



Professionswissen angehender Lehrkräfte zum Konstrukt *Nature of Science* (NOS): Entwicklung und Validierung eines Vignettentests (EKoL-NOS)

Tim Billion-Kramer^{1,2} · Hendrik Lohse-Bossenz¹ · Tobias Dörfler¹ · Markus Rehm^{1,2}

Eingegangen: 28. Mai 2019 / Angenommen: 31. März 2020 / Online publiziert: 16. April 2020
© Der/die Autor(en) 2020

Zusammenfassung

Unter der Bezeichnung *Nature of Science* (NOS) werden in der Naturwissenschaftsdidaktik Grundlagen der Erkenntnisgewinnung sowie Eigenschaften naturwissenschaftlichen Wissens diskutiert. Zentrales fachdidaktisches Anliegen ist, Lernende mit Erkenntnisprozessen naturwissenschaftlicher Disziplinen vertraut zu machen, diese aber auch kritisch zu hinterfragen. Im Fokus des Beitrags stehen die Ausprägungen des Professionswissen von Lehrkräften zum Bereich NOS und dessen situierte Erfassung. Bisher fehlen quantitativ messende Instrumente, die Professionswissen von Lehrkräften zu *Nature of Science* in Unterrichtskontexten und somit in einem möglichst realitätsnahen Bezugsrahmen erheben. Der Vignettentest EKoL-NOS soll einen Beitrag zur Bearbeitung dieser Lücke leisten. Im Test werden den Lehrkräften kurze authentische Beschreibungen von Unterrichtssituationen mit verschiedenen Handlungsalternativen präsentiert. Die Lehrkräfte bewerten die situationsspezifische Angemessenheit der Handlungsalternativen. Der vorliegende Beitrag beschreibt die theoriegestützte Entwicklung, die Dimensionalitätsprüfung auf Basis dreier zentraler *Nature of Science*-Facetten in Form von Strukturgleichungsmodellen sowie eine empirische Validierung anhand demographischer Daten, diskriminanter und konvergierender Konstrukte. Mit einer Stichprobe von $N=289$ angehenden Lehrkräften im Vorbereitungsdienst weist sowohl ein dreidimensionales Modell (je *Nature of Science*-Facette ein Faktor), als auch ein eindimensionales Modell eine sehr gute Anpassungsgüte auf. Ein Vergleich zwischen beiden Modellen zeigt keinen signifikanten Unterschied. Demographische Variablen (z. B. die kognitive Leistungsfähigkeit) lassen sich weitgehend erwartungskonform in Strukturgleichungsmodellen spezifizieren. Mit einer zweiten Stichprobe von $N=137$ wird ein Strukturgleichungsmodell mit einem diskriminanten und konvergierenden Konstrukt spezifiziert.

Schlüsselwörter Nature of Science · Vignettentest · Lehrerbildung · Professionalisierungsforschung

✉ Tim Billion-Kramer
kramer@ph-heidelberg.de

¹ Pädagogische Hochschule Heidelberg,
Keplerstr. 87, 69120 Heidelberg, Deutschland

² Klaus-Tschira-Kompetenzzentrum für frühe
naturwissenschaftliche Bildung, Forscherstation gGmbH,
Speyerer Str. 6, 69115 Heidelberg, Deutschland

Prospective Teachers' Professional Knowledge About *Nature of Science* (NOS)—Development and Validation of a Vignette Test (EKoL-NOS)

Abstract

Nature of Science (NOS) is the element of science didactics which discusses how science functions and how scientific knowledge is generated. The central goal is to familiarize learners with processes of gaining knowledge in scientific disciplines, but also to critically question them. This article focuses on professional knowledge of teachers about NOS and its situational testing. Up to now, to measure professional knowledge of teachers on Nature of Science in teaching contexts is a lack of quantitative instruments. To help to fill this gap is the intention of the test presented in this article. The test presents short authentic descriptions of teaching situations in the form of vignettes with various alternatives for instruction. These alternatives have to be rated by their situation-specific appropriateness. This paper describes the theory-based development, the dimensionality test based on three central Nature of Science facets in the form of structural equation models, and an empirical validation based on socio cognitive expectations, discriminative and convergent constructs. With a sample of $N = 289$ prospective teachers, both a three-dimensional model (one factor per Nature of Science facet) and a one-dimensional model are fitting well. A comparison between the two models shows no significant difference. Demographic variables (e.g. cognitive performance) can be specified in line with expectations; as well as discriminant and converging constructs with another sample ($N = 137$).

Keywords Nature of Science · Vignette test · Teacher formation · Professionalization research

Einleitung

Ein angemessenes Wissenschaftsverständnis wird als Teil der Allgemeinbildung innerhalb von Naturwissenschaft und Technik geprägter Gesellschaftsformen und als wichtige Voraussetzung gesellschaftlicher Teilhabe verstanden (Menthe et al. 2013; Arndt et al. 2019). Am entsprechenden bildungswissenschaftlichen Diskurs beteiligen sich die Allgemeine Didaktik (z. B. Klafki 2007), die Pädagogische Psychologie (z. B. Bromme und Kienhues 2008) sowie die an einem pragmatischen (z. B. Shen 1975) wie klassischen (z. B. Hodson 2011) Begriff von *scientific literacy* orientierten naturwissenschaftsdidaktischen Konzeptionen (vgl. Fischler et al. 2018; Sjöstrom und Eilks 2018). Es wird deutlich, dass ein adäquates Wissenschaftsverständnis Teil einer naturwissenschaftlichen Grundbildung (*scientific literacy*) ist. Für die Formulierung von Bildungszielen für den Erwerb einer naturwissenschaftlichen Grundbildung, die damit auch das Wissenschaftsverständnis der Schülerinnen und Schüler als Zielebene beinhalten, stehen seit 15 Jahren die Bildungsstandards der Kultusministerkonferenz (KMK 2005a) oder in den USA seit längerem entsprechende National Science Education Standards (National Research Council 1996) und auf das Wissenschaftsverständnis zielende Benchmarks (NGSS 2013). Doch Zielsetzungen und Diskurse über ein angemessenes Wissenschaftsverständnis bleiben bezogen auf den Schulunterricht derzeit noch wirkungsschwach und werden in der Unterrichtspraxis häufig ignoriert (Bartos und Lederman 2014). Der Ansatzpunkt zum Erwerb eines adäquaten Wissenschaftsverständnisses wäre ein wirksames Angebot und die effektive Nutzung von Lerngelegenheiten (Helmke

2012; Seidel 2014; Wilhelm et al. im Druck). Eine Schlüsselrolle auf der Seite des Angebots scheint die Lehrkraft zu spielen (Lipowsky 2006), die einerseits selbst ein adäquates Wissenschaftsverständnis benötigt und andererseits zu dessen Vermittlung Professionswissen aufbauen muss, das alle naturwissenschaftlichen Disziplinen überspannt (Rehm et al. 2016). So wird in der Naturwissenschaftsdidaktik Professionswissen zum Wissenschaftsverständnis unter der Bezeichnung *Nature of Science* (NOS) naturwissenschaftsübergreifend diskutiert und gelehrt (z. B. Lederman 2007; Höttecke und Henke 2010; McComas 2014; Allchin 2017). Es beinhaltet ein reflektiertes Verständnis über die Rolle der Naturwissenschaften im gesellschaftlichen Kontext sowie über Eigenschaften naturwissenschaftlicher Erkenntnis. Für die Seite der Nutzung von Lerngelegenheiten stehen die oben angesprochenen Bildungsstandards.

In diesem Beitrag fokussieren wir den Bereich NOS auf der Seite des Angebots von Lerngelegenheiten und thematisieren ausgewählte berufliche Anforderungen an Lehrkräfte. Diese beruflichen Anforderungen an Lehrkräfte ergeben sich in unterschiedlichen NOS-Facetten, von denen wir drei in den engeren Fokus nehmen und deren Auswahl vor dem theoretischen Hintergrund des Konstrukts NOS explizit begründen. Diese Facetten nehmen die folgenden Aspekte des Konstrukt NOS in den Blick: *die Bedeutung von Kreativität und Subjektivität für naturwissenschaftliche Erkenntnis, die Veränderung naturwissenschaftlicher Erkenntnisse im Laufe der Zeit sowie die Differenzierung zwischen Beobachtung und Schlussfolgerung im naturwissenschaftlichen Erkenntnisprozess*. Ziel des Beitrags ist es, die Entwicklung eines Testformats darzustellen, das Professionswissen zu diesen NOS-Facetten bei angehenden Lehrkräften erheben kann.

Theoretischer Hintergrund

Nature of Science

Nature of Science (NOS) gilt fachdidaktisch als naturwissenschaftsübergreifendes Konstrukt (z. B. Schwarz und Lederman 2008). Frühe Unterrichtsbemühungen und Untersuchungen zu NOS konzentrierten sich in den USA der 50er und 60er Jahre auf Erkenntniswege wissenschaftlichen Wissens, also *Scientific Inquiry* (z. B. Welch und Pella 1967). Später traten Aspekte der Wissenschaftsphilosophie, -historie und -soziologie hinzu, z. B. im Sinne von Popper (1935) oder Kuhn (1962). Zunehmend bekommt das Wissen und Reflektieren über NOS als Teil naturwissenschaftlicher Bildung auch im deutschsprachigen Raum Gewicht (z. B. KMK 2005a, 2005b, 2005c; Höttecke 2001).

In aktuellen Ansätzen beinhaltet NOS ein reflektiertes Verständnis der Rolle der Naturwissenschaften im gesellschaftlichen, technischen, ökologischen und ansatzweise historischen Kontext sowie Eigenschaften naturwissenschaftlicher Erkenntnis (z. B. Gebhard et al. 2017; Allchin 2017). Mit der Reflexion über NOS nimmt naturwissenschaftlicher Unterricht eine grundlegende allgemeinbildende Funktion wahr: Schülerinnen und Schüler sollen über Kenntnisse naturwissenschaftlicher Inhalte und Methoden verfügen, damit sie Möglichkeiten und Grenzen naturwissenschaftlichen Wissens und Handelns reflektieren können. Dabei beinhaltet diese Reflexion nicht nur naturwissenschaftliches Fachwissen, sondern auch erkenntnis- und wissenschaftstheoretisches sowie wissenschaftshistorisches und -soziologisches Wissen (Höble et al. 2004), für einen handelnden Umgang mit Naturwissenschaft im Sinne einer gesellschaftlichen Teilhabe und Meinungsbildung. Informierte Bürgerinnen und Bürger, die naturwissenschaftlichen Unterricht erhalten haben, sollten sich beispielsweise mit sogenannten socio-scientific issues, kurz SSIs (Zeidler 2015) auseinandersetzen können, also gesellschaftlich bedeutsamen und naturwissenschaftlich relevanten Problemen, wie dem Klimawandel oder dem Klonen von Tieren (Gebhard et al. 2017).

Empirisch zeigt sich jedoch, dass die Unterrichtserfolge bei den Schülerinnen und Schülern in diesem Bereich hinter dem fachdidaktischen, bzw. gesellschaftlich-politischen Anspruch zurückbleiben (Deng et al. 2011). Es zeigt sich wenig Bewusstsein für eine soziale und in Institutionen verankerte Praxis, in der Experten miteinander aushandeln, was sie als gültiges Wissen erachten. Eine Praxis, die (auch) von gegenseitiger Kontrolle geprägt ist. In der anglo-amerikanischen Naturwissenschaftsdidaktik wurden auf Basis dieser Befundlage unterschiedliche Vorschläge gemacht, um das Lernen über NOS zu strukturieren, woraus sich bestimmte berufliche Anforderungen an das fachdidaktische Wissen und Können von Lehrkräften ergeben.

Insbesondere wurde das Konstrukt NOS elementarisiert, um die Komplexität des Themas für Lernprozesse besser greifbar, und vor allem vermittelbar zu machen. Eine prominente Elementarisierungsmaßnahme stellen so genannte Aspektlisten dar.

NOS konzeptualisiert als Minimalkonsens in Aspektlisten

Seit etwa zwei Jahrzehnten werden in der Naturwissenschaftsdidaktik Listen diskutiert, welche NOS-Eigenschaften Lernende, aber auch Lehrkräfte kennen und reflektiert haben sollten (z. B. McComas 1998; Osborne et al. 2003; Lederman 2007; kritisch dazu z. B. Irzik und Nola 2011; Erduran und Dagher 2014; Allchin 2017). Die Strukturierung bzw. Elementarisierung von NOS ergibt sich aus einer fachlichen Analyse der Naturwissenschaften unter Berücksichtigung der Schülerperspektive. So untersuchten McComas und Olson (1998) in einer Textanalyse bildungspolitische Dokumente (Standards und Curricula) aus englischsprachigen Ländern auf Bildungsinhalte zu NOS. Außerdem befragten Osborne et al. (2003) in einer Delphi-Studie 23 Experten aus unterschiedlichen Fachgebieten von Fachdidaktikerinnen und Fachdidaktikern über Philosophinnen und Philosophen sowie Wissenschaftssoziologinnen und -soziologen bis hin zu Lehrkräften zu wünschenswerten Unterrichtsinhalten zu (1) naturwissenschaftlichen Methoden, (2) Eigenschaften naturwissenschaftlichen Wissens, (3) Organen und sozialen Aspekten der Naturwissenschaften. Als Ergebnis präsentierten sie zehn Aussagen, die als Konsens zwischen den Experten gewertet werden können. Lederman (2006) prüfte NOS-Aspekte im Blick auf die drei folgenden Fragestellungen:

- Ist das Wissen zu dem NOS-Aspekt für Schülerinnen und Schüler zugänglich (können sie es lernen und verstehen)?
- Besteht ein genereller Konsens über den NOS-Aspekt?
- Ist das Verständnis des NOS-Aspekts für Bürgerinnen und Bürger nützlich?

Lederman et al. (2013, S. 140) kommen durch Prüfung der Fragen zu folgendem Minimalkonsens:

scientific knowledge is tentative (subject to change), empirically-based (based on and/or derived from observations of the natural world), subjective (theory-laden), necessarily involves human inference, imagination, and creativity (involves the invention of explanations), and is socially and culturally embedded. Two additional important aspects are the distinction between observations and inferences, and the functions of, and relationships between scientific theories and laws.

Synopsen der (Minimal-)Konsenslisten haben z. B. Osborne et al. (2003, S. 713), Neumann (2011, S. 26), sowie Neumann und Kremer (2013, S. 215) zusammengestellt. Grundsätzlich versuchen diese (Minimal-)Konsenslisten das Konstrukt *NOS* für die Schule angemessen zu explizieren und zugänglich zu machen. Gleichwohl wird in der Naturwissenschaftsdidaktik inzwischen deutliche Kritik an der Elementarisierung von *NOS* in Form dieser Konsenslisten geäußert, da sie (1) die aktuelle Praxis der Naturwissenschaften, (2) mögliche Erkenntnisfehler und (3) die kulturelle Dynamik naturwissenschaftlicher Erkenntnis unzureichend abbilden würden (z. B. Allchin 2018; Hodson und Wong 2017; Erduran und Dagher 2014). Allchin und Werth (2017) sehen im transportierten Wissenschaftsverständnis der Konsenslisten die Gefahr, implizit einen naiven Szientismus zu vermitteln. Wissenschaft werde als isoliertes Vorhaben dargestellt, das unabhängig von Fragen sozialer Gerechtigkeit und der Gesellschaft agiere. Bereits die Bezeichnung „Nature of Science“ lege die Fehlvorstellung von Naturwissenschaft als „natürliches“ (und damit von gesellschaftlichen Einflüssen und Abhängigkeiten isoliertes) Unterfangen nahe. Um die gesellschaftliche Eingebundenheit wissenschaftlicher Erkenntnis stärker ins Zentrum von *NOS* zu rücken, schlägt Allchin eine begriffliche und inhaltliche Erweiterung des *NOS*-Konzepts „[f]rom Nature of Science to Nature of Science in Society“ (Allchin 2018, S. 1) vor.

Mit der Alternativkonzeption „Nature of (whole) science“ hat Allchin (2011, S. 524) zudem auf die Gefahr hingewiesen, dass das Konstrukt *NOS* durch bisherige Elementarisierungen in einzelne Facetten zergliedert werde und damit der für die Entwicklung einer *scientific literacy* notwendige ganzheitliche Charakter wissenschaftlicher Erkenntnisprozesse verloren gehe:

Whole Science, like whole food, does not exclude essential ingredients. It supports healthier understanding. Metaphorically, educators must discourage a diet of highly processed, refined “school science.” Short lists of *NOS* features should be recognized as inherently incomplete and insufficient for functional scientific literacy

Matthews (2012) plädiert in Allchins Sinne dafür, in klassischen *NOS*-Elementarisierungen den Begriff „nature“ durch „features“ zu ersetzen, da das Konstrukt nicht die *nature*, sondern die *features of science* (= *FOS*) abbilde. Normativ erwarten auch die US-amerikanischen Bildungsstandards explizit ein ganzheitliches Wissenschaftsverständnis in Allchins Sinne: „students should develop an understanding of the enterprise of science as a whole – the wondering, investigating, questioning, data collecting and analyzing“ (NGSS Lead States 2013, S. 1). Allchin (2018, S. 2) kommt zu dem Schluss: „Ironically, the ‘consensus’ list no longer enjoys a very wide consensus“. Die Kritik

weist damit auf wichtige Limitierungen traditioneller *NOS*-Konzeptualisierungen hin. Grundsätzlich bieten Konsenslisten von *NOS*-Facetten jedoch eine erste Orientierung, um den weiten Begriff von *NOS* greifbar zu machen. Sie bilden den Ausgangspunkt, um Unterrichts- und Testmaterialien zu entwickeln (Neumann und Kremer 2013). Die kritisierte Zergliederung von *NOS* in Einzelaspekte durch Konsenslisten lässt sich zudem bisher nicht empirisch belegen. Empirische Analysen deuten tendenziell eher in eine entgegengesetzte Richtung und zeigen keine Überlegenheit mehrdimensionaler Modelle von *NOS*-Facetten gegenüber einem eindimensionalen Modell (Neumann 2011; Neumann et al. 2011). Zunächst sollten traditionelle Konzeptualisierungen von *NOS* somit nicht ersetzt, sondern im Sinne aktueller Kritik erweitert werden. In diesem Sinne könnten bereits enthaltene Facetten wie die soziale und kulturelle Einbettung der Naturwissenschaften („is socially and culturally embedded“, Lederman et al. 2013, S. 140) stärker betont und ausdifferenziert werden.

Ein stärkerer Fokus auf soziale und kulturelle Aspekte naturwissenschaftlicher Erkenntnisprozesse erscheint auch vor dem Hintergrund aktuell diskutierter Neukonzeptualisierungen des Konstrukts *NOS* sinnvoll. Der *Family Resemblance Approach* (FRA) von Erduran und Dagher (2014) versucht das Konstrukt *NOS* nicht über eine Minimalkonsensliste zu explizieren, sondern in Anlehnung an Irzik und Nola (2011) sowie Wittgenstein (2003 [1953]) über Eigenschaften abzubilden, die nicht für alle naturwissenschaftlichen Disziplinen und Herangehensweisen vollkommen gleich sind, zwischen denen aber Ähnlichkeiten bestehen – so wie zwischen den Mitgliedern einer Familie. Für die Vielfalt naturwissenschaftlicher Praktiken, Methoden, Ziele und Werte definieren sie zwei Hauptdimensionen: Wissenschaft als kognitiv-epistemisches System und Wissenschaft als sozial-institutionelles System. Die beiden zentralen Systeme werden im FRA-Ansatz weiterausdifferenziert, gleichwohl werden die Dimensionen und ihre Subfacetten aber auch hier explizit als nicht voneinander trennbar, sondern als enges miteinander verflochtenes Netzwerk verstanden. Alle drei *NOS*-Ansätze – Minimalkonsens, Nature of (Whole) Science, FRA – verstehen ihre Konzeptualisierung von *NOS* somit als ganzheitlich. Für eine erste Version des Instruments erscheint gleichwohl eine Konzentration auf wesentliche Aspekte von *NOS* sinnvoll, die für das Gesamtkonstrukt *NOS* charakteristische Aspekte herausgreifen. Diese charakteristischen *NOS*-Aspekte konzentrieren sich in Erduran und Daghers Terminologie auf die kognitiv-epistemische *NOS*-Dimension und berücksichtigen insbesondere subjektive Einflüsse auf wissenschaftliches Wissen. Die soziale und kulturelle Einbettung (Lederman et al. 2013) bzw. Wissenschaft als sozial-institutionelles System (Erduran und Dagher 2014) soll in einem späteren Schritt

der Instrumentenentwicklung angegangen werden. Zu den ersten Facetten im Einzelnen:

a) Wissen über die Bedeutung von Kreativität und Subjektivität für wissenschaftliche Erkenntnis (KS)

Als von Menschen gestaltetes Vorhaben ist der Wissenschaftsprozess nicht nur objektiv, sondern beinhaltet auch subjektive und kreative Komponenten. Es steht außer Frage, dass Forschende logisches Denken beherrschen und über breites theoretisches Fachwissen verfügen müssen. Theorien und Erklärungen lassen sich jedoch nicht mechanistisch oder nur rational aus Beobachtungen oder Annahmen über die physische Welt ableiten. Es bedarf subjektiver Entscheidungen und der Vorstellungskraft und Kreativität der Forschenden, um Hypothesen, Fragestellungen und Versuchsdesigns zu entwickeln oder in Untersuchungen gewonnene Daten in (neue) Hypothesen, Theorien oder wissenschaftliche Gesetze zu überführen. Auch die Diskussion und Interpretation von (teils unerwarteten) Ergebnissen stellt einen kreativen Akt dar (McComas 2014; Lederman et al. 2013).

b) Wissen über die Veränderung wissenschaftlicher Erkenntnisse im Laufe der Zeit (VZ)

Wissenschaftliches Wissen ist offen für Veränderungen und kann nicht abschließend bewiesen werden. Es stellt stets nur Annahmen über natürliche Phänomene dar. Eine wissenschaftliche Theorie kann in seiner Genese beispielsweise als subjektive Idee einer Forscherin oder einer Forschergruppe beginnen und sich in der *scientific community* durch theoretische Begründungen und empirische Evidenzen bewähren. Grundsätzlich kann wissenschaftlich gewonnenes Wissen widerlegt oder verändert werden (Popper 1935). Und auch Paradigmen entwickeln sich weiter oder werden gegebenenfalls ersetzt. Hinzu kommt, dass auch neue Erhebungsmethoden und neue oder verbesserte Instrumente zu neuen Erkenntnissen führen können (McComas 2014; Lederman et al. 2013).

c) Wissen über die Differenzierung zwischen Beobachtung und Schlussfolgerung (BS)

Im wissenschaftlichen Erkenntnisprozess wird streng differenziert zwischen Beobachtungen einerseits sowie Erklärungen und Schlussfolgerungen andererseits. Beobachtungen stellen Beschreibungen natürlicher Phänomene dar, die durch menschliche Sinneswahrnehmung oder technische Möglichkeiten zugänglich und intersubjektiv nachvollziehbar sind, beispielsweise das Fallen von Gegenständen in Richtung Erde. Im Gegensatz dazu sind Schlussfolgerungen den menschlichen Sinnen nicht „direkt“ zugänglich. Dass Gegenstände aufgrund von Gravitation zu Boden fallen, beruht auf Annahmen, nicht auf sinnlich wahrnehmbarer Beobachtung (Lederman et al. 2013).

Unbestritten ist dagegen, dass auch Beobachtungen subjektive Anteile zugrunde liegen, bedingt z. B. durch das Vorwissen der beteiligten Forschenden oder durch die Konzeption des Beobachtungsdesigns. Eine weitere wichtige Rolle spielen Überzeugungen der Wahrnehmenden. Zu fordern sind dies die Überzeugungen über die Fragen nach dem zustande kommen naturwissenschaftlicher Erkenntnisse und welche Voraussetzungen hierfür eine Rolle spielen. Diese Überzeugungen werden von der psychologisch-ausgerichteten Lehr-Lernforschung mit dem Terminus *epistemische Überzeugungen* umschrieben (z. B. Schommer 1990; Hofer und Pintrich 1997; Bromme et al. 2016); in Bezug auf die Auswahl eines Konstrukts zu konvergierenden Validierung werden wir später darauf zurückkommen.

NOS als Facette des Professionswissens von Lehrkräften der Naturwissenschaften

Professionelle Kompetenz von Lehrkräften stellt eine wichtige Grundlage guten Unterrichtens dar (Baumert und Kunter 2011) und wird in unterschiedlichen Modellen konzeptualisiert (z. B. Baumert und Kunter 2011; Seidel und Stürmer 2014; Blömeke et al. 2015). Das innerhalb der COACTIV-Studie entwickelte Modell zur professionellen Kompetenz von Lehrkräften hat weitgehende Verbreitung gefunden (Krauss et al. 2017; Rutsch et al. 2017a). Nach Baumert und Kunter (2011) stellt das Professionswissen in diesem Rahmen eine zentrale Kompetenzfacette dar und wird im Kern mit den von Shulman (1986) beschriebenen Kategorien „CK“ (*content knowledge*=Fachwissen), „PCK“ (*paedagogical content knowledge*=fachdidaktisches Wissen) und „PK“ (*pedagogical knowledge*=pädagogisches Wissen) ausdifferenziert. Das in der Lehrerbildung erworbene Professionswissen zeigt – umgesetzt in prozedurale Routinen – einen Einfluss auf die Qualität des Unterrichtshandelns (Lipowsky 2006; Voss et al. 2011). Dieses zeigt sich wiederum als Prädiktor der Lernleistungen (z. B. Boyd et al. 2009; Kersting et al. 2012; Kunter et al. 2013; Tröger et al. 2014) und der Motivation der Lernenden (Krapp und Prenzel 2011).

Die Umsetzung des in der Lehrerbildung erworbenen Professionswissens in professionelles Unterrichtshandeln stellt allerdings einen komplexen Prozess dar. Es wird angenommen, dass prozedurale Routinen unter anderem auf dem handlungsnahen Professionswissen einer Lehrkraft beruhen (Stender 2014). Für Lehrkräfte naturwissenschaftlicher Fächer beinhaltet das Professionswissen u. a. auch Aspekte eines angemessenen Wissenschaftsverständnisses und einer angemessenen Wissenschaftsvermittlung (z. B. Bartos und Lederman 2014). Dieser Bereich ist mit jeweils eigenen Anteilen sowohl im CK als auch im PCK der Lehrkräfte naturwissenschaftlicher Fächer zu verorten: Professionswissen zu naturwissenschaftlichen Erkenntnisprozessen

stellt zunächst einen Teil des CK dar. Shulman (1986, S. 9) geht mit seiner Konzeption des CK über ein bloßes Fachwissen von Fakten und Konzepten hinaus und erwartet auch ein angemessenes Wissenschaftsverständnis im Sinne der syntaktischen Struktur einer Disziplin:

The syntactic structure of a discipline is a set of ways in which truth or falsehood, validity or invalidity, are established. [...] The teacher need not only to understand that something is so, the teacher must further understand why it is so, on what grounds its warrant can be asserted, and under what circumstances our belief in its justification can be weakened or even denied.

Diese Art des Fachwissens mediiert fachdidaktisches Wissen und Können (Baumert und Kunter 2011). Das fachdidaktische Wissen und Können verstanden als PCK beinhaltet (z. B. nach Kunter et al. 2013) im jeweiligen Fach unterrichtstypische Merkmale wie Aufgabenwissen, Erklärungswissen und Schülervorstellungen. Für den Bereich NOS bedeutet das: Lehrkräfte stehen im naturwissenschaftlichen Unterricht der berufsspezifischen Anforderung gegenüber, die Alltagsvorstellungen der Schülerinnen und Schüler zum Wissenschaftsverständnis aufzugreifen und im Verlauf des Unterrichtes sowie im Vorfeld, bei der Unterrichtsplanung adaptiv zu berücksichtigen (Feige et al. 2017). Bartos und Lederman (2014, S. 1151) sprechen von „teachers’ subject matter knowledge structures for NOS [...] communicated in their classroom practice“. Somit kann NOS im Professionswissen von Lehrkräften naturwissenschaftlicher Fächer verortet werden. Der Anforderungsbereich NOS stellt dabei eine besondere Herausforderung dar, denn er beinhaltet sowohl Teile des CKs (z. B. Wissen über die Struktur und Herangehensweise der Naturwissenschaften) als auch Teile des PCKs (z. B. Alltagsvorstellungen zu Naturwissenschaften; vgl. z. B. McComas 1998).

Professionswissen von Lehrkräften wird grundsätzlich als erlern- und veränderbar verstanden (Baumert und Kunter 2006). Das Konstrukt NOS in professionelles Unterrichtshandeln umzusetzen und Lernenden ein angemessenes Wissenschaftsverständnis zu vermitteln, stellt einen anspruchsvollen Schritt in Richtung Unterrichtswirksamkeit dar. Lehrkräfte sollten deshalb in diesem Bereich explizit ausgebildet werden und sowohl über CK und PCK zu NOS verfügen, da angenommen wird, dass die Wirksamkeit des Unterrichtes durch die explizite Ausbildung von Lehrkräften in diesem Bereich gesteigert werden kann.

Stand der Forschung zu NOS: Wissen und Können von Lehrkräften

Auch wenn NOS in der englischsprachigen Naturwissenschaftsdidaktik seit mehr als 100 Jahren diskutiert wird (vgl. z. B. Central Association of Science and Mathematics Teachers 1907), fachdidaktisch seit langem als wichtiger Inhalt naturwissenschaftlichen Unterrichts gilt (z. B. Driver et al. 1996; Lederman 2007) und seit fast 30 Jahren Bestandteil von Standards und Benchmarks ist (AAAS 1993; National Research Council 1996), zeigt die internationale Forschungslage derzeit ein unzureichendes Professionswissen bei Lehrkräften in Bezug auf NOS. Fachdidaktische Konzeptualisierungen von NOS finden sich häufig nicht in Wissenschaftsvorstellungen von (angehenden) Lehrkräften wieder (z. B. Carey und Strauss 1970; Liang et al. 2006). Weitere Befunde zeigen, dass Lehrkräfte mit angemessenen Vorstellungen zu NOS diese nicht automatisch in ihre Unterrichtspraxis integrieren bzw. übertragen (z. B. Bartos und Lederman 2014). Zudem betrachten Lehrkräfte der Naturwissenschaften NOS im Vergleich zu „konkreten“ naturwissenschaftlichen Konzepten, wie beispielsweise „Bewegung“ oder „Evolution“, nicht als gleichwertiges Ziel naturwissenschaftlicher Bildung (Lederman und Lederman 2014). Insgesamt zeigt die Befundlage anhand weniger Studien ein mangelndes Wissen und Können bei Lehrkräften für den naturwissenschaftlichen Unterricht. Dies überrascht insbesondere in Anbetracht der langen englischsprachigen fachdidaktischen Diskussion mit entsprechender Verankerung in US-amerikanischen Bildungsstandards und Benchmarks. Die Herausforderung, Lernenden an Schule und Hochschule ein adäquates NOS-Verständnis zu vermitteln, scheint somit außergewöhnlich anspruchsvoll.

Die zitierten Evidenzen beruhen allerdings auf Untersuchungen in den USA, China und der Türkei. Dortige Lehrbiographien und Studieninhalte unterscheiden sich von Deutschland, Österreich und der Schweiz. Es stellt sich somit die Frage, ob das fachdidaktische Studium in deutschsprachigen Ländern zu besseren Ergebnissen führt, schließlich wird NOS auch in der deutschsprachigen Naturwissenschaftsdidaktik seit langem zentral diskutiert (z. B. Höttecke 2001; Sodian et al. 2002; Höhle et al. 2004; Neumann und Kremer 2013; Gebhard et al. 2017). Dies könnte sich in fachdidaktischen Lehrveranstaltungen niederschlagen. Andererseits offenbart die Befundlage auch Defizite bestehender Testinstrumentarien. Zur Erhebung des CK-Anteils von NOS bei Lehrkräften liegen zahlreiche Instrumente vor, z. B.: Wisconsin Inventory of Science Processes (WISP, Carey und Strauss 1968); Nature of Science Scale (NOSS, Kimball 1967); Views of Nature of Science (VNOS, Lederman et al. 2002; dt. Priemer 2003); Knowledge Structures for NOS and SI (KS4NS, Bartos und Lederman 2014; zusammenfassend vgl. Priemer 2006). Der PCK-Anteil von

NOS wurde bisher qualitativ in Unterrichtsbeobachtungen mit kleinen Stichproben erhoben, z. B. von Lederman und Zeidler (1987), $N=18$; Bell et al. (2000), $N=13$; Bartos und Lederman (2014), $N=4$. Die Forschungslage zeigt, dass es an quantitativen Instrumentarien fehlt, um Professionswissen zu NOS im Bereich des PCKs zu erheben, um die fachdidaktische Erwartungen an Lehrkräfte und ein angemessen im Unterricht kommuniziertes Wissenschaftsverständnis im Sinne von NOS mit empirischen Realitäten bzw. dem konditionalen Handlungswissen von Lehrkräften in Bezug auf NOS abzugleichen. Notwendig erscheinen Testformate, die eine situierte Messung des Wissens und Könnens von Lehrkräften erheben (vgl. König 2015).

Situierte Erhebungen zur Erfassung von Professionswissen

Um lernwirksames Professionswissen aus Fachdidaktiken, Schulpädagogik und pädagogischer Psychologie einer empirischen Prüfung zugänglich zu machen und keine handlungsfernen Artefakte zu erheben, legen aktuelle Forschungsansätze nahe, kontextualisierte und situationspezifische Testaufgaben zu verwenden (z. B. König 2015). Hierfür bieten sich schriftliche Unterrichtsvignetten, d. h. kurze, authentische Beschreibungen von Unterrichtssituationen als Testaufgaben an, wodurch die Erfassung des Professionswissens in einem möglichst realitätsnahen Bezugsrahmen stattfindet (vgl. Shavelson 2013). Solche Testverfahren bieten die Möglichkeit Handlungsoptionen offen oder geschlossen zu erheben. Sofern vorgegebene Handlungsoptionen bewertet werden müssen, wird das Testverfahren von Oser et al. (2010) als „advokatorische Kompetenzmessung“ definiert.

Methode

Zielstellung

Aus dem oben skizzierten Desiderat quantitativ und zugleich situiert prüfender Instrumente zum Konstrukt NOS ergibt sich das folgende Ziel: Es soll ein quantitatives Instrument zur möglichst realitätsnahen und situierten Erhebung des Professionswissens von Lehrkräften zu NOS entwickelt werden. Es interessieren insbesondere die oben genannten und zentralen NOS-Facetten (a) Wissen über die Bedeutung von Kreativität und Subjektivität für wissenschaftliche Erkenntnis (KS), (b) Wissen über die Veränderung wissenschaftlicher Erkenntnisse im Laufe der Zeit (VZ) und (c) Wissen über die Differenzierung zwischen Beobachtung und Schlussfolgerung (BS). Die Modellstruktur soll in ihrer Dimensionalität geprüft und konvergent sowie diskriminant validiert werden. Dazu werden zwei Studien

mit unterschiedlichen Testbatterien und Stichproben durchgeführt.

Testentwicklung mit Unterrichtsvignetten

Mit dem EKoL-NOS-Test des Kollegs *Effektive Kompetenzdiagnose in der Lehrerbildung* (EKoL; Rutsch et al. 2017a) soll Professionswissen zum Konstrukt NOS erhoben werden. Inhaltliche Basis stellen Nature of Science-Elementarisierungen dar. Ausgehend von existierenden Konsenslisten zu NOS – die sich trotz beschriebener Kritik für eine forschungspragmatische Operationalisierung eignen – wurden drei oben beschriebene und zentrale NOS-Facetten ausgewählt, die sich stets in klassischen NOS-Elementarisierungen finden. Das Instrument soll später durch weitere Facetten ergänzt werden. Ferner wird versucht, den Umgang mit dem Konstrukt NOS situationspezifisch angemessen und möglichst nahe an realen Unterrichtssituationen zu erheben. In Anlehnung an Rehm und Bölsterli (2014) wurde versucht, zu den Facetten möglichst authentische Unterrichtsbeispiele in schriftlicher Form (Vignetten) darzustellen, die bedeutsame Unterrichtssituationen (z. B. kritische Probleme oder fruchtbare Momente) beinhalten. Mit der Verwendung von Unterrichtsvignetten soll einer zentralen methodischen Herausforderung an die empirische Erfassung professioneller Kompetenz von Lehrkräften begegnet werden: der Berücksichtigung eines authentischen Bezugsrahmens für die Kompetenzerfassung (Shavelson 2013). Im Gegensatz zu Unterrichtsbeobachtungen lassen sich Vignettentests vergleichsweise zeitökonomisch einsetzen (Rehm und Bölsterli 2014; Rutsch et al. 2017b). Zudem bieten geschlossene Items den Vorteil einer schnellen Durchführung und Auswertung (Tepner und Dollny 2014). Gleichwohl wird mit einem solchen Verfahren die Tiefe einer real durchgeführten Unterrichtsbeobachtung nicht erreicht. Außerdem fordern geschlossene Items von den Testpersonen nur eine (Wieder-)erkennungsleistung, die gegenüber einer (Re-)produktionsleistung durch halboffene Antwortformate als weniger anspruchsvoll gilt (Döring und Bortz 2016).

Herausforderungen an NOS-Vignetten

Der Konstruktcharakter von NOS hat Konsequenzen für die Testentwicklung: NOS gilt als naturwissenschaftsübergreifendes Konstrukt (z. B. Schwarz und Lederman 2008); gleichwohl ist es zunächst nicht selbstverständlich, dass Vignetten zum Modellieren in der Biologie zu denselben Ergebnissen führen, wie Vignetten zum Modellieren in der Physik oder Chemie. So wurde bei jeder Vignette bereits im einführenden Vignettenstamm darauf geachtet, das disziplinspezifische Fachwissen für einen naturwissenschaftsübergreifend einsetzbaren NOS-Test möglichst gering zu halten, bzw. in Einzelfällen auch für Lehrkräfte anderer

Naturwissenschaften verständlich aufzubereiten. Dies trifft ebenfalls auf den Anteil an pädagogischem Wissen (PK) in den Vignetten zu. Beides könnte sonst konstruktirrelevante Varianz erzeugen.

Grundsätzliches Problem des Einsatzes von Vignetten ist ferner die Gratwanderung zwischen möglichst authentischen und somit detailliert zu beschreibenden Unterrichtssequenzen einerseits und einer Fokussierung auf die zu erhebende Konstruktfacetten andererseits (z. B. Friesen et al. 2018; Kuntze 2015).

Vignettenpool

Der Aufbau der Vignetten basiert auf folgendem Muster: Nach einer Überschrift und einer Zeichnung oder einem Photo zum Unterrichtsinhalt wird den Probandinnen und Probanden eine kurze Unterrichtssequenz präsentiert, in der ein Aspekt von NOS Teil des Unterrichtsgeschehens ist. Ein solches Beispiel zur NOS-Facetten *Wissen zur Bedeutung von Subjektivität und Kreativität* zeigt Abb. 1.

Im Anschluss werden den Probandinnen und Probanden drei bis fünf Handlungsoptionen (Items) angeboten, deren situationsspezifische Angemessenheit auf einer Skala von 1 bis 6 zu bewerten ist. Damit erfolgt eine „advokatorische Kompetenzmessung“ (Oser et al. 2010). Ein solches Beispiel von Handlungsalternativen zu der Vignette aus Abb. 1 zeigt Abb. 2.

Nach diesem Schema entstand ein Pool von 14 Vignetten aus dem für die vorliegende Studie neun Vignetten übernommen wurden. Die Beschränkung auf neun Vignetten

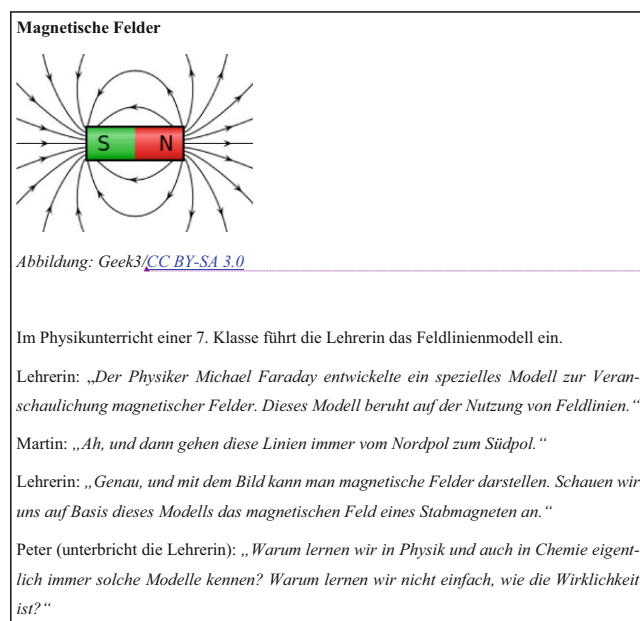


Abb. 1 Beispiel für eine Vignette zum Wissen über die Bedeutung von Subjektivität und Kreativität für wissenschaftliche Erkenntnis

stellt einen Kompromiss zwischen dem zeitlichen Aufwand einerseits, beispielsweise im Einsatz in umfangreicheren Testbatterien und andererseits einer ausreichenden Berücksichtigung der einzelnen NOS-Facetten mit jeweils drei Vignetten dar. Eine Inter-Rater-Übereinstimmung von drei Mitgliedern der Arbeitsgruppe in der Zuordnung der Vignetten zur jeweiligen NOS-Facetten ergab ein Fleiss' $k = 0,77$ (Fleiss 1971; R package *irr*, Gamer et al. 2012) und somit eine Gesamtübereinstimmung von 85,19%. Landis und Koch (1977) bewerten ein k zwischen 0,61 und 0,80 als substantiell. Die Rater haben zuvor nicht an der Vignettenentwicklung teilgenommen. Erstellt wurde der Test als Paper-and-Pencil-Instrument. Für eine Teilstichprobe wurde das Instrument in eine Online-Version (oFb/SoSciSurvey; Leiner 2016) überführt.

Testwertberechnung durch Expertennorm

Innerhalb der Vignetten erfolgt zuerst eine Punktwertvergabe auf Itemebene (vgl. Meschede et al. 2015; Rutsch et al. 2018). Hierfür wurde eine Musterlösung für jedes Item entwickelt, an der sich Naturwissenschaftsdidaktikerinnen und -didaktiker ($N=8$) beteiligten, die zu NOS an deutschsprachigen Hochschulen forschen. Auf Itemebene wird der Modalwert dieser Musterlösung mit den Antworten der Probandinnen und Probanden verglichen. Bei Übereinstimmung wird die volle Itempunktzahl ($= 2$) vergeben, liegt die Antwort eine Stelle auf der Antwortskala neben dem Modalwert, erhält die Probandin oder der Proband die halbe Punktzahl ($= 1$). Bei einzelnen Items ließ sich kein Modalwert als Musterlösung erzielen bzw. fielen die Experten-Bewertungen der Handlungsalternativen sehr unterschiedlich aus. In diesen Fällen wurde das Item nicht in die Testwertberechnung miteinbezogen. Einen Überblick zu Vignetten, ausgeschlossenen Items und Fachbezügen zeigt Tab. 1.

Der Testwert einer Vignette aggregiert sich im Sinne eines polytomen (Super-)Items aus den einzelnen Itemtestwerten der Vignette als Summenwert, der nochmals durch die Anzahl der eingeflossenen Items geteilt wird (vgl. Eckes 2015; Robitzsch und Lüdtke 2015; Rutsch et al. 2018). Die maximal erreichbaren Vignettentestwerte lagen somit pro Vignette bei 2 Punkten.

Nachteilig wirkt sich bei diesem Verfahren aus, dass bei einer sechsstufigen Skala nur jene Probanden Punkte erzielen, die auch den Modalwert treffen (bzw. ein Feld daneben). Individuelles Antwortverhalten, das tendenziell eher positiv oder negativ ausfallen könnte, wird somit nicht im Ergebnis berücksichtigt. Eine Punktevergabe durch Paarvergleiche (Artelt et al. 2009; Tepner und Dollny 2014) legt sich dagegen nicht auf einen bestimmten Referenzwert fest und toleriert Abweichungen eher (vgl. Rutsch et al. 2018). Für die Auswertung dieses Tests haben wir uns dennoch

Abb. 2 Beispiel für Handlungsalternativen zu einer Vignette zum Wissen über die Bedeutung von Subjektivität und Kreativität für wissenschaftliche Erkenntnis

Bewerten Sie die Antworten der Physiklehrerin. Sie sollte sinngemäß wie folgt antworten:						
	1	2	3	4	5	6
	trifft gar nicht zu					trifft voll und ganz zu
<i>Bitte machen Sie in jeder Zeile ein Kreuz.</i>						
„Wissenschaftliche Modelle sind sozusagen Kopien der Wirklichkeit, sie basieren auf wissenschaftlicher Beobachtung. Sie stellen aber nur diejenigen Aspekte dar, die Ihr gerade lernen sollt.“	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
„Wie die Wirklichkeit ist, können wir Menschen nicht immer erkennen. Aber Wissenschaftler können mit Modellen viele ihrer Ideen und Theorien über die Wirklichkeit verdeutlichen. Deshalb lernen wir mit Modellen.“	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
„Modelle erklären immer nur einen Teil der Wirklichkeit und dabei ändern sie sich, weil auch die Forschung Fortschritte macht und diese in neuere Modelle aufgenommen wird. Mit den Modellen versuchen wir der Wirklichkeit nahe zu kommen, aber genau abbilden können wir sie damit nicht.“	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
„Modelle helfen uns beim Lernen und Erklären. Mit Modellen können wir uns beim Lernen auf ganz bestimmte Aspekte konzentrieren, aber Du hast Recht: Modelle sind etwas anderes als die Wirklichkeit.“	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

für eine Punktevergabe auf Itemebene entschieden, da die zu bewertenden Handlungsoptionen in Theorie (und in der Punktevergabe der Experten) häufig keine eindeutige Reihenfolge der Beziehung zueinander erkennen lassen und jede Option für sich selbst beurteilt werden sollte.

Stichproben

An der Erhebung zu Studie 1 nahmen 289 Referendarinnen und Referendare des Lehramts für die Sekundarstufe 1 teil (204 weiblich). Einen detaillierten Überblick liefert Tab. 2.

Tab. 1 Vignetten, zugeordnete Fächer und ausgeschlossene Items

Vignette	Fach	Items (final)	Items ausgeschlossen ^a
Kreativität und Subjektivität 1 (KS1)	Chemie	3	0
Kreativität und Subjektivität 2 (KS2)	Biologie	4	0
Kreativität und Subjektivität 3 (KS3)	Physik	4	0
Veränderung im Laufe der Zeit 1 (VZ1)	Biologie	5	0
Veränderung im Laufe der Zeit 2 (VZ2)	Biologie	2	1
Veränderung im Laufe der Zeit 3 (VZ3)	Physik	4	0
Beobachtung und Schlussfolgerung 1 (BS1)	Sachunterricht	4	0
Beobachtung und Schlussfolgerung 2 (BS2)	Chemie	2	1
Beobachtung und Schlussfolgerung 3 (BS3)	Physik	3	0

^aGrund: Modalwert nicht eindeutig

Tab. 2 Stichprobe 1

Studierte naturwissenschaftliche Fächer	<i>n</i>	<i>W</i>	Alter <i>M</i>	Alter <i>SD</i>	Alter <i>Min</i>	Alter <i>Max</i>
Biologie ^a	146	116	26,58	2,99	23	43
Chemie ^a	39	26	26,44	3,22	23	37
Physik ^a	32	12	27,35	5,57	23	49
Biologie und Chemie	27	21	26,77	3,55	22	40
Biologie und Physik	3	1	25,33	0,58	25	26
Chemie und Physik	10	5	25,90	1,85	23	29
Kein naturwissenschaftliches Fach	24	17	26,71	3,68	24	38

^aAls einziges naturwissenschaftliches Unterrichtsfach

Tab. 3 Stichprobe 2

Studierte naturwissenschaftliche Fächer	<i>n</i>	<i>W</i>	Alter <i>M</i>	Alter <i>SD</i>	Alter <i>Min</i>	Alter <i>Max</i>
Biologie ^a	105	72	27,04	2,45	24	38
Physik ^a	22	5	26,50	2,24	24	35
Biologie und Chemie	4	3	26,25	2,87	23	30
Biologie und Physik	6	2	27,00	1,10	26	29

^aAls einziges naturwissenschaftliches Unterrichtsfach

Die Erhebungen fanden im Rahmen des Forschungs- und Nachwuchskollegs EKoL an Studienseminaren in Baden-Württemberg im Februar 2017 direkt nach Beginn des Vorbereitungsdienstes der Lehrkräfte statt.

An der Studie 2 zur diskriminanten und konvergierenden Validierung nahmen 137 Referendarinnen und Referendare der naturwissenschaftlichen Unterrichtsfächer Biologie und/oder Physik teil (82 weiblich). Einen detaillierten Überblick liefert Tab. 3. Die Erhebungen fanden von November 2017 bis Januar 2018 ebenfalls an Studienseminaren in Baden-Württemberg statt, allerdings gegen Ende des Referendariats kurz vor Beginn des Prüfungszeitraums.

Demographische Variablen

In Bezug auf demographische Variablen (kognitive Leistungsfähigkeit, Geschlecht) kann auf Basis bisheriger Studien zum PCK in den Naturwissenschaften (z. B. Krauss et al. 2017) davon ausgegangen werden, dass sich die kognitive Leistungsfähigkeit, auch in den Ergebnissen des EKoL-NOS-Tests widerspiegelt. In bisherigen naturwissenschafts-didaktischen Tests (Riese 2009; Kirschner 2013; Schödl und Göhring 2017) lässt sich z. Zt. bei männlichen Lehrkräften noch ein Wissensvorsprung beobachten, der Ergebnis von Sozialisierungseffekten sein könnte. Die entsprechenden Prozesse sollten Gegenstand weiterführender Untersuchungen sein. Ferner kann erwartet werden, dass angehende Lehrkräfte naturwissenschaftlicher Fächer in einem naturwissenschaftlichen PCK-Test bessere Leistungen erzielen, als ihre Mitreferendarinnen und -referendare ohne naturwissenschaftliches Fach; einschränkend ist jedoch zu erwarten,

dass auch angehende Lehrkräfte nicht-naturwissenschaftlicher Fächer über wissenschaftstheoretische Kenntnisse aus ihren jeweiligen Fachdisziplinen verfügen, insbesondere die drei ausgewählten Facetten gelten grundsätzlich auch für nicht-naturwissenschaftliche Wissenschaftsdisziplinen (vgl. van Dijk 2011; Gebhard et al. 2017); beispielsweise versuchen die ästhetischen Wissenschaften bei der Analyse von Gemälden, Grafiken, Skulpturen, Fotografien oder Installationen zwischen Beobachtung und Interpretation zu differenzieren (vgl. Huber 2011). Auch in der Sportwissenschaft basiert ein Untersuchungsdesign auf dem subjektiven Vorverständnis der Forschenden; Interpretation, Modellierung und Auswertung der Daten sowie die Theoriebildung erfordern auch hier schlussfolgerndes Denken und Kreativität. Studierende der Religionswissenschaft bzw. Theologie lernen im ersten Semester die *historisch-kritische Methode* kennen und damit die *tentative nature*, dass „historische Erkenntnisse nicht endgültig, sondern prinzipiell revidierbar sind.“ (Heilgenthal et al. 1999, S. 30).

Konvergierende und diskriminante Konstrukte zu NOS: Epistemische Überzeugungen und berufsbezogene Motivation

Zur diskriminanten und konvergierenden Validierung des Instruments werden zwei Konstrukte herangezogen, die an dieser Stelle kurz skizziert werden. Im Modell zur professionellen Kompetenz von Lehrkräften des Forschungsprogramms COACTIV zeigen sich berufsbezogene Motivation der Lehrkräfte und ihre selbstregulativen Fähigkeiten nahezu unkorreliert mit dem Professionswissen (Kleickmann

Tab. 4 Aufgabenschwierigkeiten und Trennschärfe

Vignette	<i>M</i>	<i>SD</i>	Min	Max	<i>P_i</i>	<i>R_{it}</i>	Missings (%)
Kreativität und Subjektivität 1 (KS1)	0,51	0,50	0	2	0,26	0,54	3,80
Kreativität und Subjektivität 2 (KS2)	0,42	0,43	0	2	0,21	0,57	3,46
Kreativität und Subjektivität 3 (KS3)	0,87	0,45	0	2	0,44	0,62	3,11
Veränderung im Laufe der Zeit 1 (VZ1)	0,50	0,46	0	2	0,25	0,46	3,80
Veränderung im Laufe der Zeit 2 (VZ2)	0,97	0,57	0	2	0,49	0,63	4,15
Veränderung im Laufe der Zeit 3 (VZ3)	0,98	0,37	0	2	0,49	0,40	4,84
Beob. und Schlussfolgerung 1 (BS1)	1,07	0,42	0	2	0,54	0,39	4,84
Beob. und Schlussfolgerung 2 (BS2)	0,96	0,54	0	2	0,48	0,29	4,50
Beob. und Schlussfolgerung 3 (BS3)	0,91	0,46	0	2	0,46	0,40	3,11

M Mittelwert, *SD* Standardabweichung, *min* tatsächlich erreichter Minimaltestwert, *max* tatsächlich erreichter Maximaltestwert, *P_i* Aufgabenschwierigkeit, *r_{it}* Korrigierte Item-Skala-Korrelation

und Anders 2011), dies trifft insbesondere auf die Facette „Einfachheit des Studiums“ zu. Zur diskriminanten Validierung wurden somit aus dem Fragebogen FEMOLA (Pohlmann und Möller 2010) Items zur entsprechenden Facette der berufsbezogenen Motivation übernommen.

Als kovergierendes Konstrukt wurden epistemische Überzeugungen ausgewählt. Die Naturwissenschaftsdidaktik diskutiert Vorstellungen über wissenschaftliches Wissen im Konstrukt NOS. Die psychologisch ausgerichtete Lehr-Lern-Forschung untersucht Vorstellungen von Wissen unter dem Label *epistemische Überzeugungen* (z.B. Schommer 1990; Hofer und Pintrich 1997; Bromme et al. 2016). Zwischen den Konstrukten bestehen somit inhaltliche Zusammenhänge. Gleichwohl lassen sich beide Konstrukte theoretisch voneinander abgrenzen (vgl. Neumann und Kremer 2013) und unterschiedlich in Kompetenzmodellen verorten. Die epistemischen Überzeugungen der Lehrkraft lassen sich nach Baumert und Kunter (2006) den Überzeugungen und Werthaltungen zuordnen, da sie individuelle Überzeugungen repräsentieren. Neumann und Kremer sprechen in diesem Zusammenhang von einer Perspektive der 1. Person. Im Gegensatz dazu definieren Neumann und Kremer für NOS eine Perspektive der 3. Person, einer naturwissenschaftlichen Erkenntnistheorie, da traditionellen NOS-Elementarisierungen ein wissenschaftstheoretischer Minimalkonsens mit normativem Charakter zu Grunde liegt, der zu lehrende Eigenschaften wissenschaftlichen Wissens beschreibt und die in Bildungsstandards verankert sind (z.B. KMK 2005a–c; NGSS 2013). Für eine konvergierende Validierung scheint jedoch trotz dieser Differenzen eine ausreichende Nähe gegeben. Der Fragebogen CAEB (= *Connotative Aspects of Epistemological Beliefs*; Stahl und Bromme 2007) versucht epistemische Überzeugungen über ein semantisches Differenzial (vgl. Osgood et al. 1957) zu erheben. Der CAEB fragt nach Assoziationen zu Konzepten oder Fächern, beispielsweise zu „Wissen in der Biologie“ oder „Wissen in der Physik“. Diese Assoziationen werden über Adjektivpaare wie *objektiv-subjektiv* oder *offen-abgeschlossen* erhoben. Dazu wird in Fragebögen

auf einer siebenstufigen Skala in Richtung des einen oder anderen Adjektivs gekreuzt. Faktorenanalytisch lassen sich zwei Dimensionen zur Struktur/Textur sowie Variabilität von Wissen abbilden, die in zahlreichen Studien repliziert werden konnten (z.B. von Porsch und Bromme 2011; Mason et al. 2014; Kramer 2017). Die Dimensionen erinnern an zwei der klassischen Dimensionen epistemischer Überzeugungen: „Simplicity“ und „Certainty“ (vgl. Schommer 1990; Hofer und Pintrich 1997).

Ergebnisse

Im Folgenden werden Ergebnisse zur Reliabilität und empirischen Validierung des Instruments vorgestellt. Die Auswertung erfolgte mit R (Version 3.5.0, R Core Team 2018) sowie dem Paket lavaan (Version 0.6.1 2018, Rosseel 2012). Dargestellt werden die Befunde mit der ersten Stichprobe ($N=289$) anhand der Aspekte (1) Aufgabenschwierigkeiten und Trennschärfe, (2) Dimensionalitätsprüfung und (3) Validitätsprüfung in Hinblick auf demographische Variablen. Die Validitätsprüfung zu diskriminanten und kovergierenden Konstrukten wurde mit einer zweiten Stichprobe ($N=137$) durchgeführt. Außerdem wurde mit dieser zweiten Stichprobe geprüft, inwiefern sich die Testergebnisse angehender Lehrkräfte der Fächer Biologie und Physik voneinander unterscheiden.

Aufgabenschwierigkeiten und Trennschärfe

Tab. 4 gibt einen Überblick zu den Aufgabenschwierigkeiten und Trennschärfe der einzelnen Vignetten. Die mittlere Aufgabenlösungshäufigkeit liegt zwischen 0,21 und 0,54. Damit kann mit den Testvignetten gut zwischen den Probandinnen und Probanden differenziert werden, wobei drei der Vignetten (KS1, KS2 und VZ1) mit Lösungshäufigkeiten von $<0,3$ als schwierig einzustufen sind. Grundsätzlich fallen den Referendarinnen und Referendaren Aufgaben zur NOS-Facette *Beobachtung und Schlussfolgerung* am leicht-

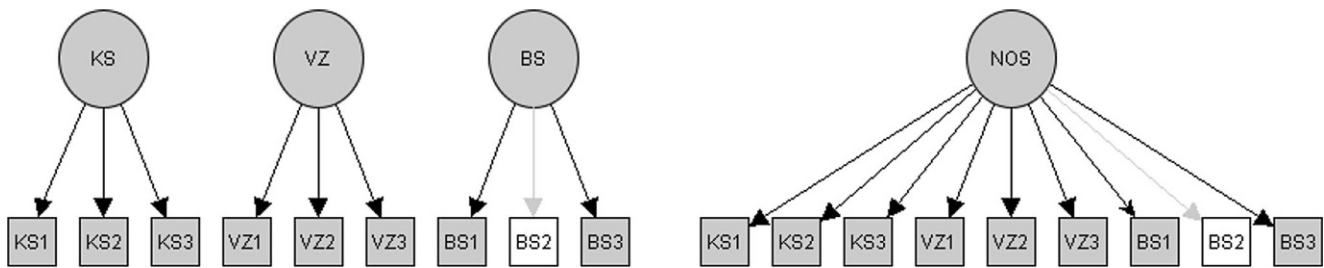


Abb. 3 Vergleich des dreidimensionalen Modells mit den Wissensfacetten zu (1) *Kreativität und Subjektivität* [KS], (2) *Veränderung wissenschaftlichen Wissens im Laufe der Zeit* [VZ] und (3) *Unterscheidung zwischen Beobachtung und Schlussfolgerung* [BS] und einem eindimensionalen Modell. Die schwachgedruckte Vignette BS2 wurde nach ersten Modellrechnungen entfernt

testen (P_i zwischen 0,46 und 0,54). Die größten Schwierigkeiten bereiten Vignetten zur NOS-Facette *Kreativität und Subjektivität* (P_i zwischen 0,21 und 0,44). Die Trennschärfen bewegen sich im mittelmäßigen (0,30 bis 0,50) bis hohen Bereich (größer als 0,50; nach Döring und Bortz 2016). Nur Vignette BS2 liegt unter 0,3; auch in der folgenden Modellprüfung ist Vignette BS2 zu hinterfragen.

Dimensionalitätsprüfung

Prüfung der Ein- bzw. Mehrdimensionalität

NOS stellt auch in den Konzeptualisierungen von Lederman (2007) und McComas (1998) bestenfalls einen Minimalkonsens und kein kontingentes Konstrukt dar. Die Dimensionalität wurde mit einem ein- und einem dreidimensionalen Strukturgleichungsmodell geprüft (siehe Abb. 3).

Grundlage stellen die neun Vignettentestwerte dar, wobei das dreidimensionale Modell zwischen den NOS-Facetten unterscheidet und das eindimensionale Modell NOS als gemeinsamen Faktor abbildet. Zunächst wies eine der Vignetten zur Facette *Wissen über die Differenzierung zwischen Beobachtung und Schlussfolgerung* in beiden Modellen inakzeptable Faktorladungen auf ($\lambda_j = 0,037$; $\lambda_j = 0,019$). Nach Entfernung der Vignette wurde mit Modellen weitergearbeitet, die auf acht Vignetten basieren. Tab. 5 zeigt, dass sowohl ein dreidimensionales Modell als auch ein eindimensionales Modell eine sehr gute Modellanpassungsgüte aufweist.

Ein Vergleich zwischen beiden Modellen zeigt keinen signifikanten Unterschied: $\Delta\chi^2(3) = 1,67$, $p = 0,642$. Damit passt ein sparsames Modell mit nur einer Dimension ebenso zu den Daten, wie ein dreidimensionales. In der Folge

wird mit dem sparsameren eindimensionalen Modell weitergearbeitet; Cronbachs α liegt für dieses Modell bei 0,631.

Validierung über demographische Daten

Mit der ersten Stichprobe wird zunächst eine Validitätsprüfung mit Variablen zu demographischen Daten in Bezug auf die kognitive Leistungsfähigkeit (approximiert über die Note des Abiturs; Baron-Boldt et al. 1988; Schödl und Göring 2017), Geschlecht und das Studium eines naturwissenschaftlichen Faches berichtet. Da die Leistung über Textvignetten erhoben wurde, wird ferner das Personenmerkmal „Deutsch als Muttersprache“ kontrolliert. Ein Strukturgleichungsmodell mit diesen Variablen (Abb. 4) weist eine sehr gute Modellpassung auf ($\chi^2 = 41,42$, $df = 60$, $n = 239$, CFI = 1,00, RMSEA = 0,000, SRMR = 0,039; Cutoff-Werte siehe Tab. 5). Die Abiturnote wirkt prädiktiv auf die Testergebnisse des Instruments ($\beta = -0,17$, $p = 0,025$). Ferner erzielen männliche Lehrkräfte etwas bessere Testergebnisse ($\beta = -0,16$, $p = 0,043$), was Ergebnisse weiterer aktueller Untersuchungen zu fachdidaktischem Wissen naturwissenschaftlicher Fächer repliziert (z. B. Schödl und Göring 2017; Kirschner 2013; Riese 2009). Der Effekt, ob ein naturwissenschaftliches Unterrichtsfach studiert wurde, wird bei der untersuchten Stichprobengröße nicht signifikant ($\beta = 0,15$, $p = 0,082$). Der Effekt von „Deutsch als Muttersprache“ wird ebenfalls nicht signifikant ($\beta = -0,07$, $p = 0,307$).

Validierung über diskriminante und konvergierende Konstrukte

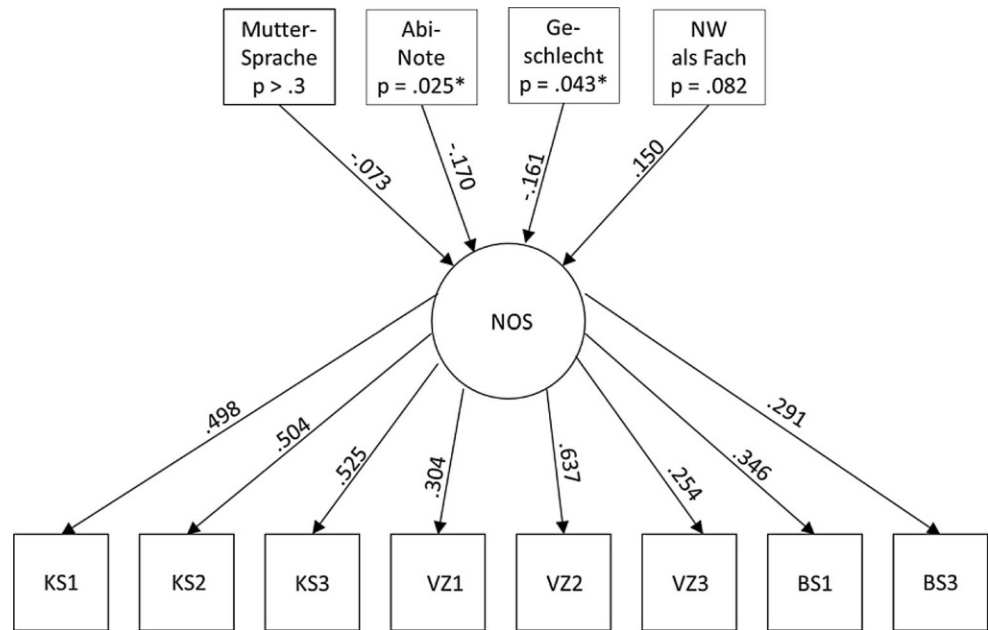
Mit einer zweiten Stichprobe aus Referendaren der Lehrämter Biologie und Physik wird versucht, das Instrument an-

Tab. 5 Modellvergleich

	χ^2	Df	n	p	CFI	RMSEA	SRMR	λ_j
Dreidimensionales Modell	25,34	17	282	0,095	0,955	0,041	0,033	0,222–0,592
Eindimensionales Modell	26,69	20	282	0,144	0,962	0,034	0,037	0,233–0,636

Cutoff-Werte nach Hu und Bentler (1999): $\chi^2 \leq 3$; df , $p \geq 0,01$; CFI (Comparative Fit Index) $\geq 0,95$; RMSEA (Root-Mean-Square Error of Approximation) $\leq 0,06$; SRMR (Standardized Root Mean Residual) $\leq 0,08$

Abb. 4 Strukturgleichungsmodell zur Validierung mit den demographischen Variablen *Muttersprache* (deutsch= 1, andere= 2), *kognitive Leistungsfähigkeit* (Note Abitur), *Geschlecht* (m= 1, w= 2) und ob ein *naturwissenschaftliches Fach* studiert wurde (nein= 0, ja= 1). An den Pfeilen werden die standardisierten Regressionskoeffizienten dargestellt



hand konvergierender und diskriminanter Konstrukte zu validieren. Tab. 6 stellt zunächst die deskriptiven Testergebnisse aus Stichprobe 2 dar.

Zur diskriminanten Validierung wird aus dem Fragebogen FEMOLA (zur Motivation der Wahl des Lehramtsstudiums; Pohlmann und Möller 2010) der Aspekt „Geringe Schwierigkeit des Studiums“ aufgenommen. Zur konvergierenden Validierung wird eine Kurzversion des CAEB (Stahl und Bromme 2007; Kurzversion aus: Kramer 2017) mit 13 Items eingesetzt. Die zweifaktorielle Struktur des CAEB mit den Faktoren „Struktur“ sowie „Variabilität“ von Wissen lässt sich replizieren. Drei der 13 CAEB-Items wurden aufgrund mangelnder Faktorladungen (<0,3) entfernt.

Weitere Konstrukte, Aspekte und Variablen wurden in diesem Modell nicht spezifiziert, um bei der Modellierung latenter Konstrukte und einer Stichprobe von deutlich unter 200 noch sinnvoll mit Strukturgleichungsmodellen arbeiten zu können (vgl. dazu Bentler und Chou 1987) (Abb. 5).

Das spezifizierte Modell weist eine gute Anpassungsgüte auf ($\chi^2 = 238,52$, $df = 203$, $n = 137$, $CFI = 0,960$, $RMSEA = 0,035$, $SRMR = 0,071$; Cutoff-Werte siehe Tab. 5). Erwartungskonform zeigt sich kein Zusammenhang zum Berufswahlmotiv „Geringe Schwierigkeit des Studiums“ ($r = -0,08$, $p = 0,180$). Der Effekt ($r = 0,17$) für die Korrelation mit dem CAEB-Faktor Variabilität wird signifikant ($p = 0,036$), die Korrelation mit dem CAEB-Faktor Struktur ($p = 0,467$) dagegen nicht. Untereinander korrelieren die zwei CAEB-Faktoren signifikant ($p = 0,010$).

Gruppenunterschiede zwischen den studierten Fächern Biologie und Physik

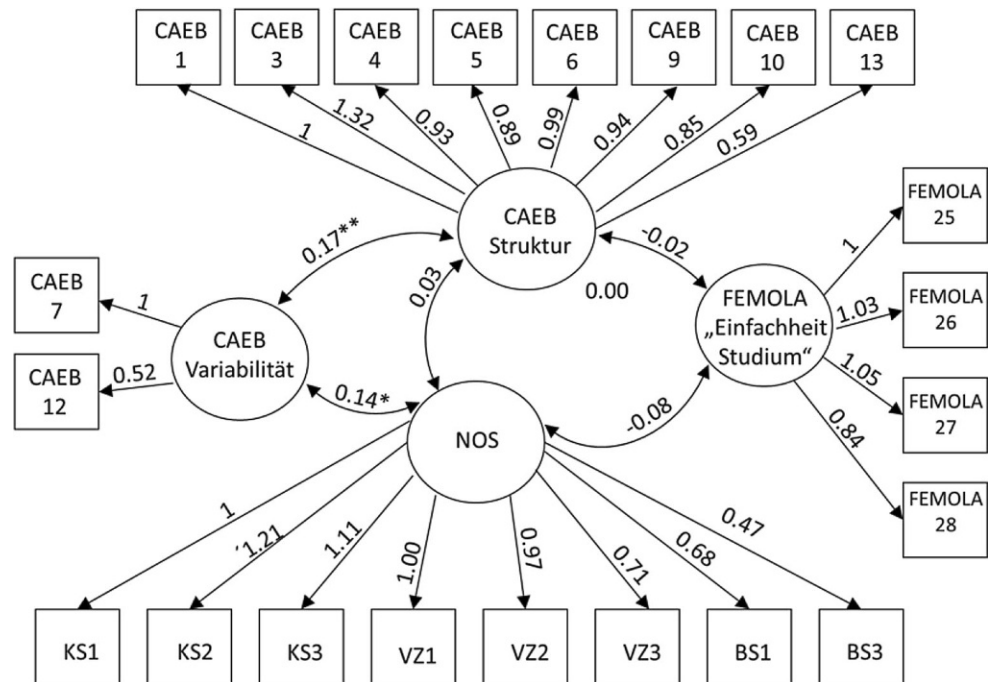
Für den NOS-Test wird zwar ein Einfluss des fachwissenschaftlichen Vorwissens angenommen, andererseits wurde versucht, durch die Vignetten- und Itemgestaltung Professionswissen zum Konstrukt NOS naturwissenschaftsübergreifend erheben zu können und das fachwissenschaftli-

Tab. 6 Deskriptive Ergebnisse Stichprobe 2

Vignette	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>P_i</i>	<i>R_{it}</i>
Kreativität und Subjektivität 1 (KS1)	0,80	0,43	0,00	2,00	0,40	0,48
Kreativität und Subjektivität 2 (KS2)	0,29	0,33	0,00	1,75	0,15	0,51
Kreativität und Subjektivität 3 (KS3)	0,73	0,39	0,00	1,50	0,37	0,52
Veränderung im Laufe der Zeit 1 (VZ1)	0,47	0,42	0,00	2,00	0,24	0,47
Veränderung im Laufe der Zeit 2 (VZ2)	0,64	0,46	0,00	1,50	0,32	0,51
Veränderung im Laufe der Zeit 3 (VZ3)	0,89	0,42	0,00	2,00	0,45	0,45
Beob. und Schlussfolgerung 1 (BS1)	1,03	0,45	0,25	2,00	0,52	0,37
Beob. und Schlussfolgerung 3 (BS3)	0,71	0,39	0,00	1,67	0,36	0,41

M Mittelwert, *SD* Standardabweichung, *min* tatsächlich erreichter Minimaltestwert, *max* tatsächlich erreichter Maximaltestwert, *P_i* Aufgabenschwierigkeit, *r_{it}* Korrigierte Item-Skala-Korrelation

Abb. 5 Strukturgleichungsmodell zur Validierung über das als diskriminant angenommene Konstrukt *Berufswahl „Geringe Schwierigkeit Studium“* und das als konvergierend angenommenen Konstrukt *Epistemische Überzeugungen*. An den einseitigen Pfeilen werden die Regressionskoeffizienten dargestellt, an den beidseitigen Pfeilen die Korrelationskoeffizienten. * = Signifikanzniveau $p=0,05$; ** = Signifikanzniveau $p=0,01$. CAEB = Connotative aspects of epistemological beliefs. FEMOLA = Fragebogen zur Erfassung der Motivation des Lehramtsstudiums.



che Vorwissens möglichst irrelevant zu halten. Um zu prüfen, inwiefern besondere fachwissenschaftliche Aspekte der Einzelvignetten die Ergebnisse beeinflussen könnten, wird mit Hilfe von *t*-Tests geprüft, inwiefern sich die Testwerte angehender Lehrkräfte der Fächer Biologie und Physik voneinander unterscheiden. Tab. 7 bildet die Ergebnisse ab. Keiner der *p*-Werte zeigt sich signifikant; der *p*-Wert für Vignette KS1 liegt allerdings bei 0,0501.

Diskussion und Ausblick

Das in diesem Artikel vorgestellte Instrument EKoL-NOS ist ein Ansatz, Professionswissen zum Konstrukt *Nature of Science* handlungsnah mit Hilfe von Vignetten zu erfassen, um Informationen über interindividuelle Kompetenzunter-

schiede von Lehrkräften zu erhalten, bzw. fachdidaktische Erwartungen mit empirischen Realitäten abzugleichen. Basis für die Testentwicklung stellen Facetten aus klassischen *Nature of Science*-Konsenslisten mit einem Schwerpunkt auf dem kognitiv-epistemischen Charakter wissenschaftlichen Wissens dar. Die Testwertberechnung erfolgt auf Basis einer Expertennorm. Da weder McComas (1998) noch Lederman (2007) ein kontingentes NOS-Bild entworfen haben, sondern (bestenfalls) einen Minimalkonsens referieren, stellt sich eine Konstruktvalidierung per se als herausfordernd dar.

Die empirische Validierung des Instruments mit einem diskriminanten Konstrukt sowie demographischen Variablen (kognitive Leistungsfähigkeit, Geschlecht) liefert zumindest erste Hinweise, Professionswissen zu NOS über quantitativ prüfende Instrumente handlungsnah abbilden zu

Tab. 7 Gruppenunterschiede in den Testwerten zwischen den studierten Fächern Biologie und Physik

Vignette	Unterrichtsbeispiel	Biologie <i>n</i> = 115		Physik <i>n</i> = 28		<i>d</i>	<i>p</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
KS1	Chemie	0,78	0,43	0,94	0,37	0,39	0,050
KS2	Biologie	0,27	0,31	0,32	0,37	0,14	0,530
KS3	Physik	0,72	0,39	0,75	0,41	0,07	0,761
VZ1	Biologie	0,47	0,42	0,46	0,40	0,01	0,971
VZ2	Biologie	0,62	0,46	0,70	0,46	0,16	0,443
VZ3	Physik	0,88	0,43	0,94	0,40	0,13	0,541
BS1	Sachunterricht	1,02	0,46	1,12	0,44	0,22	0,300
BS3	Physik	0,71	0,39	0,69	0,39	0,05	0,813

M Mittelwert, *SD* Standardabweichung, *Maximaltestwert* 2, *d* Effektstärke

können. Erwartungskonform zeigt sich kein Effekt zum diskriminanten Konstrukt der Berufswahlmotivation „Einfachheit des Lehramtsstudiums“. Die als konvergierendes Konstrukt ausgewählten epistemischen Überzeugungen, abgebildet durch das semantische Differenzial CAEB, korrelieren mit dem NOS-Vignettest zumindest in Bezug auf den Faktor Variabilität signifikant. Das ist für diesen Faktor naheliegend, als dass drei der Vignetten im NOS-Test versuchen, die Facette „Veränderbarkeit im Laufe der Zeit“ abzubilden und somit eine inhaltliche Nähe zum CAEB-Faktor aufweisen. Trotz der Signifikanz überrascht die geringe Korrelation von nur $r=0,14$; allerdings korrelieren die beiden CAEB-Faktoren auch nur auf einem Niveau von $r=0,17$ miteinander. Um von Konvergenz zu sprechen unterscheiden sich die Instrumente in den intendierten Ansprüchen und abzubildenden Konstrukten offenbar doch zu stark: Einerseits versucht der NOS-Test Professionswissen mit zum Unterricht kommunizierten Wissenschaftsverständnis zu bestimmten Facetten handlungsnah abzubilden und zu erheben; dazu gehören auch der Umgang mit Schülervorstellungen und angemessenen Aufgaben. Das semantische Differenzial CAEB ist nicht dazu konzipiert, ein differenziertes Wissenschaftsverständnis abzubilden, sondern versucht vielmehr spontane Assoziationen (CA = „connotative aspects“) zu wissenschaftlichen (Teil-)Gebieten zu erheben. PCK-Anteile wie im Vignettest sind nicht intendiert.

Eine zentrale Limitierung der vorliegenden Untersuchung liegt in der grundlegenden theoretischen Konzeption dieser ersten Version des Instruments: Die Vignetten wurden im Sinne von klassischen NOS-Elementarisierungen auf Basis der sogenannten „consensus view“ entwickelt und rekurren auf die Bedeutung subjektiver Anteile wissenschaftlichen Wissens. Aktuelle Ansätze (Allchin 2017, 2018; Erduran und Dagher 2014) fokussieren NOS stärker im Sinne von Ganzheitlichkeit und gesellschaftlicher Einbindung wissenschaftlicher Erkenntnis. Allchin nennt eine seiner Neukonzeptualisierungen „Nature of (whole) science“ (z. B. Allchin 2011, 2017), um mit der Bezeichnung *whole science* zu verdeutlichen, dass im Konstrukt NOS eben nicht voneinander isolierte „features of science“ (= FOS, Matthews 2012) elementarisiert werden sollten, sondern im Sinne einer *scientific literacy* der wissenschaftliche Erkenntnisprozess in seiner Ganzheit. Folgt man der Argumentation von Matthews und Allchin liegt mit dem vorliegenden und an klassischen NOS-Elementarisierungen orientierten Instrument ein FOS-Test und kein NOS-Test vor: Es wird versucht, Professionswissen zu ausgewählten Features (wie „subjektiver Anteile“) wissenschaftlicher Erkenntnis zu erheben und nicht die Ganzheitlichkeit des wissenschaftlichen Erkenntnisprozesses. Dass in der Dimensionalitätsprüfung ein dreidimensionales Modell mit Teilfacetten als eigenen Faktoren einem eindimensionalen Modell nicht überlegen ist, kann mit der Konzentration

auf Facetten, die im Kern auf die Subjektivität bzw. das Wechselspiel zwischen Subjektivität und Objektivität naturwissenschaftlichen Wissens rekurren erklärt werden. Die nicht signifikante Modellprüfung bestätigt somit, dass zumindest für die drei Facetten keine unterschiedlichen Dimensionen vorliegen. Aussagen für die Dimensionalität des Gesamtkonstrukts NOS können aber nicht getroffen werden. Auch die Restunsicherheit, die sich in einem Fleiss' k von „nur“ 0,77 bei neun „Superitems“ aus komplexen Vignettenstämmen und jeweils eigenen Items zeigt, kann als Hinweis interpretiert werden, dass die drei gewählten Facetten im Kern auf dieselbe Dimension rekurren, in deren Zentrum die Subjektivität bzw. das Wechselspiel zwischen Subjektivität und Objektivität naturwissenschaftlichen Wissens steht. Ein weitere Facetten prüfendes Instrument könnte an dieser Stelle zu differenzierteren Ergebnissen kommen. Erduran und Dagher (2014) elementarisieren NOS in zwei Hauptdimensionen: Naturwissenschaft als kognitiv-epistemisches System und Naturwissenschaft als sozial-institutionelles System, die sie aber ebenfalls explizit als nicht trennbar, sondern als zueinander in Beziehung stehend definieren. Die für diese erste Version des Instruments ausgewählten Facetten wissenschaftlichen Wissens konzentrieren sich primär auf Naturwissenschaft als kognitiv-epistemisches System. Für die geplante Weiterentwicklung des Instruments scheint somit eine Konzentration auf Facetten zu Naturwissenschaft als sozial-institutionellem System fruchtbar. Mit solchen Vignetten könnte geprüft werden, ob sich die von Erduran und Dagher vorgeschlagene Dimensionalisierung über einen handlungsnahen Vignettest abbilden lässt. Dies bietet sich ebenfalls für eine weitere Neukonzeptualisierung von NOS an: Auch Allchin (2018) hat inzwischen eine weitere NOS-Konzeptualisierung vorgelegt, in der er eine Weiterentwicklung von NOS zu „Nature of Science in Society“ (= NOSIS) forciert. Sowohl eine ganzheitliche Betrachtung des Wissenschaftsverständnisses, als auch die gesellschaftliche Verankerung wissenschaftlicher Erkenntnis zu erheben, war noch nicht Anspruch dieses NOS-Tests. Um auch diesen Aspekten eines angemessenen Naturwissenschaftsverständnisses gerecht zu werden, können künftige Vorhaben weitere Vignetten auf einer solchen theoretischen Grundlage ins Zentrum ihrer Forschungen stellen.

Ungeklärt bleibt, warum sich das Studium eines naturwissenschaftlichen Unterrichtsfaches gegen über der Vergleichsgruppe ohne naturwissenschaftliches Unterrichtsfach nicht in signifikant besseren Testergebnissen widerspiegelt: ob ein naturwissenschaftliches Fach studiert wurde, wird in der Modellierung mit einem p -Wert von 0,082 nicht signifikant. Interpretiert wird das Ergebnis in drei Richtungen:

Erstens: Sehr vorsichtig wird interpretiert, dass das Professionswissen zu NOS bei angehenden Lehrkräften naturwissenschaftlicher Fächer entsprechend der Testergebnisse tatsächlich nicht stärker ausgeprägt sein könnte, als bei angehenden Lehrkräften ohne naturwissenschaftliches Fach. Internationale Untersuchungen (z. B. Liang et al. 2006) liefern Hinweise, dass sich die wissenschaftstheoretische Konzeption von NOS häufig nicht im Wissenschaftsverständnis von angehenden Lehrkräften widerspiegelt; dies könnte auch im deutschsprachigen Raum der Fall sein. Die im Mittel recht geringen Testergebnisse (siehe Tab. 4 und 6) könnten (z. Zt. allerdings noch sehr vorsichtig) als Replikation internationaler Studien gedeutet werden und lassen erste Vermutungen auf eine nach wie vor geringe Bedeutung von NOS in der aktuellen Bildung naturwissenschaftlicher Lehrkräfte zu. Es müsste geprüft werden, ob der Stellenwert von NOS in der Lehrerbildung an den Hochschulen womöglich hinter anderen naturwissenschaftlichen und naturwissenschaftsdidaktischen Inhalten zurückbleibt. Die Rahmenverordnung für die Lehramtsausbildung in Baden-Württemberg, nach der die Hochschulen ihre Modulhandbücher gestalten und nach der das Gros der angehenden Lehrkräfte der Stichproben ausgebildet wurde, erwartet nur für den sogenannten naturwissenschaftlich-technischen Sachunterricht der Grundschule sowie das Fach Technik in der Sekundarstufe explizit Studieninhalte zu *Nature of Science*. Für die naturwissenschaftlichen Fächer der Sekundarstufe finden sich erkenntnistheoretische Inhalte dagegen eher implizit im Rahmen naturwissenschaftlicher Denk- und Arbeitsweisen.

Zweitens: Mit dem vorliegenden Instrument wurde versucht, Professionswissen im Bereich des CK und PCK zum Konstrukt NOS zu erfassen, also zu einem angemessenen Wissenschaftsverständnis. Sollte das Professionswissen zu CK und PCK im Bereich von NOS tatsächlich nur wenig ausgeprägt sein, da NOS im Lehramtsstudium bisher nur wenig reflektiert wurde, kann im Sinne des TIDE-Modells (Muis et al. 2006) interpretiert werden, dass die angehenden Lehrkräfte mangels differenzierter NOS-Kenntnisse (bzw. mangels angemessener fachspezifischer epistemischer Überzeugungen) ersatzweise mit generelleren epistemischen Überzeugungen antworten.

Drittens stellt sich die Frage, ob bisherige NOS-Elementarisierungen als Demarkationslinie geeignet sind, um ein angemessenes Naturwissenschaftsverständnis von einem angemessenen Verständnis anderer Wissenschaftsdisziplinen abzugrenzen. Nach van Dijk (2011) und Gebhard et al. (2017) können bisherige Konsens-Elementarisierungen der NOS-Facetten auch auf nicht-naturwissenschaftliche Wissenschaften angewandt werden. Angehende Lehrkräfte nicht-naturwissenschaftlicher Fächer haben Grundzüge pädagogischer Psychologie, der Erziehungswissenschaft sowie fachwissenschaftliche Anteile ihrer Unterrichtsfächer

studiert und sollten mit den Grundlagen wissenschaftlicher Erkenntnis in ihren jeweiligen Disziplinen vertraut sein.

Eine weitere grundsätzliche Limitation erfährt das Instrument durch die Verwendung geschlossener Items und einem advokatorischen Ansatz. Zunächst wird dadurch eine zeitökonomische Durchführung und Auswertung ermöglicht. Durch ein solches Verfahren kann jedoch nicht die Tiefe einer real durchgeführten Unterrichtsbeobachtung erreicht werden. Durch geschlossene Items wird von den Testpersonen ferner nur eine (Wieder-)Erkennungsleistung, jedoch keine (Re-)Produktionsleistung gefordert (vgl. Döring und Bortz 2016), sodass anhand der Testergebnisse nur vorsichtig auf die Performanz im tatsächlichen Unterrichtsgeschehen geschlossen werden kann. Die interne Konsistenz des eindimensionalen Modells mit einem Cronbach's α von 0,631 ist ebenfalls zu hinterfragen; für die erste Version eines neuentwickelten fachdidaktischen Instruments interpretieren wir den Wert als noch akzeptabel. Ferner ist in Bezug auf die Schwierigkeitsverteilung tendenziell ein Bodeneffekt zu erkennen. Offenbar gibt es in Hinblick auf die Zielgruppe keine leichten Items. Inzwischen wurde das Instrument in NOS-Interventionen eingesetzt; die Analyse der Ergebnisse wird Hinweise geben, ob die sich die Items auch nach Interventionen oder mit zunehmenden NOS-Kenntnissen als schwierig darstellen; daneben sollte bei der Überarbeitung des Instruments der Schwierigkeitsgrad der Items besondere Berücksichtigung finden.

Grundsätzlich lässt sich der Anspruch des Tests, NOS als ein die naturwissenschaftlichen Domänen überspannendes Konstrukt zu formulieren und dementsprechend die NOS-Vignetten möglichst frei von inhaltsbezogenem fachspezifischen *content knowledge* zu formulieren, durch die vorliegende Studie noch nicht belastbar abstützen. Die signifikante, jedoch geringe Korrelation mit dem zur konvergierenden Validitätsprüfung bemühten Instrument CAEB liefert dazu noch keine hinreichenden Belege. Um Professionswissen zum naturwissenschaftlichen Unterricht zu NOS handlungsnah und quantitativ mit Hilfe von Unterrichtsvignetten erheben zu können, bedarf es somit weiterer Prüfung.

Ein Nutzen, Professionswissen zu NOS gezielt und in Bezug auf bestimmte Aspekte zu erfassen, scheint dennoch gegeben, da die momentane Forschungslage trotz teils intensiver bildungspolitischer und fachdidaktischer Anstrengungen ein unzureichendes Professionswissen zu NOS bei zukünftigen Lehrkräften belegt. „Klassische“ NOS-Elementarisierungen haben diese Aspekte (momentan noch) sehr viel differenzierter herausgearbeitet als der Whole-Science- oder FRA-Ansatz und eignen sich deshalb insbesondere für fachdidaktische Forschungszwecke und individuelle Rückmeldungen (z. B. in der Lehrerbildung), um etwa Lernbedarfe festzustellen oder spezifischere Informationen zu Zusammenhängen mit anderen Konstrukten zu untersuchen. Im neuentwickelten und handlungsnah erhe-

benden Instrument liegen somit zu nutzende Chancen. So ist neben zusätzlichen Schritten der Validierung die Weiterentwicklung mit Vignetten zu weiteren NOS-Aspekten geplant. Aber auch über das hier vorgestellte Instrument hinaus scheint weiterer Bedarf für entsprechende Forschungsarbeiten und fachdidaktische Lehrveranstaltungen (beispielsweise mit Vignetten) nötig.

Danksagung Wir bedanken uns herzlich bei den anonymen Gutachterinnen bzw. Gutachtern für die ausführlichen Rückmeldungen und die wertvollen kritischen und konstruktiven Anregungen.

Förderung Das Forschungs- und Nachwuchskolleg EKoL (Effektive Kompetenzdiagnose in der Lehrerbildung) wurde finanziert vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg.

Funding Open Access funding provided by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- AAAS (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95(3), 518–542.
- Allchin, D. (2017). Beyond the consensus view: whole science. *Canadian Journal of Science, Mathematics and Technology Education*, 17, 18–26.
- Allchin, D., & Werth, A. (2017). The naturalizing error. *Journal for the General Philosophy of Science*, 48, 3–18.
- Allchin, D. (2018). From nature of science to nature of science in society. <http://douglasallchin.net/papers/Allchin-2018-NOSIS.pdf>. Zugegriffen: 31. Okt. 2018.
- Arndt, L., Billion-Kramer, T., Wilhelm, M., & Rehm, M. (2019). Antinomien der Naturwissenschaft – Chance zum produktiven und reflektierten Meinungsbildungsprozess im naturwissenschaftlichen Unterricht. *Chemkon*. <https://doi.org/10.1002/ckon.201900020>.
- Artelt, C., Beinicke, A., Schlagmüller, M., & Schneider, W. (2009). Diagnose von Strategiewissen beim Textverstehen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(2), 96–103.
- Baron-Boldt, J., Schuler, H., & Funke, U. (1988). Prädiktive Validität von Schulabschlussnoten: Eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 2, 79–90.
- Bartos, S. A., & Lederman, N. G. (2014). Teachers' knowledge structures for nature of science and scientific inquiry: conceptions and classroom practice. *Journal of Research in Science Teaching*, 51(9), 1150–1184. <https://doi.org/10.1002/tea.21168>.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.
- Bell, R. L., Lederman, N. G., & Abd-El-Khalick, F. (2000). Developing and acting upon one's conception of the nature of science: A follow-up study. *Journal of Research in Science Teaching*, 37(6), 563–581.
- Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16, 78–117.
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. (2015). Beyond dichotomies: competence viewed as a continuum. *Zeitschrift für Psychologie*, 223, 3–13.
- Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, 31(4), 416–440.
- Bromme, R., & Kienhues, D. (2008). Allgemeinbildung. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der pädagogischen Psychologie* (S. 619–628). Göttingen: Hogrefe.
- Bromme, R., Kienhues, D., & Stadler, M. (2016). Die kognitive Arbeitsteilung als Herausforderung für die Forschung zu epistemischen Überzeugungen. In A.-K. Mayer & T. Rosman (Hrsg.), *Denken über Wissen und Wissenschaft – Epistemologische Überzeugungen* (S. 25–38). Lengerich: Pabst.
- Carey, R. L., & Stauss, N. G. (1968). An analysis of the understanding of the nature of science by prospective secondary science teachers. *Science Education*, 52(4), 358–363.
- Carey, R. L., & Stauss, N. G. (1970). An analysis of experienced science teachers' understanding of the nature of science. *School Science and Mathematics*, 70(5), 366–376.
- Central Association for Science and Mathematics Teachers (1907). A consideration of the principles that should determine the courses in biology in secondary schools. *School Science and Mathematics*, 7, 241–247.
- Deng, F., Chen, D.-T., Tsai, C.-C., & Chai, C. S. (2011). Students' views of the nature of science: A critical review of research. *Science Education*, 95(6), 961–999. <https://doi.org/10.1002/sce.20460>.
- von Dijk, E. M. (2011). Portraying real science in science communication. *Science Education*, 95(6), 1086–1100.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer.
- Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science*. Buckingham: Open Univ. Press.
- Eckes, T. (2015). Lokale Abhängigkeit von Items im Test DaF-Leseverstehen. *Diagnostica*, 61(2), 93–106.
- Erduran, S., & Dagher, Z. R. (2014). Reconceptualizing nature of science for science education. In *Reconceptualizing the nature of science for science education. Contemporary trends and issues in science education* Bd. 43. Dordrecht: Springer. https://doi.org/10.1007/978-94-017-9057-4_1.
- Feige, E.-M., Rutsch, J., Dörfler, T., & Rehm, M. (2017). Von der Alltagsvorstellung zum fachwissenschaftlichen Konzept – Schülervorstellungen diagnostizieren und weiterentwickeln. *Naturwissenschaften im Unterricht – Chemie*, 28(159), 2–8.
- Fischler, H., Gebhard, U., & Rehm, M. (2018). Naturwissenschaftliche Bildung und Scientific Literacy. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung*. Berlin: Springer.

- Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382. <https://doi.org/10.1037/h0031619>.
- Friesen, M., Kuntze, S., & Vogel, M. (2018). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 153–177). Wiesbaden: Springer.
- Gamer, M., Lemon, J., & Singh, I.F.P. (2012). irr: Various Coefficients of Interrater Reliability and Agreement. R package Version 0.84. <https://CRAN.R-project.org/package=irr>. Zugegriffen: 2. Feb. 2019.
- Gebhard, U., Höttecke, D., & Rehm, M. (2017). Die Natur der Naturwissenschaft. In U. Gebhard, D. Höttecke & M. Rehm (Hrsg.), *Pädagogik der Naturwissenschaften* (S. 85–106). Wiesbaden: Springer VS.
- Heiligenthal, R., Lemke, F., Schieder, R., & Schneider, T.M. (1999). *Einführung in das Studium der Evangelischen Theologie*. Stuttgart: Kohlhammer.
- Helmke, A. (2012). *Unterrichtsqualität und Lehrprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer.
- Hodson, D. (2011). *Looking to the future. Building a curriculum for social activism*. Rotterdam: Sense.
- Hodson, D., & Wong, S.L. (2017). Going beyond the consensus view: Broadening and enriching the scope of NOS-oriented curricula. *Canadian Journal of Science, Mathematics and Technology Education*, 17(1), 3–17.
- Hofer, B.K. (2001). Personal epistemology research: implications for learning and teaching. *Journal of Educational Psychology Review*, 13(4), 353–383.
- Hofer, B.K., & Pintrich, P.R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140.
- Höfble, C., Höttecke, D., & Kircher, E. (Hrsg.). (2004). *Lehren und Lernen über die Natur der Naturwissenschaften*. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Höttecke, D. (2001). *Die Natur der Naturwissenschaften historisch verstehen. Fachdidaktische und wissenschaftshistorische Untersuchungen*. Berlin: Logos.
- Höttecke, D., & Henke, A. (2010). Über die Natur der Naturwissenschaften lehren und lernen – Geschichte und Philosophie im Chemieunterricht? *Naturwissenschaften im Unterricht – Chemie, Themenheft Natur der Naturwissenschaften*, (4–5), 2–7.
- Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Huber, H.D. (2011). Bildinterpretation. Der Übergang zwischen Wahrnehmung und Sprache. In K. Sachs-Hombach & R. Totzke (Hrsg.), *Bilder – Sehen – Denken. Zum Verhältnis von begrifflich-philosophischen und empirisch-psychologischen Ansätzen in der bildwissenschaftlichen Forschung* (S. 333–348). Köln: Herbert van Halem.
- Irzik, G., & Nola, R. (2011). A family resemblance approach to the nature of science for science education. *Sci & Educ*, 20(7–8), 591–607. <https://doi.org/10.1007/s11191-010-9293-4>.
- Kersting, N.B., Givvin, K.B., Thompson, B.J., Santagata, R., & Stigler, J.W. (2012). Measuring usable knowledge: teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, 49, 568–589.
- Kimball, M. E. (1967) Understanding the nature of science: A comparison of scientists and science teachers. *Journal of Research in Science Teaching*, 5(2), 110–120.
- Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften*. Berlin: Logos.
- Klafki, W. (2007). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Weinheim: Beltz.
- Kleickmann, T., & Anders, Y. (2011). Lernen an der Universität. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 305–315). Münster: Waxmann.
- KMK (2005a). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*
- KMK (2005b). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*
- KMK (2005c). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*
- König, J. (2015). Kontextualisierte Erfassung von Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 61(3), 305–309.
- Kramer, T. (2017). *Epistemologische Urteile von Lehrkräften und Lehramtsstudierenden. Empirische Befunde zu Generierung, Veränderung und Kontextsensitivität*. Berlin: Logos.
- Krapp, A., & Prenzel, M. (2011). Research on interest in science: theories, methods, and findings. *International Journal of Science Education*, 33(1), 27–50.
- Krauss, S., Lindl, A., Schlicher, A., & Tepner, O. (2017). Das Forschungsprojekt FALKO – ein einleitender Überblick. In S. Krauss, A. Lindl, A. Schlicher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R.H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenschaftstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 9–66). Münster: Waxmann.
- Kuhn, T.S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers. Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
- Kuntze, S. (2015). Expertisemerkmale von Mathematiklehrkräften und anforderungshaltige Situierungen – Fragen an Untersuchungsdesigns. In F. Caluori, H. Linneweber-Lammerskitten & C. Streit (Hrsg.), *Beiträge zum Mathematikunterricht 2015* (S. 528–531). Münster: WTM.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Lederman, N.G. (2006). Syntax of nature of science within inquiry and science instruction. In L.B. Flick & N.G. Lederman (Hrsg.), *Scientific inquiry and nature of science: Implications for teaching, learning and teacher education* (S. 301–317). Dordrecht: Springer.
- Lederman, N.G. (2007). Nature of science: past, present, and future. In S.K. Abell & N.G. Lederman (Hrsg.), *Handbook of research in science education* (S. 831–879). Mahwah: Lawrence Erlbaum.
- Lederman, N.G., & Lederman, J.S. (2014). Research on Teaching and Learning of Nature of Science. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of Research on Science Education, Volume II* (pp. 600–620). New York, NY: Routledge.
- Lederman, N. G., & Zeidler, D. L. (1987). Science teachers' conceptions of the nature of science: Do they really influence teacher behavior? *Science Education*, 71(5), 721–734.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002) Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39 (6),497–521.
- Lederman, N.G., Lederman, J.S., & Antink, A. (2013). Nature of science and scientific inquiry as contexts for the learning of

- science and achievement of scientific literacy. *International Journal of Education in Mathematics, Science and Technology*, 1(3), 138–147.
- Leiner, D.J. (2016). SoSci Survey (Version 2.6.00). <https://www.sosicisurvey.de>. Zugegriffen: Computer software.
- Liang, L.L., Chen, S., Chen, X., Kaya, O.N., Adams, A.D., Macklin, M., & Ebenezer, J. (2006). *Student understanding of science and scientific inquiry: revision and further validation of an assessment instrument*. Paper presented at the Annual Conference of the National Association for Research in Science Teaching (NARST), San Francisco.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenz und Kompetenzentwicklung von Lehrerinnen und Lehrern*. Beiheft der Zeitschrift für Pädagogik, (Bd. 51, S. 47–70).
- Mason, L., Pluchino, P., & Ariasi, N. (2014). Reading information about a scientific phenomenon on webpages varying for reliability: An eye-movement analysis. *Educational Research & Development*, 62, 663–685.
- Matthews, M. (2012). Changing the focus: from nature of science (NOS) to features of science (FOS). In M.S. Khine (Hrsg.), *Advances in nature of science research* (S. 3–26). Dordrecht: Springer.
- McComas, W.F. (1998). The principal elements of the nature of science: dispelling the myths. In W.F. McComas (Hrsg.), *The Nature of Science in Science Education*. Science & Technology Education Library, Bd. 5. Dordrecht: Springer.
- McComas, W.F. (2014). Nature of science. In W.F. McComas (Hrsg.), *The language of science education*. Rotterdam: Sense.
- McComas, W.F., & Olson, J.K. (1998). The Nature of Science in International Science Education Standards Documents. In W.F. McComas (Hrsg.), *The Nature of Science in Science Education*. Science & Technology Education Library, Bd. 5. Dordrecht: Springer.
- Menthe, J., Höttecke, D., Eilks, I., & Höble, C. (2013). *Handeln in Zeiten des Klimawandels – Bewerten Lernen als Bildungsaufgabe*. Münster: Waxmann.
- 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.
- Muis, K.R., Bendixen, L.D., & Haerle, F.C. (2006). Domain-generality and domain-specificity in personal epistemology research: philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, 18(1), 3–54.
- National Research Council (1996). *National science education standards: observe, interact, change, learn* (4. Aufl.). Washington, DC: National Academy Press.
- Neumann, I. (2011). Beyond physics content knowledge: Modeling competence regarding nature of scientific inquiry and nature of scientific knowledge. Studien zum Physik- und Chemielernen: Bd. 117. Berlin: Logos.
- Neumann, I., & Kremer, K. (2013). Nature of Science und epistemologische Überzeugungen – Ähnlichkeiten und Unterschiede. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 209–232.
- Neumann, I., Neumann, K., & Nehm, R. (2011). Evaluating instrument quality in science education: Rasch-based analyses of a nature of science test. *International Journal of Science Education*, 33(10), 1373–1405.
- NGSS (2013). The Next Generation Science Standards. <http://www.nextgenscience.org/getto-know>. Zugegriffen: 27. Dez. 2018.
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What “Ideas about Science” should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40(7), 692–720. <https://doi.org/10.1002/tea.10105>.
- Oser, F., Heinzer, S., & Salzmann, P. (2010). Die Messung der Qualität von professionellen Kompetenzprofilen von Lehrpersonen mit Hilfe der Einschätzung von Filmvignetten. *Unterrichtswissenschaft*, 38(1), 5–28.
- Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Pohlmann, B., & Möller, J. (2010). Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). *Zeitschrift für Pädagogische Psychologie*, 24(1), 73–84. <https://doi.org/10.1024/1010-0652.a000005>.
- Popper, K. (1935). *Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft*. Wien: Springer.
- Porsch, T., & Bromme, R. (2011). Effects of epistemological sensitization on source choices. *Instructional Science*, 39(6), 805–819.
- Priemer, B. (2003). Ein diagnostischer Test zu Schüleransichten über Physik und Lernen von Physik – eine deutsche Version des Tests “Views About Science Physics”. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 160–178.
- Priemer, B. (2006). Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 159–175.
- R Core Team (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rehm, M., & Bölsterli, K. (2014). Entwicklung von Unterrichtsvignetten. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 213–225). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-37827-0_18.
- Rehm, M., Brovelli, D., Wilhelm, M., & Marx, C. (2016). Effektive Lehrerinnen- und Lehrerbildung für das integrierte Fach Naturwissenschaften. *Beiträge zur Lehrerinnen-Lehrerbildung*, 34(3), 284–293.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften*. Berlin: Logos.
- Robitzsch, A., & Lüdtke, O. (2015). Kommentar zum Beitrag „Lokale Abhängigkeit von Items im Test DaF-Leseverstehen“ von Thomas Eckes. *Diagnostica*, 61(2), 107–109.
- Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Rutsch, J., Rehm, M., Vogel, M., Seidenfuß, M., & Dörfler, T. (Hrsg.). (2017a). *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen*. Wiesbaden: Springer.
- Rutsch, J., Seidenfuß, M., Vogel, M., Dörfler, T., & Rehm, M. (2017b). Fachdidaktische Unterrichtsvignetten in Forschung und Lehre: Überblick über Forschungsarbeiten und Einsatzmöglichkeiten. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 35(3), 487–505.
- Rutsch, J., Vogel, M., Rehm, M., & Dörfler, T. (2018). Modellierung der Testletstruktur bei vignetten-basierten Testverfahren mit geschlossenem Antwortformat. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 27–46). Wiesbaden: Springer.
- Schödl, A., & Göhring, A. (2017). Falko-P: Fachspezifische Lehrerkompetenzen im Fach Physik. Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R.H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 199–244). Münster: Waxmann.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498–504.

- Schwartz, R., & Lederman, N. G. (2008). What scientists say: Scientists' views of nature of science and relation to science context. *International Journal of Science Education*, 30(6), 727–771.
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60(6), 828–844.
- 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. <https://doi.org/10.3102/0002831214531321>.
- Shavelson, R. J. (2013). On an approach to testing and modeling competence. *Educational Psychologist*, 48(2), 73–86.
- Shen, B. J. (1975). Scientific literacy and the public understanding of science. In S. Day (Hrsg.), *Communication of Scientific Information*. Basel: Karger.
- Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Sjöström, J., & Eilks, I. (2018). Reconsidering different visions of scientific literacy and science education based on the concept of Bildung. In Y. Dori, Z. Mevarech & D. Baker (Hrsg.), *Cognition, Metacognition, and culture in STEM education*. Innovations in science education and technology, (Bd. 24, S. 65–88). https://doi.org/10.1007/978-3-319-66659-4_4.
- Sodian, B., Thoermer, C., Kircher, E., Grygier, P., & Günther, J. (2002). Vermittlung von Wissenschaftsverständnis in der Grundschule. *Zeitschrift für Pädagogik*, 45, 192–206.
- Stahl, E., & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, 17(6), 773–785.
- Stender, A. (2014). *Unterrichtsplanung: Vom Wissen zum Handeln: Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung*. Studien zum Physik- und Chemielernen, Bd. 170. Berlin: Logos.
- Tepner, O., & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 311–323). Berlin: Springer.
- Tröger, H., Strübe, M., Sumfleth, E., & Tepner, O. (2014). Professional knowledge of chemistry teachers video analysis of chemistry lessons. In C. P. Constantinou, N. Papadouris & A. Hadjigeorgiou (Hrsg.), *Strand. 1–16: science education research for evidence-based teaching and coherence in learning*. Proceedings of the ESERA 2013 Conference. (S. 83–88). Nicosia: European Science Education Research Association.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical and psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>.
- Welch, W. W., & Pella, M. O. (1967). The development of an instrument for inventorying knowledge of the processes of science. *Journal of Research in Science Teaching*, 5, 64–68. <https://doi.org/10.1002/tea.3660050115>.
- Wilhelm, M., Rehm, M., & Reinhardt, V. (in Druck). Qualitätsvoller Fachunterricht. In T. Billion-Kramer (Hrsg.), *Wirksamer Sachunterricht*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Wittgenstein, L. (2003). *Philosophische Untersuchungen*. Frankfurt: Suhrkamp.
- Zeidler, D. (2015). Socioscientific issues. In R. Gunstone (Hrsg.), *Encyclopedia of science education* (S. 998–1003). Dordrecht: Springer.