

Bäuerlein, Kerstin; Senn, Sascha; Fraefel, Urban
**Die Beurteilung berufspraktischer Kompetenzen am Ende der
Lehrpersonenausbildung in der Deutschschweiz**

Beiträge zur Lehrerinnen- und Lehrerbildung 37 (2019) 3, S. 401-420



Quellenangabe/ Reference:

Bäuerlein, Kerstin; Senn, Sascha; Fraefel, Urban: Die Beurteilung berufspraktischer Kompetenzen am Ende der Lehrpersonenausbildung in der Deutschschweiz - In: Beiträge zur Lehrerinnen- und Lehrerbildung 37 (2019) 3, S. 401-420 - URN: urn:nbn:de:0111-pedocs-217691 - DOI: 10.25656/01:21769

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-217691>

<https://doi.org/10.25656/01:21769>

in Kooperation mit / in cooperation with:



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

Nutzungsbedingungen

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

Terms of use

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

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

Kontakt / Contact:

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

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

BEITRÄGE ZUR LEHRERINNEN- UND LEHRERBILDUNG

Professionsethos

37. Jahrgang Heft 3/2019

Impressum

Beiträge zur Lehrerinnen- und Lehrerbildung

www.bzl-online.ch

Redaktion

Vgl. Umschlagseite vorn.

Inserate und Büro

Kontakt: Heidi Lehmann, Büro CLIP, Schreinerweg 7, 3012 Bern, Tel. 031 305 71 05,
bzl-schreibbuero@gmx.ch

Layout

Büro CLIP, Bern

Druck

Suter & Gerteis AG, Zollikofen

Abdruckerlaubnis

Der Abdruck redaktioneller Beiträge ist mit Genehmigung der Redaktion gestattet.

Abonnementspreise

Mitglieder SGL: im Mitgliederbeitrag eingeschlossen.

Nichtmitglieder SGL: CHF 80.–; Institutionen: CHF 100.–. Bei Institutionen ausserhalb der Schweiz erhöht sich der Betrag um den Versandkostenanteil von CHF 15.–.

Das Jahresabonnement dauert ein Kalenderjahr und umfasst jeweils drei Nummern.

Bereits erschienene Hefte eines laufenden Jahrgangs werden nachgeliefert.

Abonnementsmitteilungen/Adressänderungen

Schriftlich an: Giesshübel-Office/BzL, Edenstrasse 20, 8027 Zürich oder per Mail an: sgl@goffice.ch.

Hier können auch Einzelnummern der BzL zu CHF 28.–/EUR 28.– (exkl. Versandkosten) bestellt werden (solange Vorrat).

Schweizerische Gesellschaft für Lehrerinnen- und Lehrerbildung (SGL)

www.sgl-online.ch

Die Schweizerische Gesellschaft für Lehrerinnen- und Lehrerbildung SGL wurde 1992 als Dachorganisation der Dozierenden, wissenschaftlichen Mitarbeitenden und Assistierenden der schweizerischen Lehrerinnen- und Lehrerbildungsinstitute gegründet. Die SGL initiiert, fördert und unterstützt den fachlichen Austausch und die Kooperation zwischen den Pädagogischen Hochschulen bzw. universitären Instituten und trägt damit zur qualitativen Weiterentwicklung der Lehrerinnen- und Lehrerbildung bei. Sie beteiligt sich an den bildungspolitischen Diskursen und bringt die Anliegen der Lehrerinnen- und Lehrerbildung in den entsprechenden Gremien ein.

Editorial

- Annette Tettenborn mit Gastredaktor Michael Zutavern, Christian Brühwiler, Bruno Leutwyler, Sandra Moroni, Kurt Reusser, Markus Weil, Markus Wilhelm 303

Schwerpunkt

Professionsethos

- Sarah Forster-Heinzer und Martin Drahmann** Die Bedeutung der pädagogischen Verantwortung für das Berufsethos. Analyse eines konstitutiven und übergreifenden Merkmals in unterschiedlichen Ansätzen zum Ethos von Lehrerinnen und Lehrern 306
- Luciano Gasser und Yvonne Dammert** Ein Prozessmodell zur Entwicklung und Erforschung moralpädagogischer Lehrerinnen- und Lehrerbildung 322
- Alfred Weinberger und Jean-Luc Patry** Partizipative Prozesse im Unterrichtsmodell VaKE als Form eines diskursiven Berufsethos 338
- Ute Bender und Federica Valsangiacomo** Moralische und emotionale Aspekte in Gesprächen in den Bereichen «Ernährung», «Konsum» und «Nachhaltige Entwicklung» – ein Beitrag zur Hochschuldidaktik 352
- Dominik Helbling und Mirjam Schallberger** «Ethik, Religionen, Gemeinschaft» als Lernfeld für das professionelle Ethos von Lehrpersonen 369
- Jürgen Oelkers** Vorbilder, pädagogische Helden und Professionsmoral 383
- Beat A. Schwendimann** Entstehung und Bedeutung des Berufsleitbilds und der Standesregeln des LCH für die Diskussion berufsethischer Fragen des Lehrpersonenhandelns 393

Forum

- Kerstin Bäuerlein, Sascha Senn und Urban Fraefel** Die Beurteilung berufspraktischer Kompetenzen am Ende der Lehrpersonenausbildung in der Deutschschweiz 401
- Tobias Jaschke** Planungsmodell für Lehrkräftefortbildungen mit fachdidaktischen Inhalten 421

Rubriken

Buchbesprechungen

Wermke, W. & Paulsrud, D. (2019). Autonomie im Lehrerberuf in Deutschland, Finnland und Schweden. Entscheidungen, Kontrolle und Komplexität (Christoph Hess) 434

Baumgartner, M. (2017). Performanzentwicklung in der Ausbildung von Lehrkräften. Eine Interventionsstudie zur Verbesserung des Feedbacks bei angehenden Sportlehrkräften (Jürg Baumberger) 436

Steinherr, E. (2017). Werte im Unterricht. Empathie, Gerechtigkeit und Toleranz leben (Sarah Forster-Heinzer) 438

Steffens, U. & Messner, R. (Hrsg.). (2019). Unterrichtsqualität. Konzepte und Bilanzen gelingenden Lehrens und Lernens (Bernadette Gold) 440

Neuerscheinungen 443

Zeitschriftenspiegel 445

Vorschau auf künftige Schwerpunktthemen

Eine Vorschau auf die Schwerpunktthemen künftiger Hefte finden Sie auf unserer Homepage (www.bzl-online.ch). Manuskripte zu diesen Themen können bei einem Mitglied der Redaktion eingereicht werden (vgl. dazu die Richtlinien zur Manuskriptgestaltung, verfügbar auf der Homepage).

Die Beurteilung berufspraktischer Kompetenzen am Ende der Lehrpersonenausbildung in der Deutschschweiz

Kerstin Bäuerlein, Sascha Senn und Urban Fraefel

Zusammenfassung Die Studie nimmt eine Bestandsaufnahme der Verfahren vor, die in der Deutschschweiz zur Beurteilung berufspraktischer Kompetenzen am Ende der Lehrpersonenausbildung eingesetzt werden. Sie geht dabei auf Funktion, Gegenstand und Methodik der Beurteilungsverfahren ein. Zugrunde liegen Daten einer Online-Befragung der für die berufspraktische Ausbildung verantwortlichen Personen aller 15 Hochschulen der Deutschschweiz, die Regelschullehrpersonen ausbilden. Den Angaben zufolge wird auf vielfältige Weise und teilweise unter Nutzung neuerer technischer Möglichkeiten ein meist breites Spektrum an Kompetenzen erfasst, welches in der Regel von mehreren Fachpersonen anhand transparenter Kriterien beurteilt wird. Damit zeichnet sich vor dem Hintergrund des Forschungsstands und mit Blick auf zentrale Gütekriterien eine positive Entwicklung der Beurteilungsqualität seit den Nullerjahren ab.

Schlagwörter Lehrpersonenausbildung – summative Beurteilung – berufspraktische Kompetenzen – Schweiz

Procedures used in German-speaking Switzerland to assess pre-service teachers' professional competence at the end of their studies

Abstract The study gives an overview of the procedures used in German-speaking Switzerland to assess pre-service teachers' professional competence at the end of their studies. It examines the function, the object, and the methods of the assessment procedures. The overview is based on data of an online survey of the persons who are responsible for the practical training at all 15 universities of teacher education in the German-speaking part of Switzerland that prepare regular school teachers. According to the respondents, the universities usually assess a broad spectrum of skills in a variety of ways, in some cases with the help of newer technical means. These assessments are commonly carried out by several experts and mostly rest on transparent criteria. In the light of the current state of research and pertinent quality criteria, these results indicate a positive development in the quality of assessment procedures since the noughties.

Keywords teacher education – summative assessment – professional competence – Switzerland

1 Einleitung

In der einphasigen Lehrpersonenausbildung der Schweiz werden berufspraktische Kompetenzen bereits während des Studiums aufgebaut. In berufspraktischen Studien,

welche je nach Lehrdiplom 15 bis 30% des Studiums ausmachen (EDK, 2019), wird Berufspraxis nicht nur distanziert betrachtet und reflektiert, sondern auch mitgestaltet und geübt (Fraefel, 2012; Leonhard, Fraefel, Jünger, Kosinár, Reintjes & Richiger, 2016). Am Ende des Studiums muss grundlegendes pädagogisches Handeln beherrscht werden, damit im Anschluss daran direkt der Einstieg in den Lehrberuf erfolgen kann.

Die abschliessende Beurteilung der berufspraktischen Kompetenzen erfüllt in der einphasigen Ausbildung eine wichtige Gatekeeper-Funktion, denn anstellende Schulen können die Abschlussnoten zur Einschätzung der professionellen Handlungskompetenz von Bewerberinnen und Bewerbern nutzen. Zudem erhält die Hochschule anhand der Beurteilungen Informationen über die Qualität ihrer Ausbildung (vgl. Fraefel, Bäuerlein & Barabasch, 2018). Aussagekräftige Beurteilungen sind nicht nur für eine faire Beurteilung der Absolventinnen und Absolventen und zur Gewährleistung einer hohen Unterrichtsqualität im Schulfeld geboten: Aus bildungspolitischer Perspektive müssen die Institutionen der Lehrpersonenbildung plausibel darlegen, dass der Einsatz enormer Ressourcen für die Ausbildung zu den erwünschten Resultaten führt und die angehenden Lehrpersonen im Berufsfeld professionell und wirkungsvoll handeln werden.

Gegen die Beurteilungsverfahren in der Lehrpersonenausbildung richtet sich jedoch wiederkehrende Kritik, da sie sich nicht konsequent an Testgütekriterien orientieren (vgl. Bäuerlein & Fraefel, 2016; Metzger & Nüesch, 2004). Strietholt und Terhart (2009) analysierten in Deutschland über 200 Verfahren zur Beurteilung von Referendarinnen und Referendaren und bemängelten, dass die testdiagnostische Güte für keines der Instrumente geprüft worden war. Generell weisen die gemäss der genannten Studie überwiegend eingesetzten Unterrichtsbesuche vielfältige Defizite auf (vgl. Oser, Curcio & Düggeli, 2007; Wei & Pecheone, 2010). So variieren z.B. die Bedingungen, unter welchen angehende Lehrpersonen beurteilt werden, stark. Zudem wird die Fokussierung auf das Unterrichten der Breite beruflicher Anforderungen, die an Lehrpersonen gestellt werden, nicht gerecht. Strietholt und Terhart (2009) resümierten, dass die Qualität der Beurteilungsprozesse aus kompetenzdiagnostischer Perspektive problematisch bis unzureichend sei.

Luhr (2016) zufolge existieren in der deutschen Lehrpersonenausbildung nach wie vor keine adäquaten Verfahren zur Erfassung der Kompetenzen angehender Lehrpersonen. Nicht zuletzt wegen ungünstiger Abhängigkeitsverhältnisse und Rollenkonflikten werden gängige Verfahren sowohl in der Praxis als auch in der Wissenschaft kritisiert. Während in der Forschung zu Lehrpersonenkompetenzen inzwischen verschiedene diagnostische Verfahren angewendet werden, hat sich die Lizenzierung am Ende der formellen Lehrpersonenausbildung in Deutschland kaum verändert (Terhart, 2015).

Für die Schweiz liegen keine vergleichbaren Daten vor. Es gibt lediglich Hinweise darauf, dass sich die Situation in den Nullerjahren ähnlich verhielt wie in Deutschland (vgl. Fraefel & Huber, 2008). Angesichts differenter Systeme stellt sich die Problem-

lage in beiden Ländern jedoch unterschiedlich dar. Ziel der vorliegenden explorativen Studie ist daher eine Bestandsaufnahme der gegenwärtig in der Lehrpersonenausbildung der Deutschschweiz eingesetzten Verfahren zur summativen Beurteilung berufspraktischer Kompetenzen am Ende des Studiums. Das Interesse richtet sich dabei insbesondere auf zwei Fragen:

- 1) Haben vor dem Hintergrund grosser Anstrengungen in der forschungsbasierten Kompetenzdiagnostik neuere Verfahren zur Beurteilung berufspraktischer Kompetenzen Eingang in die Praxis gefunden?
- 2) Inwieweit erfüllen die derzeit in der Praxis eingesetzten Verfahren testtheoretische Gütekriterien?

2 Theoretischer Hintergrund

Im Folgenden wird basierend auf dem aktuellen Forschungsstand dargestellt,

- welches die zentralen Gütekriterien für Prüfungsverfahren sind,
- welche Funktionen die summative Beurteilung berufspraktischer Kompetenzen am Ende der Lehrpersonenausbildung erfüllen kann,
- was Gegenstand der Beurteilung sein sollte,
- nach welchen Kriterien die Beurteilung erfolgen sollte,
- wie entsprechende Kompetenzen erfasst werden können und
- welche Personen diese Kompetenzen in welcher Form beurteilen sollten, um eine faire und angemessene Beurteilung zu gewährleisten.

2.1 Gütekriterien

Für eine aussagekräftige und faire Prüfung berufspraktischer Kompetenzen liegt es nahe, die klassischen Testgütekriterien heranzuziehen (vgl. Metzger & Nüesch, 2004; Roloff, 2012a, 2012b). Da diese sich z.B. bei Unterrichtsbeurteilungen nicht immer einhalten bzw. überprüfen lassen, schlagen Bovet und Frommer (2015) für diesen Zweck alternative Gütekriterien vor, die eher auf qualitative Forschungsmethoden rekurrieren und die klassischen Gütekriterien teils ergänzen, teils substituieren können.

Objektivität bedeutet im Prüfungskontext, dass die Note unabhängig von der prüfenden Person ist (Roloff, 2012a, 2012b). Dies ist durch Standardisierung der Durchführung und Auswertung zu erreichen (Metzger & Nüesch, 2004). Alternativ können nach Bovet und Frommer (2015) verschiedene Beurteilende aus ihren subjektiven Einzelurteilen kommunikativ ein *intersubjektiv nachvollziehbares Gesamturteil* generieren, wobei die resultierenden Gesamturteile in sich sowie mit Bezug auf äussere Faktoren stimmig sein sollen.

Die *Reliabilität* steigt mit der Anzahl der Prüfenden und ist gegeben, wenn bei einer vergleichbaren Prüfung, einer Prüfungswiederholung oder einer Aufteilung der Prüfung in mehrere Teile sowie bei einer Beurteilung durch mehrere Prüfende jeweils eine

vergleichbare Note resultiert (Metzger & Nüesch, 2004; Roloff, 2012a, 2012b). Alternativ dazu kann der *Realitätsgehalt* betrachtet werden. Dieser ist hoch, wenn die Unterrichtskompetenz in diversen, den realen beruflichen Anforderungen entsprechenden Kontexten beobachtet wird (Bovet & Frommer, 2015).

Validität liegt vor, wenn die Prüfung genau die berufspraktischen Kompetenzen abbildet. Dabei sollte eine hohe inhaltliche Übereinstimmung mit festgelegten, theoretisch fundierten Kompetenzziele und Ausbildungsinhalten gegeben sein (Roloff, 2012a, 2012b). Zudem ist es wichtig, dass die beurteilten Kompetenzen repräsentativ und relevant sind für die praktischen Handlungsanforderungen (Walzik, 2012). Bovet und Frommer (2015) schlagen alternativ eine *kommunikative und ökologische Validierung* vor, wobei im Austausch mit der beurteilten Person geprüft wird, ob ihre Intentionen und ihr Handeln korrekt erfasst wurden, und bei der Beurteilung zudem Kontextbedingungen berücksichtigt werden.

Als weitere Kriterien sind *Ökonomie* (angemessene Aufwand-Nutzen-Relation), *Nutzen* (Erfüllung mindestens einer wichtigen Funktion) und *Fairness* (Chancengerechtigkeit) für die summative Prüfung berufspraktischer Kompetenzen am Ende des Studiums relevant (vgl. Metzger & Nüesch, 2004; Roloff, 2012a, 2012b). Grundsätzlich können nie alle Gütekriterien optimal erfüllt sein. Für eine faire und aussagekräftige Beurteilung gilt es, eine angemessene Balance zu finden (Walzik, 2012). Viele hochwertige forschungsbasierte Kompetenzmessverfahren nutzen anspruchsvolle qualitative und quantitative Methoden; für die Praxis sind jedoch einfach handhabbare und ressourcenschonende Verfahren erforderlich, die eine faire und angemessene Aussage darüber erlauben, inwiefern zentrale berufspraktische Kompetenzen aufgebaut wurden (vgl. Luhr, 2016).

2.2 Beurteilungsfunktion

Wie bereits angeklungen, kann die summative Beurteilung berufspraktischer Kompetenzen angehender Lehrpersonen am Ende der Ausbildung verschiedene Funktionen erfüllen (vgl. z.B. Fraefel et al., 2018; Strietholt & Terhart, 2009):

- *Evaluativ* kann die Beurteilung als Grundlage dienen, um zu prüfen, welche Kompetenzen in der Ausbildung aufgebaut wurden, und somit Hinweise auf die Ausbildungsqualität liefern.
- *Normativ* können transparente Beurteilungskriterien Ausbildenden und Studierenden als Orientierung dienen, da sie zeigen, welche Kompetenzen für die Ausübung des Lehrberufs als zentral erachtet werden.
- *Selektiv* ist die Beurteilung, wenn anhand der Resultate entschieden wird, ob eine angehende Lehrperson ein Zeugnis erhält, das den Eintritt in den Lehrberuf erlaubt.
- *Prognostisch* liefert die erteilte Note anstellenden Schulen idealiter Informationen über die zu erwartende professionelle Handlungskompetenz von Bewerberinnen und Bewerbern.
- *Formativ* kann die Beurteilung Stärken und Schwächen und somit Ansatzpunkte für die künftige Weiterentwicklung aufzeigen.

Die verschiedenen Funktionen schliessen sich nicht gegenseitig aus. Von einer Vermischung formativer und summativer Beurteilungen wird jedoch abgeraten, da dies zu Ziel- und Rollenkonflikten führen kann (vgl. Müller, 2011).

2.3 Beurteilungsgegenstand

International existieren verschiedene Kompetenz- und Standardmodelle (vgl. Frey, 2014). Im deutschsprachigen Raum dominiert derzeit ein im Rahmen der COACTIV-Studie entwickeltes, inzwischen erweitertes und spezifiziertes Modell den Diskurs (Baumert & Kunter, 2006; Rothland, Cramer & Terhart, 2018). Diesem Modell zufolge setzt sich professionelle Kompetenz aus Professionswissen, Überzeugungen, motivationalen Orientierungen und Selbstregulation zusammen. Studien nehmen häufig einzelne Tätigkeits- bzw. Kompetenzbereiche des Lehrberufs isoliert in den Blick, z.B. die diagnostische Kompetenz (von Aufschnaiter et al., 2015). Trotz zahlreicher Forschungsbemühungen bleibt das Konstrukt der professionellen Kompetenz von Lehrpersonen bislang jedoch diffus (vgl. Forster-Heinzer & Oser, 2015).

Gleichwohl ist es für die Sicherstellung der Ausbildungsqualität erforderlich, Standards im Sinne spezifischer, messbarer Kompetenzziele zu definieren (KMK, 2004). Bereits Anfang der 1990er-Jahre erarbeitete das U.S. Interstate Teacher Assessment and Support Consortium (InTASC) Standards für den Lehrberuf. Das fach- und stufenunabhängige Modell von Kernkompetenzen beeinflusste viele darauffolgende Modelle und beinhaltet in der aktuellen Version zehn Standards zu den in Tabelle 1 genannten Aspekten (InTASC, 2013). In Deutschland stellen von der Kultusministerkonferenz formulierte Standards einen verbindlichen Orientierungsrahmen für die Lehrpersonenausbildung dar (KMK, 2004). Die KMK identifizierte elf Kompetenzen, die sich in vier Bereiche gliedern lassen (vgl. Tabelle 1). Jeder Kompetenz sind wiederum drei bis fünf Standards zugeordnet. In der Schweiz gibt es keine allgemein verbindlichen Standards; vielmehr formulieren die Ausbildungsinstitute jeweils eigene Standards. Tabelle 1 führt exemplarisch jene der Pädagogischen Hochschule Zürich auf (PH Zürich, 2017). Bei allen in Tabelle 1 dargestellten Modellen liegt der Schwerpunkt auf Aspekten, die in einem Zusammenhang mit Unterricht stehen. Dies geht konform mit dem Berufsauftrag für Lehrpersonen, dem zufolge Unterricht ca. 80 bis 85% der Arbeit einer Lehrperson ausmacht (z.B. LCH, 2014). In allen drei Modellen sind jedoch auch darüber hinausgehende Aspekte vertreten, wenn auch unterschiedliche.

2.4 Beurteilungskriterien

Beurteilungskriterien basieren meist auf Qualitätsmodellen von Unterricht, die seit den 1980er-Jahren an Beachtung gewonnen haben (z.B. Porter & Brophy, 1988) und normative Vorstellungen guten Unterrichts zunehmend infrage stellen. Mittlerweile sind zahlreiche mehr oder weniger evidenzbasierte Kriterienlisten zur Unterrichtsbeurteilung (z.B. Helmke, 2015; Kunter, Baumert, Blum, Klusmann, Krauss & Neubrand, 2011; Meyer, 2018) sowie verschiedene, teilweise evaluierte fachspezifische oder generische Instrumente entstanden (vgl. Praetorius & Charalamous, 2018). Einige hochinferente

Tabelle 1: Übersicht über Kompetenzmodelle

InTASC (2013)	KMK (2004)	PH Zürich (2017)
1. Entwicklung der Lernenden	1. Unterrichten z.B. Lernsituationen gestalten	1. Fachspezifisches Wissen und Können
2. Interindividuelle Unterschiede	2. Erziehen z.B. Schwierigkeiten und Konflikte lösen	2. Lernen, Denken und Entwicklung
3. Lernumgebungen	3. Beurteilen z.B. transparente Beurteilung	3. Motivation und Interesse
4. Fachwissen	4. Innovieren z.B. Bewusstsein für das öffentliche Amt	4. Heterogenität
5. Fachdidaktik		5. Kooperation, Partizipation und soziales Umfeld
6. Diagnostik		6. Kommunikation
7. Unterrichtsplanung		7. Planung und Durchführung von Unterricht
8. Unterrichtsgestaltung		8. Diagnose und Beurteilung
9. Professionelle (Weiter-)Entwicklung		9. Sicherstellung der Qualität und professionellen Weiterentwicklung
10. Führung und Kooperation		10. Schule und Gesellschaft
		11. Schule als Organisation
		12. Beruf in der Lebensbalance

Tools zur Unterrichtsbeurteilung sind inzwischen gut handhabbar (z.B. CLASS; vgl. Pianta, Hamre & Allen, 2012). Zu anderen Bereichen berufspraktischer Kompetenzen (z.B. «Innovieren» oder «Führung und Kooperation») sind keine systematisch erprobten Kriterienraster bekannt.

Luhr (2016) stellt die Frage, inwiefern existierende Kriterienlisten geeignet seien, um den Unterricht *angehender* Lehrpersonen zu beurteilen. Er plädiert für eine Auswahl zentraler Merkmale, die alle Lehrpersonen zwingend zu realisieren in der Lage sein sollten. Darüber hinaus sollten die Kriterien aus Validitätsgründen auf die jeweiligen Kompetenzmodelle der Studiengänge abgestimmt sein (Roloff 2012a, 2012b). Unabhängig von der Auswahl erscheint es mit Blick auf die Gütekriterien erforderlich, dass einheitliche Kriterien für eine summative Abschlussbeurteilung innerhalb eines Studiengangs vorliegen.

2.5 Kompetenzerfassung

Ein fundamentales Problem der Kompetenzerfassung ist das Dilemma zwischen einer standardisierten, reliablen Messung und dem Vermeiden einer validitätsreduzierenden Standardisierung (vgl. König, 2015). Dieser Spagat ist in Bezug auf den Lehrberuf besonders schwierig, da Handlungen von Lehrpersonen komplex und situiert sind.

Schriftliche Tests, mündliche Prüfungen und Fragebögen erlauben die ökonomische Erfassung relevanten Wissens (vgl. z.B. Hohenstein, Kleickmann, Zimmermann, Köller & Möller, 2017) sowie weiterer persönlicher Ressourcen (z.B. logisches Denken, Persönlichkeit, vgl. Frey, 2006). Inwiefern daraus Schlüsse auf tatsächliches Handeln in realen Anforderungssituationen zulässig sind, ist nachzuweisen (Forster-Heinzer & Oser, 2015).

Praktische Prüfungen haben den Vorteil, dass sie eine hohe berufliche Relevanz aufweisen und Wechselwirkungen von Kompetenzen sowie überfachliche Kompetenzen sichtbar machen (den Ouden, 2019). Nachteile liegen in der Unvorhersehbarkeit der Abläufe, im hohen Prüfungsaufwand und in der Komplexität, welche eine den Gütekriterien genügende Beurteilung erschwert. Eine Beobachtung über einen längeren Zeitraum sowie der Einbezug mehrerer Lektionen und unterschiedlicher Fächer erscheinen mit Blick auf den Realitätsgehalt und die Validität günstig (vgl. Praetorius, Pauli, Rakoczy & Klieme, 2014; Praetorius, Vieluf, Saß, Bernholt & Klieme, 2016).

Traditionell werden praktische Prüfungen in der Lehrpersonenausbildung in Form von Unterrichtsbesuchen, d.h. *Verhaltensbeobachtungen in realen Anforderungssituationen*, durchgeführt (vgl. z.B. Strietholt & Terhart, 2009). Diese ermöglichen eine hohe Authentizität, sind jedoch schwierig zu standardisieren, zeitintensiv und anspruchsvoll für die Prüfenden. Beobachtungen im Arbeitsalltag durch Auszubildende (z.B. Praxislehrpersonen, Schulleitung) sind ebenfalls authentisch, jedoch in der Regel subjektiv gefärbt und noch weniger standardisiert. Zunehmend werden mittlerweile *Videoaufzeichnungen realen Unterrichts* für die Beurteilung herangezogen (z.B. Bäuerlein & Fraefel, 2016). Dabei stellen sich dann z.B. Fragen der Reaktivität, des eingeschränkten Kamerablickwinkels und des Datenschutzes (vgl. z.B. Fraefel et al., 2018; Praetorius, McIntyre & Klassen, 2017). *Verhaltensbeobachtungen in realitätsnahen Simulationen* (z.B. Rollenspiele, Computersimulationen) erlauben einen hohen Standardisierungsgrad und eine breite Anforderungsvariation, eine authentische Gestaltung ist jedoch schwierig.

Eine Alternative zur praktischen Prüfung ist die Erfassung der *Analysekompetenz*. Sie gilt als Indikator für Professionalität, da Lehrpersonen mit hoher Analyse- und Reflexionskompetenz lernwirksamer unterrichten (z.B. Kersting, Givvin, Thompson, Santagata & Stigler, 2012). Die Analyse von Videos und Vignetten gewährleistet eine gewisse Situietheit und Authentizität bei gleichzeitiger Standardisierung und Reliabilität (vgl. z.B. König, 2015; Rutsch, Rehm, Vogel, Seidenfuß & Dörfler, 2018). Es können

Unterrichtsvignetten in Text-, Video- oder Comicform oder komplette *videografierte Lektionen* eingesetzt werden (z.B. Plöger, Scholl & Seifert, 2016; Rutsch et al., 2018). Die Analyse von *Videos eigenen Unterrichts* eignet sich besonders gut, um Erfahrungen und Vorwissen bezogen auf die Unterrichtssituation zu aktivieren (Kleinknecht & Poschinski, 2014).

Videos und Unterrichtsanalysen können auch in *Portfolios* einfließen. Diese sind in Papierform oder elektronisch, rein schriftlich oder als Kombination von Text und Videos möglich. Rein schriftliche Portfolios oder textbasierte Selbsteinschätzungen scheinen jedoch eher Einstellungen und Überzeugungen widerzuspiegeln als tatsächliche eigene Performanz (Admiraal & Berry, 2016). In der Ausbildungspraxis erleben Portfolios im deutschsprachigen Raum derzeit vor allem in formativer Funktion einen Aufschwung (Feder & Cramer, 2018). Eine faire Beurteilung von Portfolios ist didaktisch anspruchsvoll (Kopf, Leopold & Seidl, 2010). *Assessment Center* und *standardisierte Beurteilungssysteme mit einer Beurteilung durch externe Fachpersonen bzw. Institute* sind vor allem aus den USA bekannt (z.B. Hutt, Gottlieb & Cohen, 2018; Porter, Youngs & Odden, 2001).

2.6 Beurteilende

Bei der Beurteilung stellen generell typische Urteilsfehler (z.B. Milde-Streng-Fehler, Halo-Effekt) ein Problem dar – insbesondere bei hochinferenten Verfahren (vgl. Lotz, Gabriel & Lipowksy, 2013). Der Vorteil speziell geschulter und erfahrener externer Beurteilender geht allerdings häufig mit dem Fehlen notwendiger Hintergrund- und Kontextinformationen einher (z.B. Praetorius, 2013). Für valide und reliable Urteile über Unterrichtsqualität ist es daher günstig, mehrere spezifisch ausgewählte, geschulte und zertifizierte Beurteilende einzusetzen (z.B. Hill, Charalambous & Kraft, 2012). Aus theoretischer Sicht erscheint es sinnvoll, wenn die Prüfenden sowohl über erziehungswissenschaftliche Kompetenzen als auch über Kenntnisse von Unterrichtssituationen verfügen (Neuenschwander, 2004). Zudem kann eine personale Trennung von formativen und summativen Feedbacks Funktions-, Ziel- und Rollenkonflikten vorbeugen (Müller, 2011).

2.7 Beurteilungsfom

Beurteilungsraster – d.h. Beurteilungsblätter mit Kriterien, Operatoren und vorgegebenem Beurteilungsformat – sollen eine differenzierte Erfassung und Beurteilung eines komplexen, multidimensionalen Geschehens erleichtern und die Urteilsreliabilität erhöhen (Messner, 2007). Strietholt und Terhart (2009) unterschieden hinsichtlich der Beurteilungsfom zwischen Freitextkommentaren (zu mehr oder weniger eng spezifizierten Bereichen), Ratingskalen und Kompetenzniveaus (Beschreibungen, die festhalten, welche Anforderungen eine Person mit einer bestimmten Kompetenz bewältigen kann und welche nicht). Die Beurteilungsfomen wurden weiter in qualitative (Freitextkommentare), quantitative (Ratingskalen und Kompetenzniveaus) und hybride (Kom-

bination von qualitativ und quantitativ) Formate unterteilt. Während Freitextkommentare eine starke Individualisierung erlauben, jedoch hinsichtlich der Fairness kritisch zu beurteilen sind, ermöglichen quantitative Varianten interindividuelle Vergleiche, sind aber schwerer zu entwickeln und weniger flexibel. Anhand einer Kombination liessen sich die Vorteile beider Formen nutzen.

3 Methode

Für die vorliegende Studie wurden im Sommer 2017 die für die berufspraktische Ausbildung verantwortlichen Personen aller 15 Ausbildungsinstitute für Regelschullehrpersonen der Deutschschweiz online angefragt, ob sie dazu bereit seien, freiwillig und potenziell anonym Angaben zu den gegenwärtig eingesetzten Verfahren zur Prüfung berufspraktischer Kompetenzen am Ende des Studiums zu machen. Von 53 kontaktierten Personen füllten 34 (64.2%) den Fragebogen aus.

Meist ist innerhalb einer Hochschule eine Person für die berufspraktische Ausbildung für eine Schulstufe bzw. einen Studiengang verantwortlich. Angaben von Personen, die für zwei Stufen verantwortlich sind, flossen pro Stufe einmal (und somit insgesamt doppelt) in die Analysen ein. An einer Hochschule ist für die Sekundarstufe II (Sek II) in jedem Fachbereich eine andere Person für die berufspraktische Ausbildung verantwortlich, weshalb hier bei sieben Hochschulen insgesamt 19 Personen angeschrieben wurden. Somit ist dieser Studiengang überrepräsentiert. Tabelle 2 zeigt, für welche Studiengänge Angaben vorliegen. Da in zwei Studiengängen¹ den Angaben zufolge keine berufspraktische Abschlussprüfung stattfindet und für sechs Studiengänge der Fragebogen nur teilweise ausgefüllt wurde, beträgt die Stichprobengrösse meist 33. Bei unvollständigen Datensätzen wurden alle verfügbaren Angaben einbezogen.

Tabelle 2: Stichprobe

Studiengang	N	Rücklauf (%)
Kindergarten/Eingangsstufe (KE)	13	100.0
Primarstufe (Primar)	11	84.6
Sekundarstufe I (Sek I)	8	100.0
Sekundarstufe II (Sek II)	9	47.4
Studiengänge	41	77.4
Hochschulen	11	73.3

¹ Zugunsten der Übersichtlichkeit wird im Folgenden die Bezeichnung «Studiengänge» verwendet, obwohl es sich bei «Sek II» zum Teil um Fachbereiche handelt.

Die Fragen bezogen sich auf das derzeit eingesetzte Prüfungsverfahren (Beurteilungsfunktion, Beurteilungsgegenstand und Beurteilungsmethode). Soweit dies möglich war, wurde ein geschlossenes Antwortformat (mit fakultativem Freitextfeld) gewählt, um eine kurze Ausfüllzeit und einen hohen Rücklauf zu erzielen. Die Antwortoptionen wurden aus den im Theorieteil dargestellten Überlegungen und Befunden abgeleitet und können den Tabellen im Ergebnisteil entnommen werden.² Die Daten wurden deskriptiv analysiert und mit Standards verglichen. In der Diskussion werden sie vor dem Hintergrund des Forschungsstands und zentraler Gütekriterien bewertet.

4 Ergebnisse

Die Ergebnisdarstellung ist analog zum theoretischen Hintergrund in die Aspekte «Beurteilungsfunktion», «Beurteilungsgegenstand», «Beurteilungskriterien», «Kompetenz-erfassung», «Beurteilende» und «Beurteilungsform» gegliedert.

4.1 Beurteilungsfunktion

Die Befragung bezog sich auf die abschliessende summative Beurteilung der berufspraktischen Kompetenzen. Häufig scheint diese Beurteilung jedoch über die Selektionsfunktion hinaus auch formativ interpretiert zu werden, da die Studierenden den Angaben zufolge meist eine über die Note hinausgehende Rückmeldung erhalten (können). In drei Studiengängen (9.1%) gibt es ausser der Note kein Feedback. In 26 Studiengängen (78.8%) ist dies hingegen der Fall, wenngleich vereinzelt nur auf Verlangen bzw. bei Nichtbestehen. In vier Studiengängen (12.1%) werden Portfolios beurteilt, die über mehrere Semester bzw. über das gesamte Studium hinweg erstellt werden, was ebenfalls auf eine Vermischung formativer und summativer Verfahren hindeutet. Darüber hinaus können die mit einer Ausnahme in allen Studiengängen existierenden verbindlichen Beurteilungskriterien im Sinne konkreter Ausbildungsziele potenziell den Zweck eines Orientierungsrahmens erfüllen, da sie den Studierenden in allen Fällen zugänglich sind.

4.2 Beurteilungsgegenstand

Auf die Frage danach, welchen Bereichen sich die in der Schlussbeurteilung geprüften Kompetenzen zuordnen lassen, wurden in der vorliegenden Studie in Anlehnung an Strietholt und Terhart (2009), die sich auf die KMK-Standards (KMK, 2004) gestützt und diese induktiv erweitert hatten, die in Abbildung 1 ersichtlichen Antwortoptionen angeboten. Unterrichtsliche Kompetenzen sind der einzige Bereich, der in allen Studiengängen geprüft wird. Fachdidaktische Kompetenzen werden bei Sek II im jeweiligen Studienfach bzw. in allen Studienfächern geprüft, bei Sek I in zwei ausgewählten oder in (allen) vier Studienfächern. Bei Primar und KE variieren Fächerzahl und Fächerauswahl zwischen den Hochschulen stärker (z.B. «in allen sechs Fachbereichen» oder «im

² Der Fragebogen kann bei der Erstautorin angefordert werden.

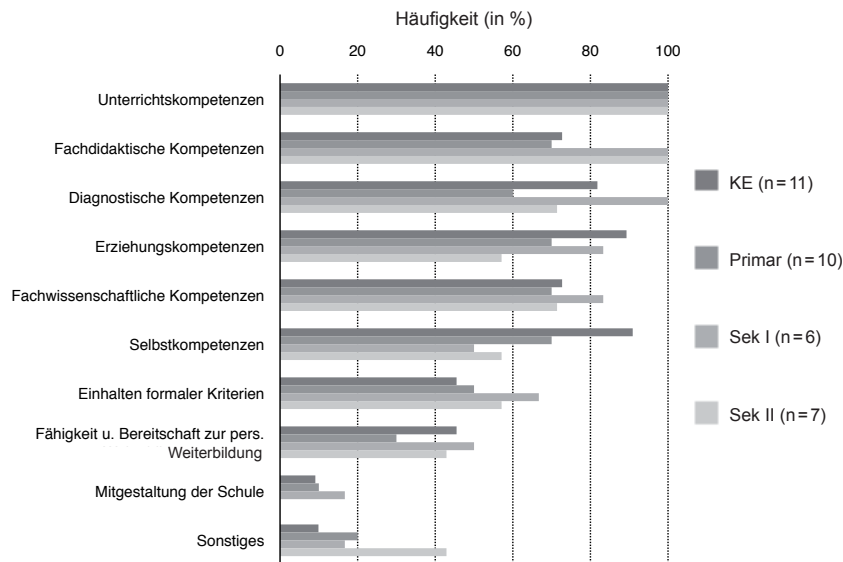


Abbildung 1: Häufigkeit, mit der die einzelnen Kompetenzbereiche in den verschiedenen Studiengängen geprüft werden (Mehrfachwahl; unter «Sonstiges» fallen z.B. Engagement oder professionelles Berufsverständnis).

Fach, in dem die Diplomalaktion stattfindet)). Für die Prüfung fachwissenschaftlicher Kompetenzen variieren Fächerzahl und Fächerauswahl analog.

Orientiert an Strietholt und Terhart (2009) wurden die im Fragebogen genannten Aspekte den vier KMK-Kompetenzbereichen «Unterrichten», «Erziehen», «Beurteilen» und «Innovieren» zugeordnet (KMK, 2004). Anschliessend wurde geprüft, in wie vielen Studiengängen die einzelnen Bereiche abgedeckt werden. Ein Bereich galt als abgedeckt, wenn mindestens eine der von der KMK im jeweiligen Bereich genannten Kompetenzen geprüft wird. Wie in Abbildung 2 zu sehen ist, ist lediglich der Kompetenzbereich «Unterrichten» in allen Studiengängen Teil der Beurteilung, wobei jeweils mindestens ein weiterer Bereich berücksichtigt wird. In zwei Dritteln der Fälle fließen sogar Aspekte aus allen vier Bereichen in die Beurteilung ein.

4.3 Beurteilungskriterien

Wie bereits erwähnt wurde, liegen laut den Angaben der Verantwortlichen mit einer Ausnahme für alle Studiengänge verbindliche Beurteilungskriterien vor. Diesen liegen in 28 Studiengängen (87.5%) zumindest u.a. wissenschaftliche Theorien zugrunde (genannt wurden z.B. Helmke, Meyer und Reusser). Ein Instrument wurde jedoch rein intuitiv bzw. erfahrungsbasiert entwickelt, und bei drei Instrumenten ist den Verantwortlichen nicht bekannt, wie sie entstanden sind bzw. worauf sie basieren.

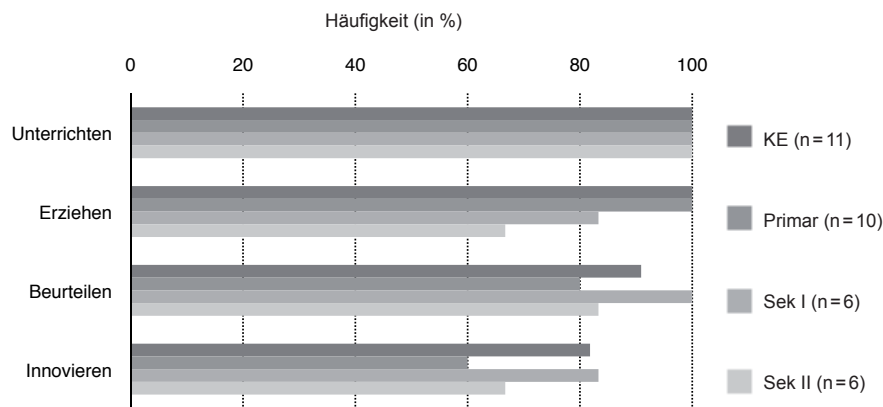


Abbildung 2: Häufigkeit, mit der die einzelnen KMK-Kompetenzbereiche geprüft werden.

4.4 Kompetenzerfassung

Die Angaben zur Frage, aus welchen Bestandteilen sich der Abschluss der berufspraktischen Ausbildung zusammensetzt, sind Abbildung 3 zu entnehmen. Die fakultativen Zusatzangaben lassen sich wie folgt zusammenfassen:

- *Live-Beobachtungen* finden bei Sek II für zwei bis drei Lektionen statt, bei Sek I für eine bis drei Lektionen. Bei Primar und KE variiert der Umfang jeweils zwischen einer Lektion und drei Halbtagen.
- Meist findet zusätzlich zum Unterrichtsbesuch ein *Reflexionsgespräch* statt, das 10 bis 90 Minuten dauert. Ablauf und Dauer sind zum Teil nicht standardisiert.
- Häufig ist die *mündliche Prüfung* an das Reflexionsgespräch gekoppelt, wobei in 15 bis 60 Minuten teils Allgemeindidaktik, teils Fachdidaktik geprüft wird.
- Die *schriftliche Prüfung* enthält überwiegend geschlossene, aber auch wenige offene Fragen.
- *Praktikumsbeurteilungen* beziehen sich hauptsächlich auf ein Praktikum und nur vereinzelt auf mehrere Praktika. Dabei handelt es sich um Blockpraktika (zwei bis acht Wochen), Tagespraktika (z.B. zwei Tage pro Woche im Zeitraum von 15 Wochen), eine Kombination aus beidem oder bis zu 50 aufaddierte Einzellektionen.
- *Portfolios* werden etwa in der Hälfte der Fälle in Papierform eingereicht und sonst in Form von Offline- oder Online-E-Portfolios. Vereinzelt enthalten die Portfolios Videos bzw. sind Videos optional.
- *Studentische Unterrichtsvideos* werden in allen Fällen von Fachpersonen beurteilt, wobei jeweils zusätzlich eine darauf basierende *schriftliche Reflexion der Studierenden* begutachtet wird. In zwei Fällen werden zwei videografierte Lektionen beurteilt, in einem Fall 30 Minuten Videomaterial, welches die Studierenden selbst begründet auswählen.
- Eine *Simulation* findet während des Semesters im Umfang von insgesamt vier Stunden statt, wobei schwierige Situationen des Schulalltags nachgestellt werden.



Abbildung 3: Häufigkeit der Prüfungsbestandteile (Mehrfachwahl).

4.5 Beurteilende

Die Anzahl der Beurteilenden streut von eins bis fünf. Der Modus liegt bei zwei Prüfenden ($N = 19$; 63.6%). Lediglich in drei Studiengängen (9.1%) prüft nur eine Person und in einem Studiengang (3.0%) sind fünf Personen an der Beurteilung beteiligt. Die Prüfenden haben in der Regel einen unterschiedlichen fachlichen Hintergrund und Bezug zu den Lehramtsstudierenden. Dozierende aus der Fachdidaktik prüfen in allen Studiengängen (vgl. Abbildung 4). Die zweithäufigste Personenkategorie sind Praxislehrpersonen, welchen vereinzelt jedoch nur eine beratende Funktion zukommt. Dabei handelt es sich teils um diejenigen Praxislehrpersonen, die die Studierenden im Schulfeld betreut haben, und teils um fremde Praxislehrpersonen.

Bei 20 der 30 Studiengänge (66.7%), in welchen mehrere Personen in die Beurteilung involviert sind, findet zur Festlegung der finalen Note eine diskursive Einigung statt. Bei sieben Studiengängen (23.3%) wird der Mittelwert aus den Beurteilungen der Prüfenden gebildet. In 14 Fällen (46.7%) sind andere Vorgehensweisen definiert, z.B. hat eine Person den Lead oder die Prüfenden vergeben Teilnoten für unterschiedliche Prüfungsbereiche, welche dann zu einer Gesamtnote integriert werden.

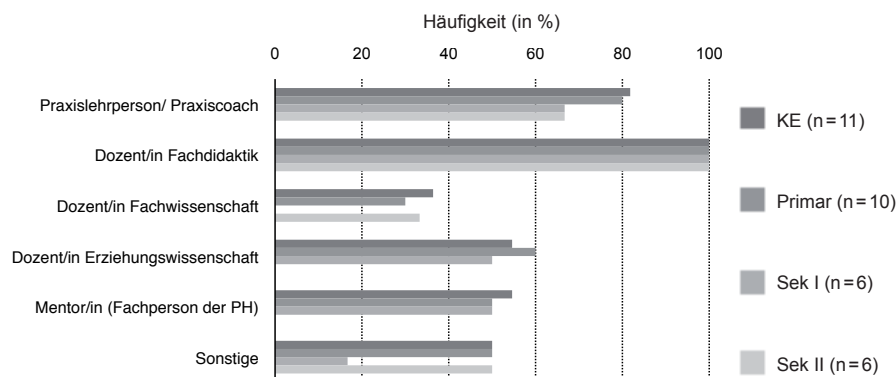


Abbildung 4: Beurteilende Personen (unter «Sonstige» fallen z.B. Personen aus dem Schulfeld mit besonderer Qualifikation oder Mitglieder des Schulinspektorats).

In 17 Studiengängen (51.5%) kennen sich Studierende und Prüfende in der Regel aus Lehrveranstaltungen oder Praktika, in drei Fällen (9.1%) ist dies in der Regel nicht der Fall und in 13 Fällen (39.4%) teilweise (z.B. prüft eine bekannte gemeinsam mit einer fremden Praxislehrperson oder es ist Zufall). In zehn Fällen (30.3%) haben die Prüfenden ausschliesslich eine beurteilende Funktion inne, während ihnen in 23 Fällen (69.7%) zusätzlich eine unterstützende Funktion im Rahmen der für die Abschlussnote relevanten Leistungen zukommt (z.B. Beratung, Zwischenfeedback oder Praktikumsbegleitung). In zwei Studiengängen (6.1%) ist die Rolle der Beurteilenden nicht festgelegt bzw. die Handhabung uneinheitlich. In allen 33 Studiengängen findet eine Schulung der Prüfenden statt, welche den Angaben zufolge zwischen einer und 60 Stunden dauert ($M = 7.21$, $SD = 13.81$, $N = 17$) und in drei Fällen freiwillig besucht wird. Vereinzelt wurde erwähnt, dass die Urteile trotz der Schulung häufig voneinander abweichen würden.

4.6 Beurteilungsform

Abbildung 5 zeigt die Antwortoptionen und die Antworthäufigkeiten bezüglich der Beurteilungsform. Den Angaben zufolge findet die Beurteilung in drei Fällen (9.1%) ausschliesslich qualitativ (Freitext) und in 13 Fällen (39.4%) ausschliesslich quantitativ (Ratingskala oder Kompetenzniveaus) statt. In 17 Studiengängen (51.5%) werden beide Formen kombiniert.

5 Diskussion

Der vorliegende Beitrag stellte in einem Überblick dar, wie die Institutionen der Lehrpersonenausbildung der Deutschschweiz die berufspraktischen Kompetenzen ihrer Studierenden am Ende des Studiums summativ prüfen. Den Datenauswertungen zufolge

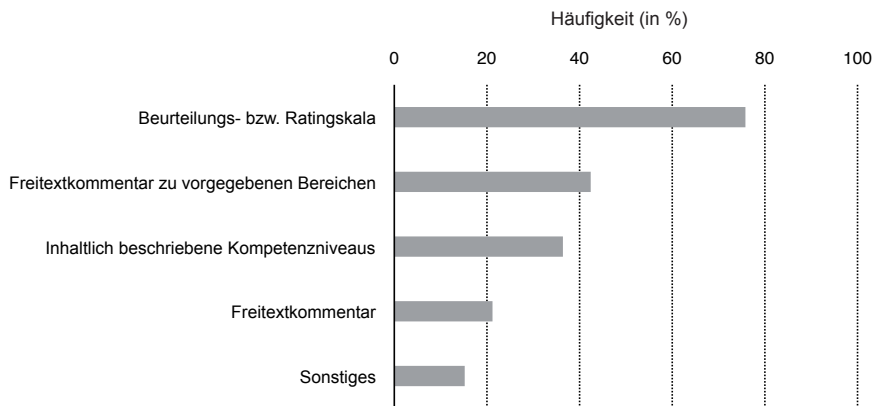


Abbildung 5: Häufigkeit der verschiedenen Beurteilungsformen (Mehrfachwahl; Gesamtstichprobe: $N = 33$; unter «Sonstiges» fällt z.B. der Hinweis, dass Beurteilungskriterien zugrunde lägen, diese aber nicht Teil des Notenblatts seien).

wird versucht, auf vielfältige Weise ein meist breites Kompetenzspektrum zu erfassen, z.B. in Form von Live- oder videobasierter Unterrichtsbeobachtung, Portfoliobegutachtung, mündlichen Prüfungen oder schriftlichen Prüfungen. Die Bewertung erfolgt grösstenteils durch mehrere Fachpersonen anhand verbindlicher und transparenter Beurteilungskriterien, welche sich überwiegend an wissenschaftlichen Theorien orientieren. Im Folgenden werden die beiden am Ende von Abschnitt 1 gestellten Fragen beantwortet, wobei zunächst der Einsatz neuerer, forschungsbasierter Methoden der Kompetenzdiagnostik betrachtet wird und anschliessend die Erfüllung der Gütekriterien durch die eingesetzten Beurteilungsverfahren im Zentrum steht.

5.1 Einsatz neuerer, forschungsbasierter Methoden

Die eingesetzten Erhebungsmethoden decken gemäss der vorliegenden Studie ein breiteres Spektrum ab als bei Strietholt und Terhart (2009), die eine gewisse Zurückhaltung bei der Anwendung neuer Methoden feststellten. So werden in der Deutschschweiz gegenwärtig teilweise u.a. Videos und E-Portfolios genutzt. Dennoch belegen die traditionellen Elemente – schriftliche Lektionsplanung, Live-Unterrichtsbeobachtung und Reflexionsgespräch – trotz methodischer Bedenken nach wie vor die Spitzenplätze. Die in der Forschung derzeit sehr verbreiteten Vignettentests und die Analyse fremder Unterrichtsvideos finden sich in der vorliegenden Studie nicht. Auch standardisierte Beurteilungssysteme mit Ratings durch externe Fachpersonen oder Assessment Center scheint es hierzulande nicht zu geben.

5.2 Erfüllung der Gütekriterien

Die Prüfungsverfahren scheinen so angelegt zu sein, dass sie mehrere *Funktionen* erfüllen können und somit mehr Funktionen als die von Strietholt und Terhart (2009)

untersuchten Verfahren, welche gemäss den Angaben der Auszubildenden in je etwa 60% der Fälle der Selektion und der Personalentwicklung und in nur rund 12% der Fälle der Orientierung dienen. Das Erfüllen multipler Funktionen ist jedoch ambivalent zu beurteilen. Einerseits spricht es für einen hohen Nutzen, andererseits kann die in einigen Studiengängen stattfindende Vermischung formativer und summativer Beurteilungsfunktionen zu Ziel- und Rollenkonflikten führen (vgl. Müller, 2011), welche die Qualität der Beurteilung beeinträchtigen können (vgl. Luhr, 2016).

Zum *Beurteilungsgegenstand* lässt sich resümieren, dass tendenziell versucht wird, ein recht breites Spektrum an Kompetenzen zu erfassen. Dies wird der Breite der von verschiedenen Seiten formulierten Standards eher gerecht als die stark auf das Unterrichten fokussierten Verfahren in der Studie von Strietholt und Terhart (2009) und ist somit mit Blick auf die Validität der Beurteilung positiv zu bewerten. Die vergleichsweise geringe Berücksichtigung des Bereichs «Innovieren» gemäss den KMK-Standards (KMK, 2004) kann jedoch angesichts der Bedeutung, die entsprechenden Kompetenzen für die Zukunft des Lehrberufs zugesprochen wird (z.B. um gesichertes Wissen in der Schule zur Entfaltung zu bringen, vgl. z.B. Gröschner, 2013), kritisch gesehen werden. Ob dieser Bereich auch in der Ausbildung weniger berücksichtigt wird, wäre eine interessante Frage für eine künftige Studie.

Mit einer Ausnahme liegen den Angaben der Verantwortlichen zufolge für alle Studiengänge verbindliche *Beurteilungskriterien* vor, welche immer allen Beteiligten zugänglich sind und meist auf wissenschaftlichen Konzepten beruhen. Dies dürfte für die Objektivität, die Transparenz und die Validität der Verfahren günstig sein. Damit heben sich die Resultate positiv von einigen Studien aus Deutschland ab. Laut Luhr (2016) beklagten Referendarinnen und Referendare häufig fehlende Beurteilungskriterien, und in einer Studie von Döbrich und Abs (2008) gaben 30% der Auszubildenden und 60% der Mentorierenden an, dass die Kriterien zur Bewertung von Referendarinnen und Referendaren (eher) uneinheitlich seien. Bei Strietholt und Terhart (2009) war allerdings selbst beim Vorliegen von Kriterien häufig unklar, wie stringent diese bei der Beurteilung berücksichtigt wurden. Bedenklich erscheint in der vorliegenden Studie hingegen, dass ein Beurteilungsinstrument rein intuitiv bzw. erfahrungsbasiert entwickelt wurde und den Verantwortlichen in drei Fällen nicht bekannt ist, wie die Kriterien zustande kamen bzw. worauf sie basieren.

Zur *Kompetenzerfassung* scheinen in der Deutschschweiz vielfältige Verfahren zum Einsatz zu kommen, die die Gütekriterien in sehr unterschiedlichem Masse erfüllen. Da, wie bereits dargestellt, alle Verfahren Vor- und Nachteile aufweisen, ist eine Bewertung hier schwierig. Mit einer Kombination verschiedener Methoden und der Integration der Informationen aus verschiedenen Quellen könnten möglicherweise die jeweiligen Vorteile genutzt und die Nachteile kompensiert werden (vgl. z.B. Busse & Kaiser, 2015). Dies wird in einigen Studiengängen z.B. in Form von Video-Portfolios versucht (vgl. Bäuerlein & Fraefel, 2016).

Die erhobenen Kompetenzen werden in der Regel durch mehrere *Beurteilende* mit unterschiedlichem fachlichem Hintergrund begutachtet. Dies ist angesichts des Forschungsstands und mit Blick auf die Objektivität und die Reliabilität der Urteile günstig (z.B. Hill et al., 2012; Neuenschwander, 2004). Auch die in zwei Dritteln der Studiengänge mit mehreren Prüfenden durchgeführten diskursiven Einigungsgespräche dürften im Sinne der Intersubjektivität positiv zu bewerten sein (vgl. Bovet & Frommer, 2015). Vor dem Hintergrund empirischer Befunde ist es zudem zu begrüßen, dass in allen Studiengängen Schulungen stattfinden (z.B. Hill et al., 2012). Vereinzelt wurde allerdings erwähnt, dass die Urteile trotz der Schulung häufig voneinander abweichen würden. Dieses Problem ist auch aus anderen Studien zur Unterrichtsbeurteilung bekannt (Bell, Jones, Qi & Lewis, 2018). Zur Frage, inwieweit sich Studierende und Prüfende bereits aus der Ausbildung kennen, ergeben die Daten kein einheitliches Bild. Da beides Vor- und Nachteile haben kann (z.B. Objektivität und Rollenklarheit vs. Verfügbarkeit von Kontextinformation), kann man diesen Befund kontrovers beurteilen (vgl. Praetorius, 2013). Einheitliche Regelungen erscheinen jedoch mit Blick auf die Fairness sinnvoll.

Bei der *Beurteilungsform* zeigt sich, dass in etwa der Hälfte der Fälle qualitative und quantitative Beurteilungen kombiniert werden, was im Vergleich dazu in der Studie von Strietholt und Terhart (2009) nur bei weniger als 4% der Verfahren der Fall war. Die Kombination ermöglicht die Nutzung der Vorteile beider Methoden. Dass zudem in der Deutschschweiz den Angaben zufolge in allen Studiengängen ein einheitliches Prozedere definiert ist, ist mit Blick auf Standardisierung und Fairness günstig. Im Gegensatz dazu blieb in der Studie von Strietholt und Terhart (2009) bei mehr als 50% der Verfahren unklar, wie genau die Beurteilung der angehenden Lehrpersonen erfolgte.

5.3 Limitationen

Bei den hier analysierten Daten handelt es sich um Selbstausskünfte der verantwortlichen Personen, welche im Sinne sozialer Erwünschtheit verfälscht sein können. Darüber hinaus erlauben die differenten Strukturen der Studiengänge und Hochschulen kaum Aussagen zur Repräsentativität. Zudem beziehen sich die Einschätzungen mit Blick auf die Gütekriterien auf die Gesamtheit der derzeit in der Deutschschweiz eingesetzten Beurteilungsverfahren. Einzelne Verfahren können selbstverständlich davon abweichen. Des Weiteren ist mit Blick auf den Vergleich der Beurteilungsgegenstände mit den KMK-Standards (KMK, 2004) zu berücksichtigen, dass in der vorliegenden Studie ein Kompetenzbereich als abgedeckt gilt, wenn mindestens ein Aspekt daraus geprüft wird. Somit wird die Berücksichtigung der Standards tendenziell überschätzt.

Hinsichtlich der mehrfach zum Vergleich herangezogenen Studie von Strietholt und Terhart (2009) ist zu beachten, dass sie sich auf die zweite Phase der deutschen Lehrpersonenausbildung bezieht und somit auf ein vom in der Deutschschweiz vorherrschenden einphasigen Modell abweichendes Ausbildungssystem. Zudem wurde sie vor zehn Jahren durchgeführt und untersuchte nicht nur summative Beurteilungen.

6 Fazit

Strietholt und Terhart (2009) analysierten in den Nullerjahren Verfahren, die in Deutschland zur Beurteilung angehender Lehrpersonen im Referendariat eingesetzt wurden, und kamen zu ernüchternden Ergebnissen. Es gibt Hinweise darauf, dass sich die Situation in der Deutschschweiz zum damaligen Zeitpunkt ähnlich darstellte (Fraefel & Huber, 2008). Die vorliegende Studie zeigt jedoch, dass sich hierzulande deutliche Fortschritte in der Qualität der Überprüfung der berufspraktischen Kompetenzen angehender Lehrpersonen abzeichnen.

Zahlreiche Fragen, deren Beantwortung zur weiteren Optimierung der Verfahren beitragen könnte, sind allerdings noch offen. So ist die Dimensionalität berufspraktischer Kompetenzen von Lehrpersonen nicht hinreichend geklärt, und Erhebungsverfahren und Messinstrumente sind nicht weit genug entwickelt und evaluiert (vgl. Neuweg, 2015). Eine zuverlässige Prüfung der berufspraktischen Kompetenzen, die stringent auf evidenzbasierten Kriterien und bewährten Erhebungsverfahren beruht, ist daher zum jetzigen Zeitpunkt nicht möglich – sofern sie überhaupt angestrebt wird. Die markante Zunahme der Forschungsbemühungen in diesem Bereich ist deshalb ebenso zu begrüßen wie die für viele Studiengänge offensichtlichen Anstrengungen, unter den gegebenen Bedingungen eine möglichst faire und aussagekräftige Beurteilung berufspraktischer Kompetenzen zu gewährleisten.

Literatur

- Admiraal, W. & Berry, A.** (2016). Video narratives to assess student teachers' competence as new teachers. *Teachers and Teaching*, 22 (1), 21–34.
- Bäuerlein, K. & Fraefel, U.** (2016). Ein Video-Portfolio zur summativen Überprüfung berufspraktischer Kompetenzen. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 34 (2), 212–218.
- Baumert, J. & Kunter, M.** (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), 469–520.
- Bell, C. A., Jones, N. D., Qi, Y. & Lewis, J. M.** (2018). Strategies for assessing classroom teaching. *Educational Assessment*, 23 (4), 229–249.
- Bovet, G. & Frommer, H.** (2015). *Praxis Lehrerberatung – Lehrerbeurteilung*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Busse, A. & Kaiser, G.** (2015). Wissen und Fähigkeiten in Fachdidaktik und Pädagogik. *Zeitschrift für Pädagogik*, 61 (3), 328–344.
- den Ouden, H.** (2019). Praktische Prüfungen. In S. Frölich-Steffen, H. den Ouden & U. Gießmann (Hrsg.), *Kompetenzorientiert prüfen und bewerten an Universitäten* (S. 153–192). Opladen: Barbara Budrich.
- Döbrich, P. & Abs, H. J.** (2008). *Evaluation der zweiten Phase der Lehrerbildung*. Frankfurt: DIPF.
- EDK.** (2019). *Rechtssammlung der EDK*. Bern: Schweizerische Konferenz der kantonalen Erziehungsdirektoren. Verfügbar unter: <http://www.edk.ch/dyn/11670.php> (21.02.2020).
- Feder, L. & Cramer, C.** (2018). Potenziale von Portfolioarbeit in der Lehrerinnen- und Lehrerbildung. *Die Deutsche Schule*, 110 (4), 354–367.
- Forster-Heinzer, S. & Oser, F.** (2015). Wer setzt das Maß? *Zeitschrift für Pädagogik*, 61 (3), 361–375.
- Fraefel, U.** (2012). Berufspraktische Studien und Schulpraktika. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 30 (2), 127–152.

- Fraefel, U., Bäuerlein, K. & Barabach, A.** (2018). Assessing teacher candidates' professional competence for evaluating teacher education programs. In V. C. X. Wang (Hrsg.), *Handbook of research on program development and assessment methodologies in K-20 education* (S. 418–442). Hershey, PA: IGI Global.
- Fraefel, U. & Huber, E.** (2008). Berufspraktische Kompetenzen zuverlässig prüfen? *Journal für Lehrerinnen- und Lehrerbildung*, 8 (1), 51–56.
- Frey, A.** (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften *Zeitschrift für Pädagogik*, 51. Beiheft, 30–46.
- Frey, A.** (2014). Kompetenzmodelle und Standards in der Lehrerbildung und im Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarbeitete Auflage) (S. 712–744). Münster: Waxmann.
- Gröschner, A.** (2013). Innovationskompetenz als Element der Lehrerausbildung. In M. Bolten & I. Bormann (Hrsg.), *Innovationen im Bildungswesen* (S. 303–327). Wiesbaden: Springer.
- Helmke, A.** (2015). *Unterrichtsqualität und Lehrerprofessionalität* (6. Auflage). Seelze: Klett Kallmeyer.
- Hill, H. C., Charalambous, C. Y. & Kraft, M. A.** (2012). When rater reliability is not enough. *Educational Researcher*, 41 (2), 56–64.
- Hohenstein, F., Kleickmann, T., Zimmermann, F., Köller, O. & Möller, J.** (2017). Erfassung von pädagogischem und psychologischem Wissen in der Lehramtsausbildung. *Zeitschrift für Pädagogik*, 63 (1), 91–113.
- Hutt, E. L., Gottlieb, J. & Cohen, J. J.** (2018). Diffusion in a vacuum: edTPA, legitimacy, and the rhetoric of teacher professionalization. *Teaching and Teacher Education*, 69, 52–61.
- InTASC.** (2013). *InTasc Model Core Teaching Standards and Learning Progressions for Teachers 1.0*. Washington: CCSSO.
- Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R. & Stigler, J. W.** (2012). Measuring usable knowledge. *American Educational Research Journal*, 49 (3), 568–589.
- Kleinknecht, M. & Poschinski, N.** (2014). Eigene und fremde Videos in der Lehrerfortbildung. *Zeitschrift für Pädagogik*, 60 (3), 471–490.
- KMK.** (2004). *Standards für die Lehrerbildung*. Bonn: Sekretariat KMK.
- König, J.** (2015). Kontextualisierte Erfassung von Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 61 (3), 305–309.
- Kopf, M., Leipold, J. & Seidl, T.** (2010). *Kompetenzen in Lehrveranstaltungen und Prüfungen*. Mainz: Johannes Gutenberg-Universität.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M.** (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften*. Münster: Waxmann.
- LCH.** (2014). *Der Berufsauftrag der Lehrerinnen und Lehrer*. Zürich: Dachverband Lehrerinnen und Lehrer Schweiz.
- Leonhard, T., Fraefel, U., Jünger, S., Kosinár, J., Reintjes, C. & Richiger, B.** (2016). Zwischen Wissenschafts- und Berufspraxis. *Zeitschrift für Hochschulentwicklung*, 11 (1), 79–98.
- Lotz, M., Gabriel, K. & Lipowksy, F.** (2013). Niedrig und hoch inferente Verfahren der Unterrichtsbeobachtung. *Zeitschrift für Pädagogik*, 59 (3), 357–380.
- Luhr, D.** (2016). *Unterrichtskompetenz erfassen und entwickeln*. Münster: Westfälische Wilhelms-Universität.
- Messner, H.** (2007). Unterrichtsbeurteilung in der berufspraktischen Ausbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 25 (1), 27–36.
- Metzger, C. & Nüesch, C.** (2004). *Fair prüfen*. St. Gallen: IWP.
- Meyer, H.** (2018). *Was ist guter Unterricht?* (13. Auflage). Berlin: Cornelsen.
- Müller, A.** (2011). Prüfungen als Lernchance. *Zeitschrift für Hochschulentwicklung*, 6 (1), 72–94.
- Neuenschwander, M. P.** (2004). Lehrerkompetenzen und ihre Beurteilung. *Journal für Lehrerinnen- und Lehrerbildung*, 4 (1), 23–29.
- Neuweg, G. H.** (2015). Kontextualisierte Kompetenzmessung. *Zeitschrift für Pädagogik*, 61 (3), 377–383.
- Oser, F., Curcio, G.-P. & Düggele, A.** (2007). Kompetenzmessung in der Lehrerbildung als Notwendigkeit. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 25 (1), 14–26.

- PH Zürich.** (Hrsg.). (2017). *Kompetenzstrukturmodell*. Zürich: Pädagogische Hochschule Zürich.
- Pianta, R., Hamre, B. & Allen, J.** (2012). Teacher-student relationships and engagement. In S. L. Christenson, A. L. Reschly & C. Wylie (Hrsg.), *Handbook of research on student engagement* (S. 365–386). New York: Springer.
- Plöger, W., Scholl, D. & Seifert, A.** (2016). «Und sie bewegt sich doch!» *Zeitschrift für Pädagogik*, 62 (1), 109–130.
- Porter, A. C. & Brophy, J. E.** (1988). Synthesis of research on good teaching. *Educational Leadership*, 45 (8), 74–85.
- Porter, A. C., Youngs, P. & Odden, A.** (2001). Advances in teacher assessment and their uses. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 259–297). Washington: AERA.
- Praetorius, A.-K.** (2013). Einschätzung von Unterrichtsqualität durch externe Beobachterinnen und Beobachter. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 31 (2), 174–185.
- Praetorius, A.-K. & Charalambous, C. Y.** (2018). Classroom observation frameworks for studying instructional quality. *ZDM*, 50 (3), 535–553.
- Praetorius, A.-K., McIntyre, N. A. & Klassen, R. M.** (2017). Reactivity effects in video-based classroom research. *Zeitschrift für Erziehungswissenschaft*, 20 (1), Supplement, 49–74.
- Praetorius, A.-K., Pauli, C., Rakoczy, K. & Klieme, E.** (2014). One lesson is all you need? *Learning and Instruction*, 31, 2–12.
- Praetorius, A.-K., Vieluf, S., Saß, S., Bernholt, A. & Klieme, E.** (2016). The same in German as in English? *Zeitschrift für Erziehungswissenschaft*, 19 (1), 191–209.
- Roloff, S.** (2012a). *Mündliche Prüfungen*. Offenburg: Hochschule Offenburg.
- Roloff, S.** (2012b). *Schriftliche Prüfungen stellen und auswerten – methodisch, effektiv, objektiv*. Offenburg: Hochschule Offenburg.
- Rothland, M., Cramer, C. & Terhart, E.** (2018). Forschung zum Lehrerberuf und zur Lehrerbildung. In R. Tippelt & B. Schmidt-Hertha (Hrsg.), *Handbuch Bildungsforschung* (S. 1011–1034). Wiesbaden: Springer.
- Rutsch, J., Rehm, M., Vogel, M., Seidenfuß, M. & Dörfler, T.** (Hrsg.). (2018). *Effektive Kompetenzdiagnose in der Lehrerbildung*. Wiesbaden: Springer.
- Strietholt, R. & Terhart, E.** (2009). Referendare beurteilen. *Zeitschrift für Pädagogik*, 55 (4), 622–645.
- Terhart, E.** (2015). Lehrerbildung als Gegenstand empirischer Forschung. In S. Gehrman, J. Helmchen, M. Krüger-Potratz & F. Ragutt (Hrsg.), *Bildungskonzepte und Lehrerbildung in europäischer Perspektive* (S. 17–35). Münster: Waxmann.
- von Aufschnaiter, C., Cappell, J., Dübbelde, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßler, R. & Wolgast, A.** (2015). Diagnostische Kompetenz. *Zeitschrift für Pädagogik*, 61 (5), 738–758.
- Walzik, S.** (2012). *Kompetenzorientiert prüfen*. Opladen: Barbara Budrich.
- Wei, R. C. & Pecheone, R. L.** (2010). Assessment for learning in preservice teacher education. In M. Kennedy (Hrsg.), *Teacher assessment and the quest for teacher quality* (S. 69–132). San Francisco: Jossey-Bass.

Autorin und Autoren

Kerstin Bäuerlein, Dr., Pädagogische Hochschule FHNW, kerstin.baeuerlein@fhnw.ch

Sascha Senn, M.A., Pädagogische Hochschule FHNW, sascha.senn@fhnw.ch

Urban Fraefel, Prof., Pädagogische Hochschule FHNW, urban.fraefel@fhnw.ch