

Reuter, Timo; Zucker, Verena; Leuchter, Miriam
Förderung des Beschreibens von prozessorientierter Diagnostik im naturwissenschaftlichen Sachunterricht – Evaluation eines auf Text- und Videovignetten basierenden Seminars für Grundschullehrerinnen- und Lehrerbildung 37 (2019) 2, S. 275-288

Beiträge zur Lehrerinnen- und Lehrerbildung 37 (2019) 2, S. 275-288



Quellenangabe/ Reference:

Reuter, Timo; Zucker, Verena; Leuchter, Miriam: Förderung des Beschreibens von prozessorientierter Diagnostik im naturwissenschaftlichen Sachunterricht – Evaluation eines auf Text- und Videovignetten basierenden Seminars für Grundschullehrerinnen- und Lehrerbildung 37 (2019) 2, S. 275-288 - In: Beiträge zur Lehrerinnen- und Lehrerbildung 37 (2019) 2, S. 275-288 - URN: urn:nbn:de:0111-pedocs-182784 - DOI: 10.25656/01:18278

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-182784>

<https://doi.org/10.25656/01:18278>

in Kooperation mit / in cooperation with:



<http://www.bzl-online.ch>

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

Zeitschrift zu Theorie und Praxis der Aus- und
Weiterbildung von Lehrerinnen und Lehrern

BEITRÄGE ZUR LEHRERINNEN- UND LEHRERBILDUNG

Forschend lernen – forschend unterrichten

Editorial

Sandra Moroni, Markus Wilhelm, Christian Brühwiler, Annette Tettenborn,
Bruno Leutwyler, Kurt Reusser, Markus Weil 147

Schwerpunkt

Forschend lernen – forschend unterrichten

Roman Suter Forschendes Lernen in der Lehrerinnen- und
Lehrerbildung – Definitionen, Begründungen und Formen 150

Peter Vetter, Markus Gerteis und Sandra Moroni Kompetenzbereich
«Forschungsmethoden»: Was sollen angehende Lehrpersonen am Ende
ihrer Ausbildung aus der Sicht von in der Forschungsausbildung tätigen
Dozierenden können? 160

Martina Homt und Stefanie van Ophuysen Forschendes Lernen in
Praxissemester und Beruf – Einstellungen und Handlungsintention
von Lehramtsstudierenden 177

Christina Egger Aufbau eines Verständnisses von Forschung im forschungs-
orientierten Lernen im Sachunterricht: Welche Rolle spielt die Lehrperson? 192

Christina Huber Die eigene Hochschule erforschen – ein Praxisbeispiel
aus der Lehrerinnen- und Lehrerbildung 208

Manuela Keller-Schneider Forschendes Lernen – das eigene Lernen
erforschen 218

Doreen Holtsch und Elisabeth Riebenbauer Forschendes Lernen in
der fachdidaktischen Ausbildung von Lehrpersonen. Selbsteinschätzungen
in Bezug auf Orientierungen, Interesse und Wissen von Studierenden im
Masterstudiengang «Wirtschaftspädagogik» 230

Friederike Runge Fallarbeit als Format Forschenden Lernens in der
Lehrpersonenbildung der Didaktik der Geisteswissenschaften 250

Victoria L. Miczajka-Rußmann und Kim Lange-Schubert
Citizen-Science-Projekte als besondere Lerngelegenheit im Kontext
des Forschenden Lernens am Beispiel der naturwissenschaftsbezogenen
Lehrerinnen- und Lehrerbildung im Grundschullehramt 263

Forum

- Timo Reuter, Verena Zucker und Miriam Leuchter** Förderung des Beschreibens von prozessorientierter Diagnostik im naturwissenschaftlichen Sachunterricht – Evaluation eines auf Text- und Videovignetten basierenden Seminars für Grundschullehrerinnenstudierende 275

Rubriken

Buchbesprechungen

- Ehmke, T., Hammer, S., Köker, A., Ohm, U. & Koch-Priewe, B. (Hrsg.). (2018). Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache (Charlotte Röhner) 289

- Münch, R. (2018). Der bildungsindustrielle Komplex. Schule und Unterricht im Wettbewerbsstaat (Jürg Frick) 291

- Wullschleger, A. (2017). Individuell-adaptive Lernunterstützung im Kindergarten. Eine Videoanalyse zur spielintegrierten Förderung von Mengen-Zahlen-Kompetenzen (Esther Brunner) 294

- Neuerscheinungen** 296

- Zeitschriftenspiegel** 298

Förderung des Beschreibens von prozessorientierter Diagnostik im naturwissenschaftlichen Sachunterricht – Evaluation eines auf Text- und Videovignetten basierenden Seminars für Grundschullehramtsstudierende

Timo Reuter, Verena Zucker und Miriam Leuchter

Zusammenfassung Die Fähigkeit, diagnostische Massnahmen einer Lehrperson in verschiedenen Unterrichtssituationen beschreiben zu können, ist ein Teilprozess der professionellen Wahrnehmung von prozessorientierter Diagnostik. Diese Fähigkeit kann als eine zentrale Voraussetzung für das Umsetzen eigener prozessorientierter Diagnostik im naturwissenschaftlichen Sachunterricht betrachtet werden. Ein vignettenbasiertes Seminar für Grundschullehramtsstudierende vermochte das Beschreiben von diagnostischen Massnahmen nur teilweise zu fördern. Der Beitrag berichtet die Ergebnisse der Seminarevaluation, diskutiert eine denkbare Weiterentwicklung des Seminars und wirft methodische Fragen zum Erhebungsinstrument auf.

Schlagwörter professionelle Wahrnehmung – Beschreiben – Vignetten – prozessorientierte Diagnostik – naturwissenschaftlicher Sachunterricht

Promotion of the description of process-oriented diagnostics in science teaching – Evaluation of a text- and video-vignette-based seminar for primary-school pre-service teachers

Abstract A teacher's ability to describe diagnostic measures in various teaching situations is a sub-process of professional vision, which, in turn, is a central prerequisite for the implementation of process-oriented diagnostics in science teaching. A vignette-based seminar for primary-school pre-service teachers was only partially successful in promoting their ability to describe diagnostic measures. The article reports on the results of the evaluation of the seminar, discusses a potential way of developing the seminar, and raises methodological questions about the instrument.

Keywords professional vision – describing – vignettes – process-oriented diagnostics – science teaching

1 Einleitung

Professionelle Wahrnehmung von bedeutsamen Lehr-Lern-Situationen im Unterricht gilt als Voraussetzung für professionelles Handeln von Lehrpersonen (Seidel & Stürmer, 2014). Professionelle Wahrnehmung umfasst als Teilprozesse das Erkennen, Beschreiben und Erklären von bedeutsamen Situationen sowie das Treffen von Vorhersagen (Schäfer & Seidel, 2015). Studien zeigen, dass professionelle Wahrnehmung, u.a.

das Beschreiben, bereits in der ersten Phase der Lehramtsausbildung gefördert werden kann (z.B. Gold, Hellermann, Burgula & Holodynski, 2016). Professionelle Wahrnehmung wird zunehmend auch im Hinblick auf prozessorientierte Diagnostik diskutiert (Hoth, Döhrmann, Kaiser, Busse, König & Blömeke, 2016), da deren Umsetzung Lehrpersonen häufig schwerfällt (z.B. Morrison & Lederman, 2003). Daher scheint es bedeutsam, die professionelle Wahrnehmung von prozessorientierter Diagnostik bereits in der ersten Phase der Lehramtsausbildung verstärkt zu fördern. Als grundlegend kann dabei insbesondere die Förderung des Beschreibens von prozessorientierter Diagnostik gesehen werden, da sich 1) im Beschreiben zeigt, welche diagnostischen Massnahmen erkannt werden und 2) das Beschreiben eine Voraussetzung für Erklärungen und Vorhersagen darstellt.

Um das Beschreiben von prozessorientierter Diagnostik zu fördern, bedarf es einer engen Theorie-Praxis-Verknüpfung, die beispielsweise mit Text- und Videovignetten realisiert werden kann (von Aufschnaiter, Selter & Michaelis, 2017). Text- und Videovignetten ermöglichen ein praxisnahes Lernen, da sie diagnostische Massnahmen einer Lehrperson in unterschiedlichen Lehr-Lern-Situationen exemplarisch abbilden können und – zumindest teilweise – der Komplexität von Unterricht gerecht werden. Die vorliegende Studie geht der Frage nach, ob bei Grundschullehramtsstudierenden das Beschreiben von Massnahmen der Diagnostik durch ein auf Text- und Videovignetten basierendes Seminar zur prozessorientierten Diagnostik gefördert werden kann.¹

2 Theoretischer Hintergrund

2.1 Bedeutung von prozessorientierter Diagnostik im naturwissenschaftlichen Sachunterricht

Dem Diagnostizieren von aktuellen Lernständen kommt im Sachunterricht eine bedeutsame Rolle zu. Lernende besitzen bereits vielfältige Konzepte zu naturwissenschaftlichen Inhalten, wenn sie in den Unterricht kommen (Harlen, 2001). Diese unterscheiden sich hinsichtlich ihrer wissenschaftlichen Adäquatheit. Einige Konzepte weisen z.B. richtige Ansätze auf und müssen lediglich ausdifferenziert werden. Die meisten Konzepte können wissenschaftlichen Ansprüchen allerdings nicht standhalten (Vosniadou, Vamvakoussi & Skopeliti, 2008). Da Konzepte einen Einfluss auf die Lernprozesse der Schülerinnen und Schüler haben, sind sie Ausgangspunkt für deren Veränderung (Möller, 2010). Die prozessorientierte Diagnostik dieser Konzepte stellt Lehrpersonen vor grosse Herausforderungen:

1. Die Lehrperson benötigt einen Überblick über die unterschiedlichen Konzepte. Da Kinder z.B. nicht alle Konzepte verbalisieren, muss sie um Möglichkeiten unterschiedlicher diagnostischer Massnahmen wissen, um prozessorientierte Diagnos-

¹ Diese Studie wurde gefördert vom BMBF im Rahmen des Projekts MoSAiK der «Qualitätsoffensive Lehrerbildung» sowie von der Deutschen Telekom Stiftung.

tik durchführen zu können. Diese Massnahmen umfassen nach Leuders, Dörfler, Leuders und Philipp (2018) a) das Analysieren von Produkten der Lernenden (z.B. Zeichnungen), b) das Beobachten (z.B. beim Umgang mit Material, Interaktion der Lernenden untereinander) und c) das Diagnostizieren im Gespräch (z.B. Vermutungen erfragen, Ergebnisse zusammenfassen lassen).

2. Die Lehrperson muss diagnostizierte Konzepte bei der weiteren Planung berücksichtigen, um für möglichst alle Lernenden herausfordernde Lernumgebungen zu schaffen (Hewson & Lemberger, 2000).
3. Die Lehrperson muss weiter beachten, dass sich durch die angebotenen Lernumgebungen die Konzepte der Kinder stetig und individuell verändern. Beispielsweise gibt es Kinder, die ihre wissenschaftlich nicht haltbaren Vorstellungen in annähernd wissenschaftliche Konzepte umstrukturieren können (Chi, 2008). Anderen Kindern wiederum gelingt diese Umstrukturierung nicht, weil die neuen Konzepte für sie nicht überzeugend sind. Die individuellen Konzepte der Kinder müssen daher während des gesamten Lernprozesses von der Lehrperson mithilfe verschiedener Massnahmen diagnostiziert werden. Solch eine prozessorientierte Diagnostik ist die Voraussetzung, um die Passung der bereitgestellten Lernangebote zu überprüfen und diese gegebenenfalls zu verändern (Bell & Cowie, 2001). Empirische Befunde zeigen allerdings, dass es Lehrpersonen häufig schwerfällt, prozessorientierte Diagnostik in ihrem Unterricht anzuwenden (Gotwals, Philhower, Cistema & Bennett, 2015; Morrison & Lederman, 2003).

2.2 Beschreiben von Diagnostik

Die lückenhafte Anwendung von prozessorientierter Diagnostik könnte damit zusammenhängen, dass Lehrpersonen Schwierigkeiten damit haben, überhaupt diagnostische Massnahmen als solche im Unterrichtsgeschehen zu erkennen. Eine Beschreibung solcher diagnostischen Massnahmen wäre somit ebenfalls nur unzureichend möglich. In diesem Beitrag wird unter dem Beschreiben eine kognitive Fähigkeit verstanden, die über die blosser Wiedergabe von Gesehenem hinausgeht. Vielmehr umfasst das Beschreiben im vorliegenden Kontext das Identifizieren und Berichten diagnostischer Massnahmen, die durch zwei parallel ablaufende Prozesse gekennzeichnet sind (vgl. Meschede, Steffensky, Wolters & Möller, 2015): 1) Ausblendung nicht diagnostischer Massnahmen einer Lehrperson (z.B. Unterstützungsmassnahmen) und 2) Konkretisierung der erkannten diagnostischen Massnahmen. Sowohl das Ausblenden als auch das Konkretisieren setzen dabei eine theoretische Wissensbasis zur prozessorientierten Diagnostik voraus sowie die Fähigkeit, dieses Wissen auch in einer spezifischen Lehr-Lern-Situation anwenden zu können. Das Beschreiben als Teilfähigkeit der professionellen Wahrnehmung wird daher gemäss der Forschungsliteratur im Folgenden auch als Indikator für anwendbares Wissen erachtet (Seidel & Stürmer, 2014). Die Verfügbarkeit und die Anwendbarkeit von Wissen sind jedoch abhängig von Professionalisierungsprozessen (angehender) Lehrpersonen (Meschede, Fiebranz, Möller & Steffensky, 2017). Lehramtsstudierenden fällt im Gegensatz zu praktizierenden Lehrpersonen teilweise bereits das Beschreiben von Unterrichtssituationen schwer (Schäfer & Seidel,

2015). Aus diesem Grund erscheint es bedeutsam, das Beschreiben von Diagnostik schon in der universitären Ausbildung zu fördern.

2.3 Förderung des Beschreibens in der universitären Lehramtsausbildung

Zahlreiche Studien sprechen dafür, dass professionelle Wahrnehmung, u.a. das Beschreiben, bei Studierenden in universitären Seminaren durch den Einsatz von Vignetten gefördert werden kann (Gold et al., 2016; Sunder, Todorova & Möller, 2016). Dies lässt sich daher auch für das Beschreiben von Diagnostik vermuten. Vignetten ermöglichen es den Studierenden, einen Teil der Unterrichtskomplexität authentisch und realitätsnah unter verschiedenen Perspektiven zu analysieren (Krammer & Reusser, 2005; Syring, Bohl, Kleinknecht, Kuntze, Rehm & Schneider, 2015). Dies kann sowohl mit fremden als auch mit eigenen Lehr-Lern-Situationen der Studierenden geschehen (Gold, Hellermann & Holodynski, 2017). Allerdings kann die in den Vignetten abgebildete Unterrichtskomplexität auch zu einer erhöhten kognitiven Belastung bei Studierenden führen (Syring et al., 2015), was sich negativ auf ihre Wahrnehmung auswirken kann (Sweller, 1994). Syring et al. (2015) zeigten, dass Lehramtsstudierende Textvignetten als kognitiv weniger belastend empfanden als Videovignetten. Sie führen dies auf die Sequenzialität und die fokussiertere Darstellung von Handlungen bei Textvignetten zurück. Doch auch Textvignetten können insbesondere für Studierende kognitiv herausfordernd sein (Gold et al., 2016), da eine mentale Repräsentation der beschriebenen Lehr-Lern-Situation anhand der textlichen Informationen gebildet werden muss (Rawson & Kintsch, 2004).

Studien zeigen, dass die Förderung der professionellen Wahrnehmung, u.a. des Beschreibens, an ihre Grenzen stösst. Beispielsweise liess sich das Beschreiben von Unterrichtsqualität (z.B. Klassenführung, Aufgabengestaltung etc.) bei angehenden Lehrpersonen im Rahmen eines Seminars nicht gleichermassen für verschiedene Dimensionen fördern (Star, Lynch & Perova, 2011). Die Ergebnisse von Star et al. (2011) weisen damit u.a. auf die Schwierigkeit hin, bei Studierenden Abgrenzungswissen (Oser, Hascher & Spychiger, 1999) – d.h. Wissen darüber, wie oder was etwas *nicht* ist – aufzubauen: Lehramtsstudierende beschrieben nach einem Seminar zwar mehr Ereignisse in den Vignetten, konnten aber nicht ausreichend zwischen bedeutsamen und nicht bedeutsamen Merkmalen unterscheiden (Star et al., 2011). Denkbar ist, dass dabei auch der kognitive Prozess der Übergeneralisierung eine Rolle spielte. Übergeneralisierung beschreibt die unangemessene Übertragung und Anwendung des in einer spezifischen Situation erworbenen Wissens auf alle anderen (ähnlichen) Situationen (Epstein, 1992). Williams, Lombrozo und Rehder (2013) zeigten in einer experimentellen Studie, dass Lernende insbesondere dann zu Übergeneralisierungen neigten, wenn sie während des Lernprozesses wiederholt zur Erklärung des Gelernten angehalten wurden.

Um eine potenzielle Förderung des Beschreibens in Seminaren überhaupt feststellen zu können, bedarf es passender Messinstrumente. In bestehenden Studien erfolgte die Messung meist anhand offener oder geschlossener Antwortformate, die die Betrach-

tenden jeweils nach der Sichtung einer Text- oder Videovignette beantworten mussten (z.B. Meschede, 2014). Aus testmethodischer Sicht birgt die Beantwortung von Fragen aus der Erinnerung heraus die Gefahr von False-Memory-Effekten. Der Begriff «False Memory» bezieht sich auf einen kognitionspsychologischen Effekt, wonach sich Menschen an Dinge oder Ereignisse erinnern, die entweder gar nicht vorkamen oder anders geschahen (Deese, 1959; Roediger & McDermott, 1995). Loftus und Palmer (1974) konnten in einer experimentellen Studie aufzeigen, wie vorformulierte Fragen menschliche Urteile über zurückliegende Ereignisse – insbesondere bei Unsicherheit – beeinflussen können. Dieser Befund erscheint vor allem im Hinblick auf ein geschlossenes Antwortformat zur Messung des Beschreibens bedeutsam zu sein.

3 Fragestellung

Bisher ist noch nicht untersucht worden, ob bei Studierenden das Beschreiben von diagnostischen Massnahmen im naturwissenschaftlichen Sachunterricht durch ein auf Text- und Videovignetten basierendes Seminar gefördert und ein dahingehender Kompetenzzuwachs bei Studierenden gemessen werden kann. Die Studie geht daher der folgenden Fragestellung nach: *Kann das Beschreiben von Massnahmen der Diagnostik bei Grundschullehrerstudierenden durch ein auf Text- und Videovignetten basierendes Seminar zur prozessorientierten Diagnostik gefördert werden?*

4 Methode

4.1 Seminarkonzept

Im Rahmen des Projekts MoSAiK («Modulare Schulpraxiseinbindung als Ausgangspunkt zur individuellen Kompetenzentwicklung») der «Qualitätsoffensive Lehrerbildung» wurde an der Universität Koblenz-Landau (Campus Landau) ein Seminar entwickelt, das das Beschreiben von Diagnostik der Studierenden mithilfe von Text- und Videovignetten schulen soll (vgl. auch Gold et al., 2017; Sunder et al., 2016). Im ersten Drittel des Seminars erarbeiteten sich die Studierenden theoretisches Grundlagenwissen zur prozessorientierten Diagnostik und zu entsprechenden Massnahmen, indem sie verschiedene Massnahmen wie z.B. das Analysieren von Produkten der Schülerinnen und Schüler, das Beobachten und das Diagnostizieren im Gespräch kennenlernten. Die Studierenden sollten danach das erworbene Wissen anwenden, indem sie Ausschnitte von text- und videobasierten Lehr-Lern-Situationen fremder Lehrpersonen theoriebasiert im Hinblick auf den Einsatz diagnostischer Massnahmen analysierten (vgl. Kramer, Hugener, Biaggi, Frommelt, Furrer Auf der Maur & Stürmer, 2016). Im zweiten Drittel planten die Studierenden eine kurze Lehr-Lern-Situation unter Berücksichtigung von diagnostischen Massnahmen, indem sie Aufgaben zu einem naturwissenschaftlichen Thema formulierten und deren diagnostisches Potenzial herausarbeiteten. Diese Aufgaben setzten die Studierenden dann in einer Schulklasse um und videografierten

die Lehr-Lern-Situation. Im letzten Drittel analysierten sie die von ihnen umgesetzten Massnahmen der Diagnostik.

4.2 Prä- und Posttestung des Beschreibens

Die Fähigkeit der Studierenden, diagnostische Massnahmen einer Lehrperson in Lehr-Lern-Situationen zu beschreiben, wurde mithilfe von zwei text- und zwei videobasierten Unterrichtssituationen (im Folgenden: Testvignetten) vor (Prätest) und nach dem Seminar (Posttest) onlinebasiert erfasst. Die Videos hatten eine Dauer von 2:42 bzw. 2:55 Minuten. Die Texte hatten eine Länge von durchschnittlich 1700 Zeichen. Jede Testvignette zeigte naturwissenschaftlichen Sachunterricht (Schwimmen und Sinken, Kerzenflamme und Verbrennung, Variablenkontrollstrategie am Beispiel der schiefen Ebene, Magnetismus). Die Testvignetten wurden eigens für die Verwendung im Instrument erstellt, bildeten bestimmte, von den Autoren intendierte Massnahmen der Diagnostik ab und wurden von Expertinnen und Experten validiert (vgl. Zucker & Leuchter, 2018). In Anlehnung an die Operationalisierung des Beschreibens durch Schäfer und Seidel (2015) sowie Meschede et al. (2017) beantworteten die Teilnehmenden unmittelbar nach der Sichtung der jeweiligen Testvignette aus ihrer Erinnerung heraus Items auf einer durch die Einschätzung von Expertinnen und Experten normierten Ratingskala (4 = «trifft zu», 3 = «trifft eher zu», 2 = «trifft eher nicht zu» und 1 = «trifft nicht zu» sowie 5 = «Ich weiss nicht»). Die Skala umfasste 28 Items, die sich u.a. auf von Leuders et al. (2018) genannte diagnostische Massnahmen bezogen (z.B. «Die Lehrkraft fordert die Schülerinnen und Schüler dazu auf, Vermutungen zu äussern»). Bei 12 Items war die beschriebene diagnostische Massnahme in der Testvignette zu erkennen (Zustimmungs-Items), bei 16 Items war dies nicht der Fall (Ablehnungs-Items). Zustimmungs-Items messen, ob die Studierenden eine in der Testvignette zu beobachtende diagnostische Massnahme der Lehrperson als solche beschreiben. Ablehnungs-Items messen, ob die Studierenden in der Testvignette *nicht* zu beobachtende diagnostische Massnahmen der Lehrperson zu erkennen glauben und beschreiben (z.B. «Die Lehrkraft fordert Begründungen ein», wenn sie dies jedoch gar nicht getan hat). Ablehnungs-Items zielen somit auf das Abgrenzungswissen der Studierenden.

Für die nachfolgenden Analysen wurden die Skalenwerte in einen Punkte-Score recodiert (3 Punkte für «trifft zu» bis 0 Punkte für «trifft nicht zu» bei Zustimmungs-Items und 3 Punkte für «trifft nicht zu» bis 0 Punkte für «trifft zu» bei Ablehnungs-Items). Die Angabe «Ich weiss nicht» wurde mit null Punkten gewertet. Die Reliabilität der recodierten Skala kann mit Cronbachs Alpha von $\alpha = .605$ im Prätest als akzeptabel und mit $\alpha = .721$ im Posttest als gut bezeichnet werden.

4.3 Design und Stichprobe

An der Studie nahmen $N = 109$ Grundschullehramtsstudierende im Master der Universität Koblenz-Landau, Campus Landau, teil, die vor und nach einem Seminar den in Abschnitt 4.2 beschriebenen Test bearbeiteten. Die Studierenden waren im siebten Fachsemester ($M = 7.31$, $SD = 1.66$), 23.27 Jahre alt ($SD = 2.143$) und zu 84% weiblich.

82 Studierende (im Folgenden «Interventionsgruppe» genannt) besuchten das in Abschnitt 4.1 beschriebene Seminar zur Förderung des Beschreibens prozessorientierter Diagnostik verteilt auf drei Gruppen bei unterschiedlichen Dozierenden im Wintersemester 2016/2017 bzw. im Sommersemester 2017. 27 Studierende (im Folgenden «Vergleichsgruppe» genannt) besuchten im Wintersemester 2017/2018 ein inhaltlich vergleichbares Seminar bei einem weiteren Dozierenden, in dem sie ebenfalls Lehr-Lern-Situationen planten und in einer Schulklasse umsetzten, jedoch weder fremde Text- oder Videovignetten noch eigene Videografien analysierten. Die beiden Gruppen unterschieden sich hinsichtlich der demografischen Merkmale nicht (alle $p > .05$). Um mögliche Effekte der Dozierenden zu minimieren und eine weitestgehend gleiche Durchführung in der Interventionsgruppe zu gewährleisten, verwendeten die Dozierenden die gleichen Foliensätze. Auch die Erläuterungen zu den jeweiligen Folien waren standardisiert und in einem Skript für die Dozentin bzw. den Dozenten festgehalten. Neben einer Vorbereitungssitzung vor Beginn des Semesters kamen die Dozierenden auch während des Semesters regelmässig zu Besprechungen zusammen. Ein Manipulation-Check wurde nicht durchgeführt.

4.4 Analyselogik

Zur Beantwortung der Forschungsfrage wurden in einem ersten Schritt Mittelwert-Scores über die vier Testvignetten für Prä- und Posttest in der Interventionsgruppe und in der Vergleichsgruppe gebildet. In einem zweiten Schritt wurde Varianzgleichheit als Voraussetzung für eine Mixed ANOVA geprüft. Gerade bei ungleichen Gruppengrößen wird insbesondere die Homogenität der Kovarianzenmatrizen als zentral erachtet (Field, 2011). Wie die Ergebnisse zeigten, lag weder bei den Fehlervarianzen zwischen den Gruppen (Levene-Test, einzelne $p < .05$) noch bei den Kovarianzenmatrizen (Box-Test, $p = .002$) Varianzgleichheit vor. Die Voraussetzungen für eine Mixed ANOVA waren demnach nicht erfüllt. Daher wurde für jede Gruppe einzeln eine messwiederholte Varianzanalyse (Prätest, Posttest) gerechnet, jedoch wurden die Gruppen nicht in einem gemeinsamen Modell gegenübergestellt. Dieses Vorgehen erlaubt es, die Entwicklung *innerhalb* der jeweiligen Gruppe statistisch zu überprüfen. Unterschiede *zwischen* den Gruppen können jedoch nur auf deskriptiver Ebene analysiert werden. Um einen möglichen Einfluss des Vignettentyps (Text, Video) sowie des Itemtyps (Zustimmung, Ablehnung) zu kontrollieren, wurden diese als weitere Faktoren innerhalb der jeweiligen Gruppe in die messwiederholte Varianzanalyse aufgenommen.

5 Ergebnisse

5.1 Interventionsgruppe

Die messwiederholte Varianzanalyse ergab in der Interventionsgruppe einen signifikanten Haupteffekt des Messzeitpunkts (MZP) ($F(1,81) = 6.818$, $p = .011$, partielles $\eta^2 = .078$). Entgegen der Erwartung sank die Punktzahl von $M = 1.93$ ($SD = 0.221$) im Prätest auf $M = 1.79$ ($SD = 0.294$) im Posttest. Die Studierenden zeigten nach dem

Seminar somit eine schlechtere Leistung beim Beschreiben von diagnostischen Massnahmen als zuvor. Jedoch legte ein Interaktionseffekt von MZP und Itemtyp ($F(1,81) = 22.518, p < .001$, partielles $\eta^2 = .218$) nahe, dass sich die Punktzahl bei Items, denen zuzustimmen war, anders entwickelt hatte als bei Items, die abzulehnen waren. Um diese Wechselwirkung zu untersuchen, wurden paarweise Vergleiche für Zustimmungs- und Ablehnungs-Items vorgenommen. Das Signifikanzniveau wurde nach Bonferroni auf $p < .025$ angepasst (auch für die im Folgenden beschriebenen paarweisen Vergleiche). Tabelle 1 gibt einen Überblick über die Mittelwerte. Ein Vergleich der Prätest-Werte zeigt zunächst, dass die Studierenden bei der Einschätzung von Ablehnungs-Items schlechter abschnitten als bei der Einschätzung von Zustimmungs-Items. Im Prätest-Posttest-Vergleich verringerte sich der Punktwert bei den Ablehnungs-Items signifikant ($t(81) = 5.457, p < .001, d = 0.74$). Hingegen stieg der Punktwert bei Zustimmungs-Items ($t(81) = 2.000, p = .025, d = 0.30$). Demnach verschlechterten sich die Teilnehmenden insofern, als sie in den Items beschriebene diagnostische Massnahmen zu erkennen glaubten, obwohl diese in der Testvignette nicht vorgekommen waren. Dieser Effekt trat zudem bei Videovignetten ausgeprägter hervor ($d = 0.77$) als bei Textvignetten ($d = 0.27$), wie eine signifikante Dreifachinteraktion aus MZP, Itemtyp und Vignetentyp zeigte ($F(1,81) = 8.690, p = .004$, partielles $\eta^2 = .097$). Hingegen verbesserten sich die Studierenden im Beschreiben von diagnostischen Massnahmen, die von der Lehrperson in der Testvignette tatsächlich gezeigt worden waren.

Tabelle 1: Mittelwerte der Interventions- und Vergleichsgruppe im Prätest-Posttest-Vergleich

	Interventionsgruppe (n = 109)		Vergleichsgruppe (n = 27)	
	Zustimmung (12 Items)	Ablehnung (16 Items)	Zustimmung (12 Items)	Ablehnung (16 Items)
Punktwert ^a				
$M(SD)_{prä}$	2.10 (0.324)	1.80 (0.430)	2.10 (0.464)	1.60 (0.487)
$M(SD)_{post}$	2.20 (0.320)	1.48 (0.452)	2.14 (0.452)	1.40 (0.437)
<i>t</i> -Tests				
<i>p</i> (1-seitig)	.025	< .001	.334	.018
<i>d</i> (Cohen)	0.30	0.74	0.09	0.42

Anmerkungen: Nach Bonferroni korrigiertes Signifikanzniveau liegt bei $p < .025$; ^a Werte von 0 bis maximal 3 Punkte.

5.2 Vergleichsgruppe

Die messwiederholte Varianzanalyse zeigte in der Vergleichsgruppe keinen Haupteffekt des MZP ($F(1,26) = 0.734, p = .399$, partielles $\eta^2 = .027$). Jedoch trat ein Interaktionseffekt des MZP mit dem Itemtyp auf ($F(1,26) = 4.940, p = .035$, partielles $\eta^2 = .160$). Wie in der Interventionsgruppe lag auch in der Vergleichsgruppe der im Prätest erzielte

Punktwert bei Ablehnungs-Items erstens deutlich unter dem Wert der Zustimmung-Items und verringerte sich zweitens im Prätest-Posttest-Vergleich signifikant ($t(26) = 2.206$, $p = .018$, $d = 0.42$). Für Zustimmung-Items konnte jedoch kein Anstieg des Punktwerts im Prätest-Posttest-Vergleich festgestellt werden. Wie die Studierenden der Interventionsgruppe glaubten auch die Studierenden der Vergleichsgruppe nach dem Seminar häufiger, bestimmte diagnostische Massnahmen in den Testvignetten erkannt zu haben, obwohl die Lehrpersonen diese Massnahmen im Video nicht gezeigt hatten. Die Studierenden der Vergleichsgruppe wurden aber nicht besser darin, die von der Lehrperson gezeigten diagnostischen Massnahmen zu beschreiben.

6 Diskussion

In dieser Studie wurde die Frage untersucht, inwiefern bei Grundschullehramtsstudierenden das Beschreiben von Massnahmen der Diagnostik durch ein auf Text- und Videovignetten basierendes Seminar zur prozessorientierten Diagnostik gefördert werden kann. Zusammenfassend kann diesbezüglich Folgendes festgehalten werden:

1. Im Prätest gelang es sowohl Studierenden der Interventionsgruppe als auch Studierenden der Vergleichsgruppe besser, in den Testvignetten vorkommende Massnahmen der Diagnostik zu beschreiben (Zustimmungs-Items), als zu erkennen, dass bestimmte Massnahmen nicht vorgekommen waren (Ablehnungs-Items).
2. Sowohl Studierende der Interventionsgruppe als auch Studierende der Vergleichsgruppe verschlechterten sich im Prätest-Posttest-Vergleich, wenn nach Massnahmen der Diagnostik gefragt wurde, die *nicht* in den Testvignetten vorgekommen waren (Ablehnungs-Items). Bei Studierenden der Interventionsgruppe war dieser Effekt bei Videovignetten ausgeprägter als bei Textvignetten.
3. Studierende der Interventionsgruppe verbesserten sich beim Beschreiben von Massnahmen der Diagnostik, die in den Testvignetten vorgekommen waren (Zustimmungs-Items), nicht aber Studierende der Vergleichsgruppe.

Die Ergebnisse weisen auf gegenläufige Effekte hin: Einerseits verbesserte sich bei der Interventionsgruppe das Beschreiben, wenn die im Item genannte diagnostische Massnahme in der Testvignette vorgekommen war und dementsprechend beschrieben werden konnte (Effekt 1). Dies legt den Schluss nahe, dass bei Studierenden der Interventionsgruppe das Beschreiben von diagnostischen Massnahmen durch den Einsatz von Text- und Videovignetten zumindest teilweise gefördert werden konnte. Die fortwährende Analyse fremder sowie eigener Unterrichtssituationen ging mit wiederholten Erklärungen zur prozessorientierten Diagnostik einher. Dies hat die Studierenden offenbar insoweit beeinflusst, als sie dahingehend auftretende Unterrichtsmerkmale nach dem Seminar besser beschreiben konnten als zuvor. Andererseits verschlechterte sich das Beschreiben bei allen Studierenden, wenn die im Item genannte diagnostische Massnahme nicht in der Testvignette vorgekommen war und dementsprechend nicht zu beschreiben war (Effekt 2). Offenbar gelang es den Studierenden nach dem jeweiligen

Seminar noch weniger als zuvor, in den Testvignetten vorkommende diagnostische Massnahmen der Lehrperson von nicht vorkommenden abzugrenzen. Einen ähnlichen Befund zeigte die Untersuchung von Star et al. (2011), wonach es Studierenden nach einem einsemestrigen Seminar zur professionellen Wahrnehmung nicht ausreichend gelang, in einer einstündigen Unterrichtssequenz zwischen bedeutsamen und nicht bedeutsamen Unterrichtsmerkmalen zu differenzieren. Als mögliche Erklärungen für diesen Befund erscheinen sowohl Übergeneralisierungen (Epstein, 1992) als auch False-Memory-Effekte (Deese, 1959; Roediger & McDermott, 1995) plausibel.

In der Interventionsgruppe wurden die Studierenden wiederholt zur Erklärung des Gelernten angehalten, was nach Williams et al. (2013) Übergeneralisierungen fördern kann. Die Studierenden bekamen im Seminar ein Konzept von prozessorientierter Diagnostik immer wieder anhand exemplarischer Unterrichtssituationen in Form von Text- und Videovignetten erklärt und waren aufgefordert, selbst Erklärungen zu geben. Es spricht viel dafür, dass es den an der Intervention teilnehmenden Studierenden u.a. deshalb erschwert wurde, «echte» diagnostische Massnahmen der Lehrperson (z.B. Fragen auf inhaltlicher Ebene) von anderen Massnahmen (z.B. Fragen auf organisatorischer Ebene) abzugrenzen (Oser et al., 1999). Diese Schwierigkeit zeigte sich in der Testsituation bei Videovignetten deutlicher als bei Textvignetten. Die in den Videovignetten gezeigten Lehrperson-Schulkind-Interaktionen entsprachen transkribiert durchschnittlich 2000 Zeichen gegenüber 1700 Zeichen in den Textvignetten. Dies dürfte bei der Verarbeitung der Videovignetten wie bei Syring et al. (2015) zu einer höheren kognitiven Belastung der Studierenden geführt haben. Diese Belastung hat womöglich zu einer stärker an Oberflächenmerkmalen ausgerichteten Wahrnehmung und Verarbeitung der Inhalte geführt und damit die Übergeneralisierung befördert.

Ob der Anstieg bei den Zustimmungs-Items in der Interventionsgruppe eine tatsächliche Verbesserung der Studierenden im Beschreiben von prozessorientierter Diagnostik abbildet oder aber auch einer Übergeneralisierung geschuldet ist, kann nicht gesichert beantwortet werden. Bei einer umfassenden Übergeneralisierung (jede Massnahme der Lehrperson wird als Diagnostik wahrgenommen) hätte der Anstieg jedoch deutlich grösser ausfallen müssen, als dies der Fall war. Dies spricht dafür, dass es in der Interventionsgruppe nicht nur zu Übergeneralisierungen, sondern auch zu einem verbesserten Konzept von prozessorientierter Diagnostik kam.

Zusätzlich zu den beschriebenen Übergeneralisierungen können auch False-Memory-Effekte aufgetreten sein, was die Verschlechterung bei Ablehnungs-Items in der Vergleichsgruppe erklären würde. Das in der vorliegenden Studie eingesetzte Instrument verlangte von den Studierenden, die in den Testvignetten gezeigten Unterrichtssituationen anhand geschlossener Items aus ihrer Erinnerung heraus zu beantworten. Die geschlossenen Items suggerierten bestimmte Massnahmen der Lehrperson, die im Falle der Ablehnungs-Items aber nicht zutrafen. Sowohl bei den Studierenden der Interventionsgruppe als auch bei den Studierenden der Vergleichsgruppe können in diesen Fällen

tendenziell falsche Erinnerungen ausgelöst worden sein (Loftus & Palmer, 1974). Dies würde erklären, warum Ablehnungs-Items bereits im Prätest von den Studierenden beider Gruppen deutlich seltener richtig beantwortet worden waren als Zustimmung-Items. Die Studierenden hatten zu diesem Zeitpunkt kein stabiles Konzept von prozessorientierter Diagnostik, was eine verzerrte Erinnerung aufgrund von Unsicherheit begünstigt haben dürfte. In der Vergleichsgruppe konnten die Konzepte darüber hinaus nicht verbessert werden. Die Unsicherheit scheint vielmehr noch gestiegen zu sein, was die Verschlechterung bei Ablehnungs-Items in der Vergleichsgruppe erklären würde. Die Teilnehmenden der Interventionsgruppe formten – wie bereits dargelegt – im Seminar vermutlich übergeneralisierte Konzepte von prozessorientierter Diagnostik, die eine falsche Erinnerung bei Ablehnungs-Items im Posttest womöglich noch verstärkten.

Auf inhaltlicher Ebene weisen die Ergebnisse somit darauf hin, dass das Beschreiben von diagnostischen Massnahmen durch das Seminar zur prozessorientierten Diagnostik bei den Studierenden noch nicht in erwünschter Weise gefördert werden konnte, da es zu Übergeneralisierungen kam. Auf methodischer Ebene verdeutlicht die vorliegende Studie eine potenzielle Schwierigkeit geschlossener Items bei der Messung von Teilprozessen der professionellen Wahrnehmung wie dem Beschreiben: Die im Item genannte Massnahme kann einen False-Memory-Effekt auslösen. Vor diesem Hintergrund scheinen einerseits Veränderungen im Seminarkonzept und andererseits Anpassungen im Erhebungsinstrument nötig. Um einer Übergeneralisierung entgegenzuwirken, bietet es sich an, im Seminar Texte und Videos zu verwenden, anhand derer sowohl vorhandene als auch nicht vorhandene diagnostische Massnahmen thematisiert, analysiert und verglichen werden können. Vergleichsprozesse wurden als entscheidender Mechanismus identifiziert, der das Lernen in komplexen Problemlösungsbereichen unterstützt (vgl. Alfieri, Nokes-Malach & Schunn, 2013). Dabei scheint eine Fokussierung auf stark kontrastierende Videos effizienter zu sein als eine Konzentration auf Ähnlichkeiten (Hirstein, Denn, Jurkowski & Lipowsky, 2017). In einer Folgestudie soll in einem Prä-Post-Vergleichsgruppendesign deshalb untersucht werden, ob die vergleichende Analyse fremder und eigener Lehr-Lern-Situationen lernwirksamer ist als das hier untersuchte Seminar. Um mögliche durch die Aussagen geschlossener Items induzierte False-Memory-Effekte bei der Messung der professionellen Wahrnehmung und ihrer Teilprozesse zu verringern, kann das Messinstrument um offene Items ergänzt werden (vgl. z.B. van Es & Sherin, 2008). Die Studierenden würden direkt im Anschluss an das Betrachten der Testvignette ohne vorformulierte – und womöglich suggestive – Aussagen gefragt, ob diagnostische Massnahmen seitens der Lehrperson zu erkennen waren und, wenn ja, aufgefordert, diese Massnahmen so genau wie möglich zu beschreiben.

Einschränkend muss die vergleichsweise kleine Fallzahl (und damit geringere Teststärke) der Vergleichsgruppe gegenüber der Interventionsgruppe angemerkt werden. Auch lag keine Varianzgleichheit vor, sodass Interventionsgruppe und Vergleichsgruppe

nicht als Faktor in ein varianzanalytisches Gesamtmodell (Mixed ANOVA) aufgenommen werden konnten, sondern jeweils für sich betrachtet wurden. Gruppenvergleiche konnten daher nur auf deskriptiver Ebene erfolgen. Auch konnte nicht geprüft werden, wie mögliche Übergeneralisierungen und False-Memory-Effekte zusammenwirken und welchen Anteil der jeweilige kognitive Prozess an der beobachteten Verschlechterung der Studierenden ausmachte. Dies müsste in streng kontrollierten experimentellen Laborstudien erforscht werden. Im Rahmen einer quasiexperimentellen Feldstudie wäre aber z.B. zu überprüfen, ob eine Verschlechterung auch allein durch die wiederholte Bearbeitung der Testvignetten auftreten würde. Hierfür wäre zusätzlich eine passive Kontrollgruppe nötig, die in gleichem Zeitabstand wie die Seminargruppen nur das Testinstrument durchführt. Eine wirklich passive Kontrollgruppe dürfte im universitären Kontext jedoch nur schwer zu realisieren sein.

Alles in allem zeigt die vorliegende Untersuchung sowohl das Potenzial als auch unerwartete Schwierigkeiten eines auf Text- und Videovignetten basierenden Seminars zur Förderung des Beschreibens prozessorientierter Diagnostik und der Messung der angezielten Kompetenzen bei angehenden Lehrpersonen.

Literatur

- Alferi, L., Nokes-Malach, T.J. & Schunn, C.D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, 48 (2), 87–113.
- Bell, B. & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85 (5), 536–553.
- Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Hrsg.), *International handbook of research on conceptual change* (S. 61–82). New York: Routledge.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58 (1), 17–22.
- Epstein, S. (1992). Coping ability, negative self-evaluation, and overgeneralization: Experiment and theory. *Journal of Personality and Social Psychology*, 62 (5), 826–836.
- Field, A. (2011). *Discovering statistics using SPSS*. Los Angeles: Sage.
- Gold, B., Hellermann, C., Burgula, K. & Holodynski, M. (2016). Fallbasierte Unterrichtsanalyse – Effekte von video- und textbasierter Fallanalyse auf kognitive Belastung, aufgabenspezifisches Interesse und die professionelle Unterrichtswahrnehmung von Grundschullehramtsstudierenden. *Unterrichtswissenschaft*, 44 (4), 322–338.
- Gold, B., Hellermann, C. & Holodynski, M. (2017). Effekte videobasierter Trainings zur Förderung der Selbstwirksamkeitsüberzeugungen über Klassenführung im Grundschulunterricht. *Zeitschrift für Erziehungswissenschaft*, 20 (Supplement 1), 115–136.
- Gotwals, A. W., Phillhower, J., Cisterna, D. & Bennett, S. (2015). Using video to examine formative assessment practices as measures of expertise for mathematics and science teachers. *International Journal of Science and Mathematics Education*, 13 (2), 405–423.
- Harlen, W. (2001). *Primary Science: Taking the plunge*. Portsmouth, NH: Heinemann.

- Hewson, P. & Lemberger, J. (2000). Status as the hallmark of conceptual learning. In R. Millar, J. Leach & J. Osborne (Hrsg.), *Improving science education* (S. 110–125). Buckingham: Open University Press.
- Hirstein, A., Denn, A.-K., Jurkowski, S. & Lipowsky, F. (2017). Entwicklung der professionellen Wahrnehmungs- und Beurteilungsfähigkeit von Lehramtsstudierenden durch das Lernen mit kontrastierenden Videofällen – Anlage und erste Ergebnisse des Projekts KONTRAST. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 35 (3), 472–486.
- Hoth, J., Döhrmann, M., Kaiser, G., Busse, A., König, J. & Blömeke, S. (2016). Diagnostic competence of primary school mathematics teachers during classroom situations. *ZDM – The International Journal on Mathematics Education*, 48 (1–2), 41–54.
- Krammer, K., Hugener, I., Biaggi, S., Frommelt, M., Furrer Auf der Maur, G. & Stürmer, K. (2016). Videos in der Ausbildung von Lehrkräften: Förderung der professionellen Unterrichtswahrnehmung durch die Analyse von eigenen bzw. fremden Videos. *Unterrichtswissenschaft*, 44 (4), 357–372.
- Krammer, K. & Reusser, K. (2005). Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 23 (1), 35–50.
- Leuders, T., Dörfler, T., Leuders, J. & Philipp, K. (2018). Diagnostic competence of mathematics teachers – Unpacking a complex construct. In T. Leuders, K. Philipp & J. Leuders (Hrsg.), *Diagnostic competence of mathematics teachers – Unpacking a complex construct in teacher education and teacher practice* (S. 3–31). Cham: Springer.
- Loftus, E. F. & Palmer, J. C. (1974). Reconstruction of automobile destruction. An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13 (5), 585–589.
- Meschede, N. (2014). *Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht*. Berlin: Logos.
- Meschede, N., Fiebranz, N., 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.
- Meschede, N., Steffensky, M., Wolters, M. & Möller, K. (2015). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht. Theoretische Beschreibung und empirische Erfassung. *Unterrichtswissenschaft*, 43 (4), 317–335.
- Möller, K. (2010). Lernen von Naturwissenschaft heisst: Konzepte verändern. In P. Labudde (Hrsg.), *Fachdidaktik Naturwissenschaft. 1.–9. Schuljahr* (S. 57–72). Bern: Haupt.
- Morrison, J. A. & Lederman, N. G. (2003). Science teachers' diagnosis and understanding of students' preconceptions. *Science Education*, 87 (6), 849–867.
- Oser, F., Hascher, T. & Spychiger, M. (1999). Lernen aus Fehlern. Zur Psychologie des «negativen» Wissens. In W. Althof (Hrsg.), *Fehlerwelten. Vom Fehlermachen und Lernen aus Fehlern. Beiträge und Nachträge zu einem interdisziplinären Symposium aus Anlaß des 60. Geburtstags von Fritz Oser* (S. 11–41). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Rawson, K. A. & Kintsch, W. (2004). Exploring encoding and retrieval effects of background information on text memory. *Discourse Processes*, 38 (3), 323–344.
- Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21 (4), 803–814.
- 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.
- Seidel, T. & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in pre-service teachers. *American Educational Research Journal*, 51 (4), 739–771.
- Star, J. R., Lynch, K. & Perova, N. (2011). Using video to improve preservice mathematics teachers' abilities to attend to classroom features. In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Hrsg.), *Mathematics teacher noticing. Seeing through teachers' eyes* (S. 117–133). New York: Routledge.
- Sunder, C., Todorova, M. & Möller, K. (2016). Förderung der professionellen Wahrnehmung bei Bachelorstudierenden durch Fallanalysen. Lohnt sich der Einsatz von Videos bei der Repräsentation der Fälle? *Unterrichtswissenschaft*, 44 (4), 339–356.

- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4 (4), 295–312.
- Syring, M., Bohl, T., Kleinknecht, M., Kuntze, S., Rehm, M. & Schneider, J. (2015). Videos oder Texte in der Lehrerbildung? Effekte unterschiedlicher Medien auf die kognitive Belastung und die motivational-emotionalen Prozesse beim Lernen mit Fällen. *Zeitschrift für Erziehungswissenschaft*, 18 (4), 667–685.
- van Es, E. A. & Sherin, M. G. (2008) Mathematics teachers' «learning to notice» in the context of a video club. *Teaching and Teacher Education*, 24 (2), 244–276.
- von Aufschnaiter, C., Selter, C. & Michaelis, J. (2017). Nutzung von Vignetten zur Entwicklung von Diagnose- und Förderkompetenzen. Konzeptionelle Überlegungen und Beispiele aus der MINT-Lehrerbildung. In C. Selter, S. Hußmann, C. Höfle, C. Knipping & K. Lengnink (Hrsg.), *Diagnose und Förderung heterogener Lerngruppen. Theorien, Konzepte und Beispiele aus der MINT-Lehrerbildung* (S. 85–105). Münster: Waxmann.
- Vosniadou, S., Vamvakoussi, X. & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Hrsg.), *International handbook of research on conceptual change* (S. 3–34). New York: Routledge.
- Williams, J. J., Lombrozo, T. & Rehder, B. (2013). The hazards of explanation. Overgeneralization in the face of exceptions. *Journal of Experimental Psychology: General*, 142 (4), 1006–1014.
- Zucker, V. & Leuchter, M. (2018). Die Fähigkeit von Studierenden im Grundschullehramt, Formative Assessment im naturwissenschaftlichen Sachunterricht zu erkennen. In U. Franz, H. Giest, A. Hartinger, A. Heinrich-Dönges & B. Reinthoffer (Hrsg.), *Handeln im Sachunterricht* (S. 175–182). Bad Heilbrunn: Klinkhardt.

Autor und Autorinnen

Timo Reuter, Dr., Universität Koblenz-Landau, reutertimo@uni-landau.de
Verena Zucker, Universität Koblenz-Landau, zucker@uni-landau.de
Miriam Leuchter, Prof. Dr., Universität Koblenz-Landau, leuchter@uni-landau.de