 104 Adrian Voßkühler: Blickbewegungsmessung an Versuchsaufbauten. *Studien zur Wahrnehmung, Verarbeitung und Usability von physikbezogenen Experimenten am Bildschirm und in der Realität*
ISBN 978-3-8325-2548-4 47.50 EUR
- 105 Verena Tobias: Newton'sche Mechanik im Anfangsunterricht. *Die Wirksamkeit einer Einführung über die zweidimensionale Dynamik auf das Lehren und Lernen*
ISBN 978-3-8325-2558-3 54.00 EUR
- 106 Christian Rogge: Entwicklung physikalischer Konzepte in aufgabenbasierten Lernumgebungen
ISBN 978-3-8325-2574-3 45.00 EUR
- 107 Mathias Ropohl: Modellierung von Schülerkompetenzen im Basiskonzept Chemische Reaktion. *Entwicklung und Analyse von Testaufgaben*
ISBN 978-3-8325-2609-2 36.50 EUR
- 108 Christoph Kulgemeyer: Physikalische Kommunikationskompetenz. *Modellierung und Diagnostik*
ISBN 978-3-8325-2674-0 44.50 EUR
- 109 Jennifer Olszewski: The Impact of Physics Teachers' Pedagogical Content Knowledge on Teacher Actions and Student Outcomes
ISBN 978-3-8325-2680-1 33.50 EUR
- 110 Annika Ohle: Primary School Teachers' Content Knowledge in Physics and its Impact on Teaching and Students' Achievement
ISBN 978-3-8325-2684-9 36.50 EUR
- 111 Susanne Mannel: Assessing scientific inquiry. *Development and evaluation of a test for the low-performing stage*
ISBN 978-3-8325-2761-7 40.00 EUR
- 112 Michael Plomer: Physik physiologisch passend praktiziert. *Eine Studie zur Lernwirksamkeit von traditionellen und adressatenspezifischen Physikpraktika für die Physiologie*
ISBN 978-3-8325-2804-1 34.50 EUR
- 113 Alexandra Schulz: Experimentierspezifische Qualitätsmerkmale im Chemieunterricht. *Eine Videostudie*
ISBN 978-3-8325-2817-1 40.00 EUR
- 114 Franz Boczianowski: Eine empirische Untersuchung zu Vektoren im Physikunterricht der Mittelstufe
ISBN 978-3-8325-2843-0 39.50 EUR
- 115 Maria Ploog: Internetbasiertes Lernen durch Textproduktion im Fach Physik
ISBN 978-3-8325-2853-9 39.50 EUR

- 116 Anja Dhein: Lernen in Explorier- und Experimentiersituationen. *Eine explorative Studie zu Bedeutungsentwicklungsprozessen bei Kindern im Alter zwischen 4 und 6 Jahren*
ISBN 978-3-8325-2859-1 45.50 EUR
- 117 Irene Neumann: Beyond Physics Content Knowledge. *Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*
ISBN 978-3-8325-2880-5 37.00 EUR
- 118 Markus Emden: Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. *Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*
ISBN 978-3-8325-2867-6 38.00 EUR
- 119 Birgit Hofmann: Analyse von Blickbewegungen von Schülern beim Lesen von physikbezogenen Texten mit Bildern. *Eye Tracking als Methodenwerkzeug in der physikdidaktischen Forschung*
ISBN 978-3-8325-2925-3 59.00 EUR
- 120 Rebecca Knobloch: Analyse der fachinhaltlichen Qualität von Schüleräußerungen und deren Einfluss auf den Lernerfolg. *Eine Videostudie zu kooperativer Kleingruppenarbeit*
ISBN 978-3-8325-3006-8 36.50 EUR
- 121 Julia Hostenbach: Entwicklung und Prüfung eines Modells zur Beschreibung der Bewertungskompetenz im Chemieunterricht
ISBN 978-3-8325-3013-6 38.00 EUR
- 122 Anna Windt: Naturwissenschaftliches Experimentieren im Elementarbereich. *Evaluation verschiedener Lernsituationen*
ISBN 978-3-8325-3020-4 43.50 EUR
- 123 Eva Kölbach: Kontexteinflüsse beim Lernen mit Lösungsbeispielen
ISBN 978-3-8325-3025-9 38.50 EUR
- 124 Anna Lau: Passung und vertikale Vernetzung im Chemie- und Physikunterricht
ISBN 978-3-8325-3021-1 36.00 EUR
- 125 Jan Lamprecht: Ausbildungswege und Komponenten professioneller Handlungskompetenz. *Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*
ISBN 978-3-8325-3035-8 38.50 EUR
- 126 Ulrike Böhm: Förderung von Verstehensprozessen unter Einsatz von Modellen
ISBN 978-3-8325-3042-6 41.00 EUR
- 127 Sabrina Dollny: Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften
ISBN 978-3-8325-3046-4 37.00 EUR
- 128 Monika Zimmermann: Naturwissenschaftliche Bildung im Kindergarten. *Eine integrative Längsschnittstudie zur Kompetenzentwicklung von Erzieherinnen*
ISBN 978-3-8325-3053-2 54.00 EUR

- 129 Ulf Saballus: Über das Schlussfolgern von Schülerinnen und Schülern zu öffentlichen Kontroversen mit naturwissenschaftlichem Hintergrund. *Eine Fallstudie*
ISBN 978-3-8325-3086-0 39.50 EUR
- 130 Olaf Krey: Zur Rolle der Mathematik in der Physik. *Wissenschaftstheoretische Aspekte und Vorstellungen Physiklernender*
ISBN 978-3-8325-3101-0 46.00 EUR
- 131 Angelika Wolf: Zusammenhänge zwischen der Eigenständigkeit im Physikunterricht, der Motivation, den Grundbedürfnissen und dem Lernerfolg von Schülern
ISBN 978-3-8325-3161-4 45.00 EUR
- 132 Johannes Börlin: Das Experiment als Lerngelegenheit. *Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*
ISBN 978-3-8325-3170-6 45.00 EUR
- 133 Olaf Uhden: Mathematisches Denken im Physikunterricht. *Theorieentwicklung und Problemanalyse*
ISBN 978-3-8325-3170-6 45.00 EUR
- 134 Christoph Gut: Modellierung und Messung experimenteller Kompetenz. *Analyse eines large-scale Experimentiertests*
ISBN 978-3-8325-3213-0 40.00 EUR
- 135 Antonio Rueda: Lernen mit ExploMultimedial in kolumbianischen Schulen. *Analyse von kurzzeitigen Lernprozessen und der Motivation beim länderübergreifenden Einsatz einer deutschen computergestützten multimedialen Lernumgebung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-3218-5 45.50 EUR
- 136 Krisztina Berger: Bilder, Animationen und Notizen. *Empirische Untersuchung zur Wirkung einfacher visueller Repräsentationen und Notizen auf den Wissenserwerb in der Optik*
ISBN 978-3-8325-3238-3 41.50 EUR
- 137 Antony Crossley: Untersuchung des Einflusses unterschiedlicher physikalischer Konzepte auf den Wissenserwerb in der Thermodynamik der Sekundarstufe I
ISBN 978-3-8325-3275-8 40.00 EUR
- 138 Tobias Viering: Entwicklung physikalischer Kompetenz in der Sekundarstufe I. *Validierung eines Kompetenzentwicklungsmodells für das Energiekonzept im Bereich Fachwissen*
ISBN 978-3-8325-3277-2 37.00 EUR
- 139 Nico Schreiber: Diagnostik experimenteller Kompetenz. *Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*
ISBN 978-3-8325-3284-0 39.00 EUR
- 140 Sarah Hundertmark: Einblicke in kollaborative Lernprozesse. *Eine Fallstudie zur reflektierenden Zusammenarbeit unterstützt durch die Methoden Concept Mapping und Lernbegleitbogen*
ISBN 978-3-8325-3251-2 43.00 EUR

- 141 Ronny Scherer: Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. *Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*
ISBN 978-3-8325-3312-0 43.00 EUR
- 142 Patricia Heitmann: Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. *Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*
ISBN 978-3-8325-3314-4 37.00 EUR
- 143 Jan Fleischhauer: Wissenschaftliches Argumentieren und Entwicklung von Konzepten beim Lernen von Physik
ISBN 978-3-8325-3325-0 35.00 EUR
- 144 Nermin Özcan: Zum Einfluss der Fachsprache auf die Leistung im Fach Chemie. *Eine Förderstudie zur Fachsprache im Chemieunterricht*
ISBN 978-3-8325-3328-1 36.50 EUR
- 145 Helena van Vorst: Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie
ISBN 978-3-8325-3321-2 38.50 EUR
- 146 Janine Cappell: Fachspezifische Diagnosekompetenz angehender Physiklehrkräfte in der ersten Ausbildungsphase
ISBN 978-3-8325-3356-4 38.50 EUR
- 147 Susanne Bley: Förderung von Transferprozessen im Chemieunterricht
ISBN 978-3-8325-3407-3 40.50 EUR
- 148 Cathrin Blaes: Die übungsgestützte Lehrerpräsentation im Chemieunterricht der Sekundarstufe I. *Evaluation der Effektivität*
ISBN 978-3-8325-3409-7 43.50 EUR
- 149 Julia Suckut: Die Wirksamkeit von piko-OWL als Lehrerfortbildung. Eine Evaluation zum Projekt *Physik im Kontext* in Fallstudien
ISBN 978-3-8325-3440-0 45.00 EUR
- 150 Alexandra Dorschu: Die Wirkung von Kontexten in Physikkompetenztestaufgaben
ISBN 978-3-8325-3446-2 37.00 EUR
- 151 Jochen Scheid: Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: *Ein Beitrag zur Entwicklung der Aufgabenkultur*
ISBN 978-3-8325-3449-3 49.00 EUR
- 152 Tim Plasa: Die Wahrnehmung von Schülerlaboren und Schülerforschungszentren
ISBN 978-3-8325-3483-7 35.50 EUR
- 153 Felix Schoppmeier: Physikkompetenz in der gymnasialen Oberstufe. *Entwicklung und Validierung eines Kompetenzstrukturmodells für den Kompetenzbereich Umgang mit Fachwissen*
ISBN 978-3-8325-3502-5 36.00 EUR

- 154 Katharina Groß: Experimente alternativ dokumentieren. *Eine qualitative Studie zur Förderung der Diagnose- und Differenzierungskompetenz in der Chemielehrerbildung*
ISBN 978-3-8325-3508-7 43.50 EUR
- 155 Barbara Hank: Konzeptwandelprozesse im Anfangsunterricht Chemie. *Eine quasixperimentelle Längsschnittstudie*
ISBN 978-3-8325-3519-3 38.50 EUR
- 156 Katja Freyer: Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie
ISBN 978-3-8325-3544-5 38.00 EUR
- 157 Alexander Rachel: Auswirkungen instruktionaler Hilfen bei der Einführung des (Ferro-)Magnetismus. *Eine Vergleichsstudie in der Primar- und Sekundarstufe*
ISBN 978-3-8325-3548-3 43.50 EUR
- 158 Sebastian Ritter: Einfluss des Lerninhalts Nanogrößeneffekte auf Teilchen- und Teilchenmodellvorstellungen von Schülerinnen und Schülern
ISBN 978-3-8325-3558-2 36.00 EUR
- 159 Andrea Harbach: Problemorientierung und Vernetzung in kontextbasierten Lernaufgaben
ISBN 978-3-8325-3564-3 39.00 EUR
- 160 David Obst: Interaktive Tafeln im Physikunterricht. *Entwicklung und Evaluation einer Lehrerfortbildung*
ISBN 978-3-8325-3582-7 40.50 EUR
- 161 Sophie Kirschner: Modellierung und Analyse des Professionswissens von Physiklehrkräften
ISBN 978-3-8325-3601-5 35.00 EUR
- 162 Katja Stief: Selbstregulationsprozesse und Hausaufgabenmotivation im Chemieunterricht
ISBN 978-3-8325-3631-2 34.00 EUR
- 163 Nicola Meschede: Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. *Theoretische Beschreibung und empirische Erfassung*
ISBN 978-3-8325-3668-8 37.00 EUR
- 164 Johannes Maximilian Barth: Experimentieren im Physikunterricht der gymnasialen Oberstufe. *Eine Rekonstruktion übergeordneter Einbettungsstrategien*
ISBN 978-3-8325-3681-7 39.00 EUR
- 165 Sandra Lein: Das Betriebspraktikum in der Lehrerbildung. *Eine Untersuchung zur Förderung der Wissenschafts- und Technikbildung im allgemeinbildenden Unterricht*
ISBN 978-3-8325-3698-5 40.00 EUR
- 166 Veranika Maiseyenka: Modellbasiertes Experimentieren im Unterricht. *Praxistauglichkeit und Lernwirkungen*
ISBN 978-3-8325-3708-1 38.00 EUR

- 167 Christoph Stolzenberger: Der Einfluss der didaktischen Lernumgebung auf das Erreichen geforderter Bildungsziele am Beispiel der W- und P-Seminare im Fach Physik
ISBN 978-3-8325-3708-1 38.00 EUR
- 168 Pia Altenburger: Mehrebenenregressionsanalysen zum Physiklernen im Sachunterricht der Primarstufe. *Ergebnisse einer Evaluationsstudie.*
ISBN 978-3-8325-3717-3 37.50 EUR
- 169 Nora Ferber: Entwicklung und Validierung eines Testinstruments zur Erfassung von Kompetenzentwicklung im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-3727-2 39.50 EUR
- 170 Anita Stender: Unterrichtsplanung: Vom Wissen zum Handeln.
Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung
ISBN 978-3-8325-3750-0 41.50 EUR
- 171 Jenna Koenen: Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen
ISBN 978-3-8325-3785-2 43.00 EUR
- 172 Teresa Henning: Empirische Untersuchung kontextorientierter Lernumgebungen in der Hochschuldidaktik. *Entwicklung und Evaluation kontextorientierter Aufgaben in der Studieneingangsphase für Fach- und Nebenfachstudierende der Physik*
ISBN 978-3-8325-3801-9 43.00 EUR
- 173 Alexander Pusch: Fachspezifische Instrumente zur Diagnose und individuellen Förderung von Lehramtsstudierenden der Physik
ISBN 978-3-8325-3829-3 38.00 EUR
- 174 Christoph Vogelsang: Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. *Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*
ISBN 978-3-8325-3846-0 50.50 EUR
- 175 Ingo Brebeck: Selbstreguliertes Lernen in der Studieneingangsphase im Fach Chemie
ISBN 978-3-8325-3859-0 37.00 EUR
- 176 Axel Eghtessad: Merkmale und Strukturen von Professionalisierungsprozessen in der ersten und zweiten Phase der Chemielehrerbildung. *Eine empirisch-qualitative Studie mit niedersächsischen Fachleiter_innen der Sekundarstufenlehrkräfte*
ISBN 978-3-8325-3861-3 45.00 EUR
- 177 Andreas Nehring: Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung
ISBN 978-3-8325-3872-9 39.50 EUR
- 178 Maike Schmidt: Professionswissen von Sachunterrichtslehrkräften. Zusammenhangsanalyse zur Wirkung von Ausbildungshintergrund und Unterrichtserfahrung auf das fachspezifische Professionswissen im Unterrichtsinhalt „Verbrennung“
ISBN 978-3-8325-3907-8 38.50 EUR

- 179 Jan Winkelmann: Auswirkungen auf den Fachwissenszuwachs und auf affektive Schülermerkmale durch Schüler- und Demonstrationsexperimente im Physikunterricht
ISBN 978-3-8325-3915-3 41.00 EUR
- 180 Iwen Kobow: Entwicklung und Validierung eines Testinstrumentes zur Erfassung der Kommunikationskompetenz im Fach Chemie
ISBN 978-3-8325-3927-6 34.50 EUR
- 181 Yvonne Gramzow: Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik. Modellierung und Testkonstruktion
ISBN 978-3-8325-3931-3 42.50 EUR
- 182 Evelin Schröter: Entwicklung der Kompetenzerwartung durch Lösen physikalischer Aufgaben einer multimedialen Lernumgebung
ISBN 978-3-8325-3975-7 54.50 EUR
- 183 Inga Kallweit: Effektivität des Einsatzes von Selbsteinschätzungsbögen im Chemieunterricht der Sekundarstufe I. *Individuelle Förderung durch selbstreguliertes Lernen*
ISBN 978-3-8325-3965-8 44.00 EUR
- 184 Andrea Schumacher: Paving the way towards authentic chemistry teaching. *A contribution to teachers' professional development*
ISBN 978-3-8325-3976-4 48.50 EUR
- 185 David Woitkowski: Fachliches Wissen Physik in der Hochschulausbildung. *Konzeptualisierung, Messung, Niveaubildung*
ISBN 978-3-8325-3988-7 53.00 EUR
- 186 Marianne Korner: Cross-Age Peer Tutoring in Physik. *Evaluation einer Unterrichtsmethode*
ISBN 978-3-8325-3979-5 38.50 EUR
- 187 Simone Nakoinz: Untersuchung zur Verknüpfung submikroskopischer und makroskopischer Konzepte im Fach Chemie
ISBN 978-3-8325-4057-9 38.50 EUR
- 188 Sandra Anus: Evaluation individueller Förderung im Chemieunterricht. *Adaptivität von Lerninhalten an das Vorwissen von Lernenden am Beispiel des Basiskonzeptes Chemische Reaktion*
ISBN 978-3-8325-4059-3 43.50 EUR
- 189 Thomas Roßbegalle: Fachdidaktische Entwicklungsforschung zum besseren Verständnis atmosphärischer Phänomene. *Treibhauseffekt, saurer Regen und stratosphärischer Ozonabbau als Kontexte zur Vermittlung von Basiskonzepten der Chemie*
ISBN 978-3-8325-4059-3 45.50 EUR
- 190 Kathrin Steckenmesser-Sander: Gemeinsamkeiten und Unterschiede physikbezogener Handlungs-, Denk- und Lernprozesse von Mädchen und Jungen
ISBN 978-3-8325-4066-1 38.50 EUR
- 191 Cornelia Geller: Lernprozessorientierte Sequenzierung des Physikunterrichts im Zusammenhang mit Fachwissenserwerb. *Eine Videostudie in Finnland, Deutschland und der Schweiz*
ISBN 978-3-8325-4082-1 35.50 EUR

- 192 Jan Hofmann: Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme
ISBN 978-3-8325-4104-0 38.50 EUR
- 193 Andreas Dickhäuser: Chemiespezifischer Humor. *Theoriebildung, Materialentwicklung, Evaluation*
ISBN 978-3-8325-4108-8 37.00 EUR
- 194 Stefan Korte: Die Grenzen der Naturwissenschaft als Thema des Physikunterrichts
ISBN 978-3-8325-4112-5 57.50 EUR
- 195 Carolin Hülsmann: Kurswahlmotive im Fach Chemie. Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe
ISBN 978-3-8325-4144-6 49.00 EUR
- 196 Caroline Körbs: Mindeststandards im Fach Chemie am Ende der Pflichtschulzeit
ISBN 978-3-8325-4148-4 34.00 EUR
- 197 Andreas Vorholzer: Wie lassen sich Kompetenzen des experimentellen Denkens und Arbeitens fördern? *Eine empirische Untersuchung der Wirkung eines expliziten und eines impliziten Instruktionsansatzes*
ISBN 978-3-8325-4194-1 37.50 EUR
- 198 Anna Katharina Schmitt: Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung
ISBN 978-3-8325-4228-3 39.50 EUR
- 199 Christian Maurer: Strukturierung von Lehr-Lern-Sequenzen
ISBN 978-3-8325-4247-4 36.50 EUR
- 200 Helmut Fischler, Elke Sumfleth (Hrsg.): Professionelle Kompetenz von Lehrkräften der Chemie und Physik
ISBN 978-3-8325-4523-9 34.00 EUR
- 201 Simon Zander: Lehrerfortbildung zu Basismodellen und Zusammenhänge zum Fachwissen
ISBN 978-3-8325-4248-1 35.00 EUR
- 202 Kerstin Arndt: Experimentierkompetenz erfassen.
Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie
ISBN 978-3-8325-4266-5 45.00 EUR
- 203 Christian Lang: Kompetenzorientierung im Rahmen experimentalchemischer Praktika
ISBN 978-3-8325-4268-9 42.50 EUR
- 204 Eva Cauet: Testen wir relevantes Wissen? *Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten*
ISBN 978-3-8325-4276-4 39.50 EUR
- 205 Patrick Löffler: Modellanwendung in Problemlöseaufgaben. *Wie wirkt Kontext?*
ISBN 978-3-8325-4303-7 35.00 EUR

- 206 Carina Gehlen: Kompetenzstruktur naturwissenschaftlicher Erkenntnisgewinnung im Fach Chemie
ISBN 978-3-8325-4318-1 43.00 EUR
- 207 Lars Oettinghaus: Lehrerüberzeugungen und physikbezogenes Professionswissen. *Vergleich von Absolventinnen und Absolventen verschiedener Ausbildungswege im Physikreferendariat*
ISBN 978-3-8325-4319-8 38.50 EUR
- 208 Jennifer Petersen: Zum Einfluss des Merkmals Humor auf die Gesundheitsförderung im Chemieunterricht der Sekundarstufe I. *Eine Interventionsstudie zum Thema Sonnenschutz*
ISBN 978-3-8325-4348-8 40.00 EUR
- 209 Philipp Straube: Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik
ISBN 978-3-8325-4351-8 35.50 EUR
- 210 Martin Dickmann: Messung von Experimentierfähigkeiten. *Validierungsstudien zur Qualität eines computerbasierten Testverfahrens*
ISBN 978-3-8325-4356-3 41.00 EUR
- 211 Markus Bohlmann: Science Education. Empirie, Kulturen und Mechanismen der Didaktik der Naturwissenschaften
ISBN 978-3-8325-4377-8 44.00 EUR
- 212 Martin Draude: Die Kompetenz von Physiklehrkräften, Schwierigkeiten von Schülerinnen und Schülern beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-4382-2 37.50 EUR
- 213 Henning Rode: Prototypen evidenzbasierten Physikunterrichts. *Zwei empirische Studien zum Einsatz von Feedback und Blackboxes in der Sekundarstufe*
ISBN 978-3-8325-4389-1 42.00 EUR
- 214 Jan-Henrik Kechel: Schülerschwierigkeiten beim eigenständigen Experimentieren. *Eine qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*
ISBN 978-3-8325-4392-1 55.00 EUR
- 215 Katharina Fricke: Classroom Management and its Impact on Lesson Outcomes in Physics. *A multi-perspective comparison of teaching practices in primary and secondary schools*
ISBN 978-3-8325-4394-5 40.00 EUR
- 216 Hannes Sander: Orientierungen von Jugendlichen beim Urteilen und Entscheiden in Kontexten nachhaltiger Entwicklung. *Eine rekonstruktive Perspektive auf Bewertungskompetenz in der Didaktik der Naturwissenschaft*
ISBN 978-3-8325-4434-8 46.00 EUR
- 217 Inka Haak: Maßnahmen zur Unterstützung kognitiver und metakognitiver Prozesse in der Studieneingangsphase. *Eine Design-Based-Research-Studie zum universitären Lernzentrum Physiktreff*
ISBN 978-3-8325-4437-9 46.50 EUR

- 218 Martina Brandenburger: Was beeinflusst den Erfolg beim Problemlösen in der Physik?
Eine Untersuchung mit Studierenden
ISBN 978-3-8325-4409-6 42.50 EUR
- 219 Corinna Helms: Entwicklung und Evaluation eines Trainings zur Verbesserung der Erklärqualität von Schülerinnen und Schülern im Gruppenpuzzle
ISBN 978-3-8325-4454-6 42.50 EUR
- 220 Viktoria Rath: Diagnostische Kompetenz von angehenden Physiklehrkräften. *Modellierung, Testinstrumentenentwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik*
ISBN 978-3-8325-4456-0 42.50 EUR
- 221 Janne Krüger: Schülerperspektiven auf die zeitliche Entwicklung der Naturwissenschaften
ISBN 978-3-8325-4457-7 45.50 EUR
- 222 Stefan Mutke: Das Professionswissen von Chemiereferendarinnen und -referendaren in Nordrhein-Westfalen. *Eine Längsschnittstudie*
ISBN 978-3-8325-4458-4 37.50 EUR
- 223 Sebastian Habig: Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren
ISBN 978-3-8325-4467-6 40.50 EUR
- 224 Sven Liepertz: Zusammenhang zwischen dem Professionswissen von Physiklehrkräften, dem sachstrukturellen Angebot des Unterrichts und der Schülerleistung
ISBN 978-3-8325-4480-5 34.00 EUR
- 225 Elina Platova: Optimierung eines Laborpraktikums durch kognitive Aktivierung
ISBN 978-3-8325-4481-2 39.00 EUR
- 226 Tim Reschke: Lesegeschichten im Chemieunterricht der Sekundarstufe I zur Unterstützung von situationalem Interesse und Lernerfolg
ISBN 978-3-8325-4487-4 41.00 EUR
- 227 Lena Mareike Walper: Entwicklung der physikbezogenen Interessen und selbstbezogenen Kognitionen von Schülerinnen und Schülern in der Übergangsphase von der Primar- in die Sekundarstufe. *Eine Längsschnittanalyse vom vierten bis zum siebten Schuljahr*
ISBN 978-3-8325-4495-9 43.00 EUR
- 228 Stefan Anthofer: Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden
ISBN 978-3-8325-4498-0 39.50 EUR
- 229 Marcel Bullinger: Handlungsorientiertes Physiklernen mit instruierten Selbsterklärungen in der Primarstufe. *Eine experimentelle Laborstudie*
ISBN 978-3-8325-4504-8 44.00 EUR
- 230 Thomas Amenda: Bedeutung fachlicher Elementarisierungen für das Verständnis der Kinematik
ISBN 978-3-8325-4531-4 43.50 EUR

- 231 Sabrina Milke: Beeinflusst *Priming* das Physiklernen?
Eine empirische Studie zum Dritten Newtonschen Axiom
ISBN 978-3-8325-4549-4 42.00 EUR
- 232 Corinna Erfmann: Ein anschaulicher Weg zum Verständnis der elektromagnetischen Induktion. *Evaluation eines Unterrichtsvorschlags und Validierung eines Leistungsdiagnoseinstruments*
ISBN 978-3-8325-4550-5 49.50 EUR
- 233 Hanne Rautenstrauch: Erhebung des (Fach-)Sprachstandes bei Lehramtsstudierenden im Kontext des Faches Chemie
ISBN 978-3-8325-4556-7 40.50 EUR
- 234 Tobias Klug: Wirkung kontextorientierter physikalischer Praktikumsversuche auf Lernprozesse von Studierenden der Medizin
ISBN 978-3-8325-4558-1 37.00 EUR
- 235 Mareike Bohrmann: Zur Förderung des Verständnisses der Variablenkontrolle im naturwissenschaftlichen Sachunterricht
ISBN 978-3-8325-4559-8 52.00 EUR
- 236 Anja Schödl: FALKO-Physik – Fachspezifische Lehrerkompetenzen im Fach Physik. *Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften*
ISBN 978-3-8325-4553-6 40.50 EUR
- 237 Hilda Scheuermann: Entwicklung und Evaluation von Unterstützungsmaßnahmen zur Förderung der Variablenkontrollstrategie beim Planen von Experimenten
ISBN 978-3-8325-4568-0 39.00 EUR
- 238 Christian G. Strippel: Naturwissenschaftliche Erkenntnisgewinnung an chemischen Inhalten vermitteln. *Konzeption und empirische Untersuchung einer Ausstellung mit Experimentierstation*
ISBN 978-3-8325-4577-2 41.50 EUR
- 239 Sarah Rau: Durchführung von Sachunterricht im Vorbereitungsdienst. *Eine längsschnittliche, videobasierte Unterrichtsanalyse*
ISBN 978-3-8325-4579-6 46.00 EUR
- 240 Thomas Plotz: Lernprozesse zu nicht-sichtbarer Strahlung. *Empirische Untersuchungen in der Sekundarstufe 2*
ISBN 978-3-8325-4624-3 39.50 EUR
- 241 Wolfgang Aschauer: Elektrische und magnetische Felder. *Eine empirische Studie zu Lernprozessen in der Sekundarstufe II*
ISBN 978-3-8325-4625-0 50.00 EUR
- 242 Anna Donhauser: Didaktisch rekonstruierte Materialwissenschaft. *Aufbau und Konzeption eines Schülerlabors für den Exzellenzcluster Engineering of Advanced Materials*
ISBN 978-3-8325-4636-6 39.00 EUR

- 243 Katrin Schüller: Lernen mit Lösungsbeispielen im Chemieunterricht. *Einflüsse auf Lernerfolg, kognitive Belastung und Motivation*
ISBN 978-3-8325-4640-3 42.50 EUR
- 244 Timo Fleischer: Untersuchung der chemischen Fachsprache unter besonderer Berücksichtigung chemischer Repräsentationen
ISBN 978-3-8325-4642-7 46.50 EUR
- 245 Rosina Steininger: Concept Cartoons als Stimuli für Kleingruppendiskussionen im Chemieunterricht. *Beschreibung und Analyse einer komplexen Lerngelegenheit*
ISBN 978-3-8325-4647-2 39.00 EUR
- 246 Daniel Rehfeldt: Erfassung der Lehrqualität naturwissenschaftlicher Experimentalpraktika
ISBN 978-3-8325-4590-1 40.00 EUR
- 247 Sandra Puddu: Implementing Inquiry-based Learning in a Diverse Classroom: Investigating Strategies of Scaffolding and Students' Views of Scientific Inquiry
ISBN 978-3-8325-4591-8 35.50 EUR
- 248 Markus Bliersbach: Kreativität in der Chemie. *Erhebung und Förderung der Vorstellungen von Chemielehramtsstudierenden*
ISBN 978-3-8325-4593-2 44.00 EUR
- 249 Lennart Kimpel: Aufgaben in der Allgemeinen Chemie. *Zum Zusammenspiel von chemischem Verständnis und Rechenfähigkeit*
ISBN 978-3-8325-4618-2 36.00 EUR
- 250 Louise Bindel: Effects of integrated learning: explicating a mathematical concept in inquiry-based science camps
ISBN 978-3-8325-4655-7 37.50 EUR
- 251 Michael Wenzel: Computereinsatz in Schule und Schülerlabor. *Einstellung von Physiklehrkräften zu Neuen Medien*
ISBN 978-3-8325-4659-5 38.50 EUR
- 252 Laura Muth: Einfluss der Auswertephase von Experimenten im Physikunterricht. *Ergebnisse einer Interventionsstudie zum Zuwachs von Fachwissen und experimenteller Kompetenz von Schülerinnen und Schülern*
ISBN 978-3-8325-4675-5 36.50 EUR
- 253 Annika Fricke: Interaktive Skripte im Physikalischen Praktikum. *Entwicklung und Evaluation von Hypermedien für die Nebenfachausbildung*
ISBN 978-3-8325-4676-2 41.00 EUR
- 254 Julia Haase: Selbstbestimmtes Lernen im naturwissenschaftlichen Sachunterricht. *Eine empirische Interventionsstudie mit Fokus auf Feedback und Kompetenzerleben*
ISBN 978-3-8325-4685-4 38.50 EUR
- 255 Antje J. Heine: Was ist Theoretische Physik? *Eine wissenschaftstheoretische Betrachtung und Rekonstruktion von Vorstellungen von Studierenden und Dozenten über das Wesen der Theoretischen Physik*
ISBN 978-3-8325-4691-5 46.50 EUR

- 256 Claudia Meinhardt: Entwicklung und Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen von (angehenden) Physiklehrkräften in physikdidaktischen Handlungsfeldern
ISBN 978-3-8325-4712-7 47.00 EUR
- 257 Ann-Kathrin Schlüter: Professionalisierung angehender Chemielehrkräfte für einen Gemeinsamen Unterricht
ISBN 978-3-8325-4713-4 53.50 EUR
- 258 Stefan Richtberg: Elektronenbahnen in Feldern. Konzeption und Evaluation einer webbasierten Lernumgebung
ISBN 978-3-8325-4723-3 49.00 EUR
- 259 Jan-Philipp Burde: Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells
ISBN 978-3-8325-4726-4 57.50 EUR
- 260 Frank Finkenberg: Flipped Classroom im Physikunterricht
ISBN 978-3-8325-4737-4 42.50 EUR
- 261 Florian Treisch: Die Entwicklung der Professionellen Unterrichtswahrnehmung im Lehr-Lern-Labor Seminar
ISBN 978-3-8325-4741-4 41.50 EUR
- 262 Desiree Mayr: Strukturiertheit des experimentellen naturwissenschaftlichen Problemlöseprozesses
ISBN 978-3-8325-4757-8 37.00 EUR
- 263 Katrin Weber: Entwicklung und Validierung einer Learning Progression für das Konzept der chemischen Reaktion in der Sekundarstufe I
ISBN 978-3-8325-4762-2 48.50 EUR
- 264 Hauke Bartels: Entwicklung und Bewertung eines performanznahen Videovignetten-tests zur Messung der Erklärfähigkeit von Physiklehrkräften
ISBN 978-3-8325-4804-9 37.00 EUR
- 265 Karl Marniok: Zum Wesen von Theorien und Gesetzen in der Chemie. *Begriffsanalyse und Förderung der Vorstellungen von Lehramtsstudierenden*
ISBN 978-3-8325-4805-6 42.00 EUR
- 266 Marisa Holzapfel: Fachspezifischer Humor als Methode in der Gesundheitsbildung im Übergang von der Primarstufe zur Sekundarstufe I
ISBN 978-3-8325-4808-7 50.00 EUR
- 267 Anna Stolz: Die Auswirkungen von Experimentiersituationen mit unterschiedlichem Öffnungsgrad auf Leistung und Motivation der Schülerinnen und Schüler
ISBN 978-3-8325-4781-3 38.00 EUR
- 268 Nina Ulrich: Interaktive Lernaufgaben in dem digitalen Schulbuch eChemBook. *Einfluss des Interaktivitätsgrads der Lernaufgaben und des Vorwissens der Lernenden auf den Lernerfolg*
ISBN 978-3-8325-4814-8 43.50 EUR

- 269 Kim-Alessandro Weber: Quantenoptik in der Lehrerfortbildung. *Ein bedarfsgeprägtes Fortbildungskonzept zum Quantenobjekt Photon mit Realexperimenten*
ISBN 978-3-8325-4792-9 55.00 EUR
- 270 Nina Skorsetz: Empathisierer und Systematisierer im Vorschulalter. *Eine Fragebogen- und Videostudie zur Motivation, sich mit Naturphänomenen zu beschäftigen*
ISBN 978-3-8325-4825-4 43.50 EUR
- 271 Franziska Kehne: Analyse des Transfers von kontextualisiert erworbenem Wissen im Fach Chemie
ISBN 978-3-8325-4846-9 45.00 EUR
- 272 Markus Elsholz: Das akademische Selbstkonzept angehender Physiklehrkräfte als Teil ihrer professionellen Identität. *Dimensionalität und Veränderung während einer zentralen Praxisphase*
ISBN 978-3-8325-4857-5 37.50 EUR
- 273 Joachim Müller: Studienerfolg in der Physik. *Zusammenhang zwischen Modellierungskompetenz und Studienerfolg*
ISBN 978-3-8325-4859-9 35.00 EUR
- 274 Jennifer Dörschelln: Organische Leuchtdioden. *Implementation eines innovativen Themas in den Chemieunterricht*
ISBN 978-3-8325-4865-0 59.00 EUR
- 275 Stephanie Strelow: Beliefs von Studienanfängern des Kombi-Bachelors Physik über die Natur der Naturwissenschaften
ISBN 978-3-8325-4881-0 40.50 EUR
- 276 Dennis Jaeger: Kognitive Belastung und aufgabenspezifische sowie personenspezifische Einflussfaktoren beim Lösen von Physikaufgaben
ISBN 978-3-8325-4928-2 50.50 EUR
- 277 Vanessa Fischer: Der Einfluss von Interesse und Motivation auf die Messung von Fach- und Bewertungskompetenz im Fach Chemie
ISBN 978-3-8325-4933-6 39.00 EUR
- 278 René Dohrmann: Professionsbezogene Wirkungen einer Lehr-Lern-Labor-Veranstaltung. *Eine multimethodische Studie zu den professionsbezogenen Wirkungen einer Lehr-Lern-Labor-Blockveranstaltung auf Studierende der Bachelorstudiengänge Lehramt Physik und Grundschulpädagogik (Sachunterricht)*
ISBN 978-3-8325-4958-9 40.00 EUR
- 279 Meike Bergs: Can We Make Them Use These Strategies? *Fostering Inquiry-Based Science Learning Skills with Physical and Virtual Experimentation Environments*
ISBN 978-3-8325-4962-6 39.50 EUR
- 280 Marie-Therese Hauerstein: Untersuchung zur Effektivität von Strukturierung und Binnendifferenzierung im Chemieunterricht der Sekundarstufe I. *Evaluation der Strukturierungshilfe Lernleiter*
ISBN 978-3-8325-4982-4 42.50 EUR

- 281 Verena Zucker: Erkennen und Beschreiben von formativem Assessment im naturwissenschaftlichen Grundschulunterricht. *Entwicklung eines Instruments zur Erfassung von Teilfähigkeiten der professionellen Wahrnehmung von Lehramtsstudierenden*
ISBN 978-3-8325-4991-6 38.00 EUR
- 282 Victoria Telser: Erfassung und Förderung experimenteller Kompetenz von Lehrkräften im Fach Chemie
ISBN 978-3-8325-4996-1 50.50 EUR
- 283 Kristine Tschirschky: Entwicklung und Evaluation eines gedächtnisorientierten Aufgabendesigns für Physikaufgaben
ISBN 978-3-8325-5002-8 42.50 EUR
- 284 Thomas Elert: Course Success in the Undergraduate General Chemistry Lab
ISBN 978-3-8325-5004-2 41.50 EUR
- 285 Britta Kalthoff: Explizit oder implizit? *Untersuchung der Lernwirksamkeit verschiedener fachmethodischer Instruktionen im Hinblick auf fachmethodische und fachinhaltliche Fähigkeiten von Sachunterrichtsstudierenden*
ISBN 978-3-8325-5013-4 37.50 EUR
- 286 Thomas Dickmann: Visuelles Modellverständnis und Studienerfolg in der Chemie. *Zwei Seiten einer Medaille*
ISBN 978-3-8325-5016-5 44.00 EUR
- 287 Markus Sebastian Feser: Physiklehrkräfte korrigieren Schülertexte. *Eine Explorationsstudie zur fachlich-konzeptuellen und sprachlichen Leistungsfeststellung und -beurteilung im Physikunterricht*
ISBN 978-3-8325-5020-2 49.00 EUR
- 288 Matylda Dudzinska: Lernen mit Beispielaufgaben und Feedback im Physikunterricht der Sekundarstufe 1. *Energieerhaltung zur Lösung von Aufgaben nutzen*
ISBN 978-3-8325-5025-7 47.00 EUR
- 289 Ines Sonnenschein: Naturwissenschaftliche Denk- und Arbeitsprozesse Studierender im Labor
ISBN 978-3-8325-5033-2 52.00 EUR
- 290 Florian Simon: Der Einfluss von Betreuung und Betreuenden auf die Wirksamkeit von Schülerlaborbesuchen. *Eine Zusammenhangsanalyse von Betreuungsqualität, Betreuermerkmalen und Schülerlaborzielen sowie Replikationsstudie zur Wirksamkeit von Schülerlaborbesuchen*
ISBN 978-3-8325-5036-3 49.50 EUR
- 291 Marie-Annette Geyer: Physikalisch-mathematische Darstellungswechsel funktionaler Zusammenhänge. *Das Vorgehen von SchülerInnen der Sekundarstufe 1 und ihre Schwierigkeiten*
ISBN 978-3-8325-5047-9 46.50 EUR
- 292 Susanne Digel: Messung von Modellierungskompetenz in Physik. *Theoretische Herleitung und empirische Prüfung eines Kompetenzmodells physikspezifischer Modellierungskompetenz*
ISBN 978-3-8325-5055-4 41.00 EUR

- 293 Sönke Janssen: Angebots-Nutzungs-Prozesse eines Schülerlabors analysieren und gestalten. *Ein design-based research Projekt*
ISBN 978-3-8325-5065-3 57.50 EUR
- 294 Knut Wille: Der Productive Failure Ansatz als Beitrag zur Weiterentwicklung der Aufgabenkultur
ISBN 978-3-8325-5074-5 49.00 EUR
- 295 Lisanne Kraeva: Problemlösestrategien von Schülerinnen und Schülern diagnostizieren
ISBN 978-3-8325-5110-0 59.50 EUR
- 296 Jenny Lorentzen: Entwicklung und Evaluation eines Lernangebots im Lehramtsstudium Chemie zur Förderung von Vernetzungen innerhalb des fachbezogenen Professionswissens
ISBN 978-3-8325-5120-9 39.50 EUR
- 297 Micha Winkelmann: Lernprozesse in einem Schülerlabor unter Berücksichtigung individueller naturwissenschaftlicher Interessenstrukturen
ISBN 978-3-8325-5147-6 48.50 EUR
- 298 Carina Wöhlke: Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte
ISBN 978-3-8325-5149-0 43.00 EUR
- 299 Thomas Schubatzky: Das Amalgam Anfangs-Elektrizitätslehreunterricht. *Eine multiperspektivische Betrachtung in Deutschland und Österreich*
ISBN 978-3-8325-5159-9 50.50 EUR
- 300 Amany Annaggar: A Design Framework for Video Game-Based Gamification Elements to Assess Problem-solving Competence in Chemistry Education
ISBN 978-3-8325-5150-6 52.00 EUR
- 301 Alexander Engl: CHEMIE PUR – Unterrichten in der Natur: *Entwicklung und Evaluation eines kontextorientierten Unterrichtskonzepts im Bereich Outdoor Education zur Veränderung der Einstellung zu „Chemie und Natur“*
ISBN 978-3-8325-5174-2 59.00 EUR
- 302 Christin Marie Sajons: Kognitive und motivationale Dynamik in Schülerlaboren. *Kontextualisierung, Problemorientierung und Autonomieunterstützung der didaktischen Struktur analysieren und weiterentwickeln*
ISBN 978-3-8325-5155-1 56.00 EUR
- 303 Philipp Bitzenbauer: Quantenoptik an Schulen. *Studie im Mixed-Methods Design zur Evaluation des Erlanger Unterrichtskonzepts zur Quantenoptik*
ISBN 978-3-8325-5123-0 59.00 EUR

- 304 Malte Ubben: Typisierung des Verständnisses mentaler Modelle mittels empirischer Datenerhebung am Beispiel der Quantenphysik
ISBN 978-3-8325-5181-0 43.50 EUR
- 305 Wiebke Kuske-Janßen: Sprachlicher Umgang mit Formeln von LehrerInnen im Physikunterricht am Beispiel des elektrischen Widerstandes in Klassenstufe 8
ISBN 978-3-8325-5183-4 47.50 EUR

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder per Fax (030 - 42 85 10 92) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Hans Niedderer, Helmut Fischler und Elke Sumfleth

Die Reihe umfasst inzwischen eine große Zahl von wissenschaftlichen Arbeiten aus vielen Arbeitsgruppen der Physik- und Chemiedidaktik und zeichnet damit ein gültiges Bild der empirischen physik- und chemiedidaktischen Forschung in Deutschland.

Die Herausgeber laden daher Interessenten zu neuen Beiträgen ein und bitten sie, sich im Bedarfsfall an den Logos-Verlag oder an ein Mitglied des Herausgeberteams zu wenden.

Kontaktadressen:

Prof. Dr. Hans Niedderer
Institut für Didaktik der Naturwissenschaften,
Abt. Physikdidaktik, FB Physik/Elektrotechnik,
Universität Bremen,
Postfach 33 04 40, 28334 Bremen
Tel. 0421-218 2484/4695, e-mail:
niedderer@physik.uni-bremen.de

Prof. Dr. Helmut Fischler
Didaktik der Physik, FB Physik, Freie Universität Berlin,
Arnimallee 14, 14195 Berlin
Tel. 030-838 56712/55966, e-mail:
fischler@physik.fu-berlin.de

Prof. Dr. Elke Sumfleth
Didaktik der Chemie,
Fachbereich Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen
Tel. 0201-183 3757/3761, e-mail:
elke.sumfleth@uni-essen.de

Die professionelle Unterrichtswahrnehmung ist eine zentrale und situative Fähigkeit, die (angehende) Lehrkräfte befähigt, Lernprozesse im Unterricht zu erkennen, mithilfe ihres professionellen Wissens zu interpretieren und anschließend lernförderliche Handlungspläne abzuleiten. Der vorliegenden Arbeit liegt die Forschungsfrage zugrunde, ob diese Fähigkeit mit einem physikspezifischen Fokus mithilfe eines Instruments - bestehend aus inszenierten Unterrichtsvideos und einem teilweise geschlossenen Fragebogen - valide erfasst werden kann. Ein solches Instrument könnte dazu dienen, herauszufinden, ob die professionelle Unterrichtswahrnehmung domänenspezifisch ist (Hypothese 1) und ob diese situative Fähigkeit entlang der universitären Ausbildung eine Progression erfährt (Hypothese 2).

Die Ergebnisse der Arbeit ergeben 19 Items (EAP Reliabilität liegt bei 0.698), mit denen eine physikspezifische professionelle Unterrichtswahrnehmung gemessen werden kann. Weitere Analysen (zur Korrelation mit dem physikdidaktischen Wissen und der Abhängigkeit von studiumsbezogenen Daten) stützen die Konstruktvalidität des Tests. In Bezug auf Hypothese 1 hat sich wie erwartet gezeigt, dass die physikspezifische professionelle Unterrichtswahrnehmung signifikant von dem studierten Fach und dem Studiengang abhängt. Anders als erwartet ergeben die Ergebnisse zur Hypothese 2, dass die professionelle Unterrichtswahrnehmung aber nicht signifikant vom Studienfortschritt abhängt.

Logos Verlag Berlin

ISBN 978-3-8325-5149-0