



Entwicklung und Validierung der „Mannheimer Skalen zur Erfassung von Handlungswissen, Reflexion und Berufsbezug (HAREBE)“ für situierte Lehrveranstaltungen an Hochschulen

Anne-Sophie Waag , Benedict C. O. Fehringer und Stefan Münzer

Bildungspsychologie, Universität Mannheim

Zusammenfassung: Situierte Formate der Hochschullehre stellen authentische Kontexte bereit, in denen Studierende praktische Erfahrung sammeln, Handlungsoptionen abwägen, komplexe Problemkonstellationen wissenschaftlich analysieren und Aspekte der späteren beruflichen Tätigkeit kennenlernen können. Bislang fehlen geeignete Lehrevaluationsinstrumente für solche Lehrformate. Die vorliegende Arbeit beschreibt die Entwicklung und Validierung sechs ökonomisch einsetzbarer Skalen für die Einschätzung von Handlungswissen, Reflexion und Berufsbezug in der Hochschullehre. Drei Skalen erfassen dabei die Selbsteinschätzung von studentischen Kompetenzen. Drei weitere, parallel entwickelte Skalen erfassen wahrgenommene Merkmale des Lernangebots. Die Skalen wurden theoriebasiert entwickelt und auf Basis empirischer Datenerhebungen überarbeitet sowie inhaltlich, faktoriell und kriterial validiert. Die internen Konsistenzen sind gut. Die postulierte Modellstruktur wird durch eine konfirmatorische Faktorenanalyse bestätigt. Die Kriteriumsvalidität wird mittels verschiedener personenbezogener Merkmale und durch konstruktnahe Skalen überprüft.

Schlüsselwörter: Situiertes Lernen, Lehrveranstaltungsevaluation, Handlungswissen, Reflexion, Berufsbezug

Development and Validation of the “Mannheim Scales for the Acquisition of Action Knowledge, Reflection and Occupational Relevance” (HAREBE) for Situated Courses at Universities

Abstract: Situated teaching formats in higher education provide an authentic setting in which students can gain practical experience. They learn to weigh different options for action, to scientifically analyze complex problem constellations, and to become acquainted with aspects of later professional activities. To date, however, there was a lack of suitable evaluation instruments for such teaching formats. The present paper describes the development and validation of six economically applicable scales for the assessment of action knowledge, reflection, and occupational relevance in higher education. Three scales address the self-assessment of student competencies regarding these constructs; three further scales address the perceived characteristics of the given learning opportunities. The scales were developed and revised using several data collections. A cognitive interview was applied to examine the content validity of the scales. The factorial and criterion validity was examined as well. The scales show good reliability. Confirmatory factor analysis supports the postulated model structure, and various personal characteristics and construct-related scales verified the criterion validity.

Keywords: situated learning, course evaluation, action knowledge, reflection, occupational relevance

Der Wissenschaftsrat hat die Berufsbefähigung und Persönlichkeitsentwicklung von Studierenden als zentrale Ziele kompetenzorientierter Hochschulbildung hervorgehoben (Wissenschaftsrat, 2008). Studierende sollen fachliche und überfachliche Kompetenzen erwerben, um professionell auf herausfordernde berufliche Anforderungssituationen reagieren zu können (Schindler, 2004; Schubarth et al., 2012).

Um die Anwendung wissenschaftlicher Erkenntnisse und Methoden auf praxisnahe Probleme zu fördern, wurden seit der Bologna-Reform in der universitären Lehre eine Reihe von Formaten entwickelt, die über den tradierten Rahmen hinausgehen (Wildt, 2013). Dazu zählen u. a. forschendes und problembasiertes Lernen sowie Service Learning. Diese Lernformate sind insofern *situiert*, als dass Studierende in einem berufsnahen, authentischen

Anwendungskontext echte Handlungsoptionen abwägen und Handlungsentscheidungen treffen können (Rank, Gebauer, Hartinger & Fölling-Albers, 2012). Der wissenschaftliche Anspruch manifestiert sich in gezielt geschaffenen Gelegenheiten zur wissenschaftlichen Analyse und Reflexion realer Probleme in der universitären Lehre.

Die vorliegende Arbeit leistet einen Beitrag zur Diagnose von Handlungswissen (HA), Reflexion (RE) und Berufsbezug (BE) als typische und spezifische Facetten situierter Lehr-Lernformate. Konkrete Handlungsweisen können erlernt und ausprobiert werden (HA), praktische Erfahrungen werden vor dem Hintergrund theoretischen Wissens und wissenschaftlicher Analysemethoden diskutiert (RE) und eine Auseinandersetzung mit Anforderungen möglicher Berufsfelder findet statt (BE).

Um die Ausprägung dieser Facetten in den jeweiligen situierten Lehrveranstaltungen diagnostizieren zu können, wurden sechs ökonomisch einsetzbare Skalen entwickelt, psychometrisch überprüft und validiert. Die HAREBE-Skalen liegen zum einen als Selbsteinschätzungsskalen zur Bestimmung studentischer Kompetenzen und zum anderen als Evaluationsskalen zur Bestimmung von hochschulischen Lerngelegenheiten vor. Mithilfe der sechs Skalen können situierte Lehrveranstaltungen ergänzend und ökonomisch evaluiert werden.

Ergebnis- und Gestaltungsmerkmale situierter Lehrveranstaltungen

Situierte Lehrveranstaltungen unterscheiden sich in ihrer Gestaltung und in den erwünschten und antizipierten Ergebnissen von klassischen Lehrveranstaltungen (Wildt, 2013). Bei der Gestaltung der Formate stehen studienorientierte Lerngelegenheiten im Vordergrund, welche im Ergebnis die fachliche und überfachliche Kompetenzentwicklung der Studierenden fördern sollen (Schaper, Reis, Wildt, Horvath & Bender, 2012). Als für situierte Lehrveranstaltungen charakteristische und erwünschte Lernergebnisse werden insbesondere gesteigerte Handlungs-, Reflexions-, aber auch Berufswahlkompetenzen genannt (Reinders, 2016). Darunter verstehen wir im Kontext situierten Lernens Folgendes:

- *Handlungskompetenz* ist die Fähigkeit, „aus einem begrenzten Regelsystem heraus (Wissensbasis) eine unendliche Vielzahl von situationsadäquaten Handlungen generieren zu können“ (Schaper et al., 2012, S. 16). Das benötigte prozedurale Wissen umfasst verschiedene Praktiken, Techniken, Methoden und Strategien, die in spezifischen Situationen eingesetzt werden können (Kiel & Rost, 2002). Studierende mit einer ausgeprägten Handlungskompetenz fühlen sich in berufspraktischen, herausfordernden Situationen sicher, geeignete

Handlungstechniken einsetzen zu können bzw. gute Handlungsoptionen zu finden.

- *Reflexionskompetenz* beschreibt die Fähigkeit, theoretisches Wissen und praktische Erfahrungen miteinander in Bezug zu setzen (Reinders, 2016; Schön, 1991; Sliwka, 2009). Dieser Bezug besteht zum einen in der Übersetzung theoretischer Modelle in situativ angemessene Handlungspraktiken und zum anderen in der kritischen Betrachtung von Handlungen und Handlungsfolgen. Dazu gehören die theoriegeleiteten Bewertungen der Handlungen und ihrer Folgen sowie die Begründung von besseren Handlungsalternativen.
- *Berufswahlkompetenz* beschreibt die Fähigkeit, eine Berufsentscheidung mit möglichst hoher individueller Passung zu treffen. Dafür müssen die persönlichen Interessen und Wünsche in Bezug auf ein berufliches Tätigkeitsfeld erkannt und das eigene berufsrelevante Wissen und Können festgestellt werden (Butz, 2008). Ferner sollen „Möglichkeiten, Bedarfe und Anforderungen“ (Butz, 2008, S. 50) der interessierenden Berufsfelder erkundet werden. Ziel ist ein Abgleich beider Bereiche, um eine möglichst hohe Übereinstimmung zu finden.

Studien zeigen, dass Kompetenzentwicklungen in den genannten Bereichen keine zwangsläufige Folge situierter Lehrveranstaltungen sind, sondern eine explizite Förderung in der Lehrveranstaltung benötigen (Reinders, 2016; Wagner, Klein, Klopp & Stark, 2014). Spezifische Lerngelegenheiten regen die gewünschten Entwicklungen an und unterstützen diese gezielt. Wir nehmen dementsprechend an, dass sich Lehrveranstaltungen durch jeweils niedrigere oder höhere Handlungs-, Reflexions- und Berufsorientierung auszeichnen können. Darunter verstehen wir im Kontext situierter Lehre Folgendes:

- *Handlungsorientierung* entsteht, wenn Studierende in authentischen Kontexten Gelegenheit erhalten, Handlungsentscheidungen abzuwägen und umzusetzen. Lerngelegenheiten mit Handlungsbezug können auch in konventionellen Lehrveranstaltungen durch Fallbeispiele, Simulationen und Rollenspiele erfolgen. In situierteren Lehrveranstaltungen ist jedoch typischerweise ein realer und authentischer Kontext gegeben, in dem die Handlung ausgeführt und an z.T. komplexe und nicht vollständig kontrollierbare Konstellationen von Einflüssen angepasst wird.
- *Reflexionsorientierung* zeigt sich in Lerngelegenheiten, die einen expliziten Bezug zwischen i.d.R. herausfordernden, praktischen Erfahrungen und passendem theoretischem Wissen herstellen. Entscheidend für die Förderung der Reflexionskompetenz ist das Eingrenzen und Analysieren der Praxissituation mithilfe theoretischer Modelle und wissenschaftlicher Methoden so-

wie die theoriegeleitete Begründung der abgeleiteten Handlungsoptionen.

- *Berufsorientierung* kann sowohl durch informative Lehrbausteine als auch durch praktische Tätigkeiten ermöglicht werden. Die Berufsorientierung einer Lehrveranstaltung zeigt sich u. a. darin, dass Studierende Gelegenheit erhalten, sich mit konkreten berufsbezogenen Aspekten zu befassen und ihre Eindrücke in der Lehrveranstaltung thematisieren zu können. Die Lerngelegenheiten regen Studierende zur Auseinandersetzung mit entsprechenden Passungsüberlegungen an.

Erfassung von überfachlichen Kompetenzfacetten und Gestaltungsmerkmalen situierter Lehrveranstaltungen

Zur Untersuchung überfachlicher Kompetenzen hat sich die Erfassung mittels Selbstbeurteilungsverfahren bewährt (Braun, Gusy, Leidner & Hannover, 2008; Klieme, Artelt & Stanat, 2002). Eines der ersten deutschsprachigen, kompetenzorientierten Selbstbeurteilungsverfahren ist das *Berliner Evaluationsinstrument für selbsteingeschätzte studentische Kompetenzen* (BEvaKomp; Braun et al., 2008). Es erfasst verschiedene fachübergreifende Kompetenzbereiche, adressiert jedoch nicht die für situierte Lehrveranstaltungen relevanten Kompetenzfacetten.

Die Sichtung spezifischerer Skalen zeigte weitere Einschränkungen. Zur Erfassung von Handlungswissen liegen fachspezifische Tests aus der Medizin (*Stroke Action Test*, STAT; Roebbers, Razum, Kutschmann & Wagner, 2008) oder dem Lehramt (*Pädagogisches Unterrichtswissen*, PUW; König & Blömeke, 2010) vor. Zur Erfassung der Theorie-Praxis-Reflexion sind insbesondere Skalen für das Lehramt (*Theorie-Praxis-Reflexion*; Reinders, Hillesheim & Sebald, 2015, *Reflexion über Probleme im Unterricht*; Linninger, 2016) vorhanden. Für die Berufswahlkompetenz ist der *Fragebogen Berufswahlkompetenz* (FBK; Lipowski, Kaak, Kracke & Holstein, 2015) hervorzuheben. Die Einschränkungen beziehen sich auf folgende Punkte: (1) Es werden eher persönliche Überzeugungen und Einstellungen als Kompetenzen erfasst (z. B. Skala Theorie-Praxis-Reflexion), (2) Skalen sind domänenspezifisch formuliert und daher nicht in anderen Fächern einsetzbar (z. B. STAT oder PUW), und (3) die Anzahl der Items ist hoch (z. B. FBK).

Gestaltungsmerkmale von Lehrveranstaltungen werden i. d. R. durch die klassischen Lehrveranstaltungsevaluationen erfasst. Zu den bekannteren deutschsprachigen Instrumenten gehören das *Trierer Inventar zur Lehrveranstaltung* (TRIL; Gollwitzer & Schlotz, 2003), das *Heidelberger Inventar zur Lehrveranstaltungsevaluation* (HILVE; Rindermann & Amelang, 1994) sowie die *Internet-*

basierte studentische Evaluation von Lehrveranstaltungen – Kurzform (InstEvaL-K; Erdfelder, Pesch & Stahl, 2004). Diese Instrumente erfassen überwiegend allgemeine Gestaltungsmerkmale und Input-Faktoren wie die *Struktur der Lehrveranstaltung* oder *Merkmale der Dozierenden* (Zumbach, Spinath, Schahn, Friedrich & Kögel, 2007), die in klassischen Seminaren und Vorlesungen bedeutsam sind. Sie fragen aber nicht danach, ob in der Lehrveranstaltung die Entwicklung von Handlungskompetenz unterstützt, Theorie-Praxis-Reflexion gefördert oder Berufsbezug hergestellt wurde. Daher können situierte Lehrveranstaltungen mit diesen Instrumenten allein nicht angemessen evaluiert werden.

Die Skalen zur Erfassung von Handlungswissen, Reflexion und Berufsbezug

Im Folgenden wird die Entwicklung, Überprüfung und Validierung von sechs Skalen beschrieben, die überfachlich und ökonomisch für hochschulische Lehrveranstaltungsevaluationen einsetzbar sind. Dabei thematisieren drei Skalen die studentischen Handlungs-, Reflexions- und Berufswahlkompetenzen (HA-KOMP, RE-KOMP, BE-KOMP) als überfachliche Selbsteinschätzungen und drei weitere Skalen die Handlungs-, Reflexions- und Berufsorientierung (HA-LERN, RE-LERN, BE-LERN) als Gestaltungsmerkmale situierter Lehrveranstaltungen. Mithilfe der Skalen können situierte Lehrveranstaltungen an Hochschulen weiterentwickelt und standardisierte Evaluationen ergänzt werden.

Validierungskriterien und Annahmen

Die Validierung ist an das Rahmenkonzept zur Validität der *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association & National Council on Measurement in Education, 2014) angelehnt. Die Standards stellen Informationen und Kriterien bereit, um die Qualität psychologischer Testverfahren zu gewährleisten und zu beurteilen (Plake & Wise, 2014). Für die Beurteilung der Validität wurden Aspekte (1) der internen Struktur, (2) der Beziehung von Skalenkonstrukten zu personenspezifischen Merkmalen sowie (3) der Beziehung von Skalenkonstrukten mit konstruktnahen und -fernen Variablen herangezogen (Standards 1.13 bis 1.18; AERA et al., 2014). Darüber hinaus wurden Prätest-Posttest-Unterschiede (vor/nach einer situierten Lehrveranstaltung) als

Hinweis auf Veränderung durch gezielte Lernerfahrungen berücksichtigt.

Interne Struktur. Die Analyse der internen Struktur eines Tests soll zeigen, inwiefern die theoretisch vermuteten Beziehungen zwischen Items und Faktoren mit jenen in den Daten gefundenen übereinstimmen (AERA et al., 2014). In den Standards 1.13 bis 1.15 (*Evidence Regarding Internal Structure*) werden Aussagen dazu gefordert, welche Implikationen für Interpretation und Anwendung die interne Struktur bezogen auf das gesamte Ergebnis (1.13), Teilergebnisse (1.14) oder Einzelergebnisse (1.15) hat.

Für unsere sechs Skalen postulieren wir auf Grundlage der theoretischen Annahmen und des Entwicklungsprozesses eine hierarchische Modellstruktur: Wir nehmen an, dass Kompetenzeinschätzungen auf der einen Seite (HAREBE-KOMP) und Gestaltungsmerkmale des Lernangebots auf der anderen Seite (HAREBE-LERN) übergeordnete Konstrukte mit den jeweiligen untergeordneten Konstrukten HA-, RE-, BE-KOMP bzw. -LERN bilden. Wir gehen außerdem aufgrund der parallelen Itementwicklung davon aus, dass Kompetenz- und Lernangebotsskalen innerhalb jeden Bereichs korrelieren (z. B. HA-KOMP mit HA-LERN). Für die Anwendung können Summenwerte für einzelnen Skalen (z. B. HA-KOMP) und für die Gesamtskala (z. B. HAREBE-KOMP) gebildet werden. Außerdem können inhaltlich aufeinander bezogene Skalen (z. B. HA-KOMP und HA-LERN) zueinander in Beziehung gesetzt werden. Die Betrachtung von Ergebnissen auf Itemebene ist nicht vorgesehen.

Beziehungen zu personen- und gruppenspezifischen Merkmalen. Um die Interpretation der Skalenergebnisse zu prüfen, können die Konstrukte mit anderen Variablen in Beziehung gesetzt werden. Im Standard 1.16 (*Evidence Regarding Relationships With Conceptually Related Constructs*) werden theoretisch motivierte Begründungen für die Auswahl der Variablen begründet (AERA et al., 2014). In den folgenden Abschnitten stellen wir daher die Validierungsvariablen und deren erwartete Beziehung zu den betreffenden Konstrukten vor. Eine Übersicht aller herangezogenen Variablen, inklusive Angaben zu deren Reliabilität (Standard 1.17, AERA et al., 2014), findet sich im elektronischen Supplement ESM 1.

Das *deklarative Fachwissen* hängt positiv mit Handlungskompetenz zusammen, wofür es zunehmend empirische Evidenz bspw. aus der Lehrerbildungsforschung gibt (Baumert & Kunter, 2013; Voss, Kunina-Habenicht, Hoehne & Kunter, 2015). Auch bildet das fachlich-theoretische Wissen die Grundlage für erfolgreiche Reflexionsprozesse (Bosse, 2011; Goeze, Hetfleisch & Schrader, 2013). Mit zunehmendem Fachwissen steigt zudem die Kompetenz, die eigene Passung für spezifische Berufe beurteilen zu können (Lipowski, Kaak & Kracke, 2016; Porath, 2014). Basierend auf diesen Überlegungen wird

angenommen, dass höheres Fachwissen positiv mit allen drei Kompetenzbereichen zusammenhängt.

Das *Fachinteresse* hängt u. a. positiv mit der Informationsverarbeitung sowie der Leistung von Studierenden zusammen, wie die Motivationsforschung zeigt (Hidi & Harackiewicz, 2000). Ferner postulieren einschlägige Kompetenzmodelle (Blömeke, Zlatkin-Troitschanskaia, Kuhn & Fege, 2013) motivationale Orientierungen wie das Interesse als eine von drei Kompetenzdimensionen (Gronostay & Manzel, 2015). Es ist daher anzunehmen, dass das Fachinteresse einen positiven Einfluss auf die Kompetenzeinschätzung in allen drei Bereichen hat.

Studienrelevante Praktika können einen positiven Einfluss auf den (über-)fachlichen studentischen Kompetenzerwerb haben (Gröschner, Schmitt & Seidel, 2013; Häscher, 2006). Wir nehmen daher an, dass Studierende mit studienbezogenen Praktikumserfahrungen ihre Handlungs-, Reflexions- und Berufswahlkompetenzen höher einschätzen als Studierende ohne diese Erfahrungen, was dem Konzept der *known-groups* Validierung von Cronbach und Meehl (1955) entspricht.

Beziehungen zu konstrukt-nahen und -fernen Variablen. Zur konvergenten und divergenten Validierung werden konstrukt-nahe oder -ferne Variablen herangezogen (AERA et al., 2014). Unter Bezug auf den o.g. Standard 1.16 begründen wir die erwarteten Beziehungen zu den gewählten Variablen wie folgt:

Für die Skala Handlungskompetenz wird zur konvergenten Validierung die *Skala zur allgemeinen Selbstwirksamkeitserwartung* (SWE; Schwarzer & Jerusalem, 1999) herangezogen. Je mehr Studierende sich zutrauen, mit Herausforderungen umgehen zu können, desto höher sollten sie ihre Handlungskompetenz wahrnehmen.

Für die Skala Reflexionskompetenz werden die Skalen *Reflexionsfähigkeit* und *Theorie-Praxis-Reflexion* (Reinders et al., 2015) zur konvergenten Validierung herangezogen, da sie eine inhaltliche Passung aufweisen und bereits im Kontext situierter Lehrveranstaltungen erprobt wurden (Reinders, 2016). Da die Validierung in der vorliegenden Studie in der Lehrerbildung erfolgt, ist der Fachbezug dieser Skalen nicht problematisch. Ein bedeutsamer positiver Zusammenhang wird erwartet.

Für die Skala Berufswahlkompetenz wurden drei Skalen aus dem *Fragebogen Berufswahlkompetenz* (FBK; Lipowski, Kaak, Kracke & Hostein, 2015) zur konvergenten Validierung ausgewählt. Die Facetten Selbstwissen, Bedingungswissen und Planungskompetenz bilden relevante Inhalte für das hier einzugrenzende Konstrukt ab, die sich auch in unserer Konstruktdefinition finden. Ein bedeutsamer positiver Zusammenhang wird erwartet.

Für die Skala Berufsorientierung wurde zur konvergenten Validierung die Skala *Berufstransfer* (Reinders et al., 2015) herangezogen. Sie erfasst das wahrgenommene

Verhältnis der Seminarinhalte zum späteren Berufsleben. Ein bedeutsamer positiver Zusammenhang wird erwartet.

Wir nehmen außerdem an, dass die Skalen zur Erfassung der studentischen Wahrnehmung von Lerngelegenheiten in Bezug auf Handlungs-, Reflexions- und Berufsorientierung in einer Lehrveranstaltung auch in Beziehung mit weiteren, klassischen Lehrveranstaltungsmerkmalen stehen. Hierzu wurden Items zur Bewertung der Lehrveranstaltungsstruktur sowie der Dozierendenmerkmale des *Instruments zur Evaluation von Lehrveranstaltungen – Kurzform* (InstEvaL-K, Erdfelder, Stahl & Pesch, 2004) herangezogen. Ein bedeutsamer positiver Zusammenhang mit den HAREBE-Lernangebotskalen wird erwartet.

Wir erwarten, dass der Besuch einer situierten Lehrveranstaltung einen Einfluss auf die Beurteilung der eigenen Handlungs-, Reflexions- und Berufswahlkompetenzen hat. Eine Variable, von der wir dies jedoch nicht im besonderen Maße annehmen, ist die Leistungsmotivation; hier erfasst mit der Subskala *Hoffnung auf Erfolg* (Lang & Fries, 2006). Überprüft wurde diese Annahme in einem Prä-Posttest-Design anhand der Wirkung auf die gemessenen Konstrukte durch die Teilnahme an einer situierten Lehrveranstaltung. Während für die HAREBE Kompetenzskalen ein signifikanter Zuwachs erwartet wird, gilt dies nicht für die Skala *Hoffnung auf Erfolg*. Die situierte Lehrveranstaltung sollte sich also spezifisch auf die HAREBE-Skalen auswirken, nicht aber auf die Motivationskala.

Itempool

Zur Entwicklung der je drei Kompetenz- und Lernangebotskalen wurden auf Basis der Konstruktdefinitionen Itempools generiert. Diese bestanden aus insgesamt 23 Items für die Kompetenz- und 27 Items für die Lernangebotskonstrukte. Diese 50 Items wurden anschließend an einer Stichprobe von 126 Studierenden (82 Frauen, 36 Männer, acht ohne Angabe), durchschnittlich 21 Jahre alt ($M = 21.01$, $SD = 2.27$) in einem Paper-Pencil-Verfahren erhoben. Ziel war eine Itemreduktion auf ökonomische fünf Items je Skala. Alle Items waren auf einer vierstufigen Rating-Skala (1 = *trifft nicht zu* bis 4 = *trifft sehr zu*) zu beantworten.

Die Itemreduktion erfolgte mithilfe konfirmatorischer Faktorenanalysen erster Ordnung jeweils für die drei Kompetenz- sowie die drei Lernangebotskalen bei paralleler Prüfung der internen Konsistenzen. Schrittweise wurden pro Skala jeweils die Items mit den niedrigsten Faktorladungen ausgeschlossen. Übrig blieben fünf Items pro Skala, die die Basis für die folgenden psychometrischen Untersuchungen darstellten.

Erste psychometrische Überprüfung, Inhaltsvalidierung und Überarbeitung

Methode

In einer weiteren Datenerhebung wurden die sechs Skalen einer ersten psychometrischen Prüfung unterzogen. Die Stichprobe bestand aus 246 Studierenden (149 Frauen, 73 Männer, 1 divers, 23 o.A.). Alle Studierenden wurden im Rahmen von universitären, situierten Lehrveranstaltungen in den Fächern Psychologie, Mathematik, BWL, Geschichte und Lehramt befragt. Das Durchschnittsalter betrug 22 Jahre ($M = 22.17$, $SD = 3.03$); sie besuchten im Mittel das 4. Fachsemester ($M = 4.11$, $SD = 2.91$, $Min = 1$, $Max = 12$). Die Datenerhebung erfolgte in der jeweils letzten Veranstaltungssitzung im Sommersemester 2018. Die Daten wurden sowohl analog als auch online (*SoSci Survey*; Leiner, 2019) erhoben. Mit einer Teilstichprobe aus 12 Studierenden der verschiedenen Fächer wurde außerdem ein *kognitives Interview* durchgeführt (für mehr Informationen siehe ESM 2).

Die Reliabilitätsanalysen wurden mit dem Statistikprogramm R (Version 4.0.2, RCore Team, 2014), der Entwicklungsumgebung RStudio (RStudio Team, 2019) und dem Softwarepaket psych (Revelle, 2020) durchgeführt. Die postulierte hierarchische Struktur der Fragebogendaten wurde mithilfe einer konfirmatorischen Faktorenanalyse in R und dem Softwarepaket lavaan (Rosseel, 2012) getestet. Es wurden folgende Spezifikationen für das zu untersuchende Modell getroffen: (1) Die Faktoren zweiter Ordnung HAREBE-KOMP und HAREBE-LERN sind orthogonal zueinander, (2) die jeweiligen Faktoren erster Ordnung (HA-, RE- und BE-KOMP sowie HA-, RE- und BE-LERN) sind untereinander korreliert, (3) die Faktoren erster Ordnung der gleichen Facette (HA-KOMP und -LERN, RE-KOMP und -LERN sowie BE-KOMP und -LERN) sind ebenfalls korreliert.

Da die manifesten Variablen nicht normalverteilt waren, wurde zur Anpassung des Modells der MLM Schätzer mit robusten Standardfehlern und Sartorra-Bentler skaliertes Statistik eingesetzt. Dieser kann nur bei vollständigen Datensätzen eingesetzt werden, weshalb die fehlenden Werte listenweise entfernt wurden (Rosseel et al., 2019). Um die Passung eines geschätzten Modells auf die empirischen Daten zu überprüfen, werden verschiedene Fit-Indices empfohlen (Hu & Bentler, 2009). Zur Überprüfung der einzelnen Modellfits wurden die Fit-Indices *Comparative Fit Index (CFI)*, *Root Mean Square Error of Approximation (RMSEA)* und *Standardized Root Mean Square Residual (SRMR)* herangezogen. Die noch akzeptablen Cut-off-Werte liegen je nach Quelle bei $CFI > .90$,

Tabelle 1. Kennung und Wortlaut von je einem Beispielitem der sechs HAREBE Skalen

Kennung	Wortlaut
HA_K1	In praxisbezogenen Situationen weiß ich was zu tun ist.
RE_K1	Nach schwierigen Situationen in der Praxis erarbeite ich mithilfe theoretischer Modelle neue Handlungsmöglichkeiten.
BE_K1	Ich weiß, welche Inhalte ich für einen möglichen späteren Beruf noch lernen muss.
HA_L1	Ich habe im Seminar Vorgehensweisen gelernt, die ich in der praktischen Tätigkeit direkt anwenden konnte.
RE_L1	Während der praktischen Tätigkeit habe ich bereits Situationen erlebt, in denen theoretisches Wissen aus dem Seminar hilfreich war.
BE_L1	Die im Seminar gelernten Inhalte stehen im direkten Bezug zu möglichen Berufsfeldern.

$RMSEA < .08$ und $SRMR < .09$ (Kline, 2005; Schermelleh-Engel, Moosbrugger & Müller, 2003).

Ergebnisse

Reliabilitätsschätzungen und Faktorstruktur. Die Reliabilitätsschätzungen zeigten interne Konsistenzen zwischen $a = .73$ und $a = .85$. Die konfirmatorische Faktorenanalyse zeigte eine akzeptable Modellpassung ($CFI = .929$, $RMSEA = .043$, $SRMR = .070$).

Aufgrund der Ergebnisse aus den kognitiven Interviews (siehe ESM 2) wurden die Items sprachlich überarbeitet, hinsichtlich semantischer Komplexität reduziert und ausdifferenziert. Die Antwortoptionen wurden von vier auf sechs erhöht und die Randpositionen in der Benennung weniger extrem beschrieben (*trifft zu* und *trifft nicht zu*). In Tabelle 1 ist je Skala ein Beispielitem der überarbeiteten Version mit Wortlaut aufgeführt (für alle Items siehe ESM 3).

Validierung der Faktorstruktur und Zusammenhänge mit ausgewählten Konstrukten

Methode

Es wurde eine neue Stichprobe aus 201 Studierenden (146 Frauen, 55 Männer) eines gymnasialen Lehramtsstudiengangs online (*SoSci Survey*; Leiner, 2019) erhoben. Das Durchschnittsalter betrug 21 Jahre ($M = 20.96$, $SD = 2.46$). Ein Großteil studierte im 2. Semester ($M = 2.62$, $SD = 2.14$). Circa 30 Prozent hatten bereits ein studienrelevantes Praktikum absolviert. Die Daten einer Teilstichprobe von 94 Studierenden wurden im Rahmen des situierten Begleitseminars zum schulischen Orientierungspraktikum im zweiten Semester Lehramt erhoben. Die Kompetenzskalen wurden hier als Prä- und Posttest erhoben, die Lernangebotsskalen als Posttest.

Verwendete Variablen zur Validierung. Die Variablen, welche zur Validierung der HAREBE Skalen herangezogen wurden, wurden oben im Abschnitt zu den Validierungsannahmen und Kriterien beschrieben und im ESM 2 eingehend vorgestellt.

Durchführung. Zunächst wurden die Reliabilitäten sowie die skaleninternen und -übergreifenden Korrelationsmuster überprüft. Anschließend wurde die postulierte hierarchische Struktur der Fragebogendaten mithilfe einer konfirmatorischen Faktorenanalyse getestet.

Aufgrund der hohen Interkorrelationen innerhalb der Lernangebotsskalen wurde zur weiteren Absicherung der postulierten Faktorstruktur ein Alternativmodell getestet. Auf der Seite der Kompetenzfaktoren blieb dieses unverändert; auf der Seite der Lernangebotsfaktoren wurde statt der drei Faktoren erster Ordnung (HA-, RE-, BE-LERN) und dem Gesamtfaktor zweiter Ordnung (HAREBE-LERN) lediglich ein Faktor HAREBE-LERN gebildet. Das Alternativmodell wurde so spezifiziert, dass die Faktoren HAREBE-KOMP und HAREBE-LERN weiterhin orthogonal blieben, die Kompetenzfaktoren erster Ordnung (HA-, RE-, BE-KOMP) aber jeweils mit dem Lernangebotsfaktor HAREBE-LERN korrelierten.

Die Validierungsskalen (siehe ESM 1) wurden z -standardisiert, um die verschiedenen Ratingskalierungen zu berücksichtigen. Wie im Standard 1.18 (AERA et al., 2014) empfohlen, wurden zur Untersuchung der Beziehung von Konstrukten und Validierungsvariablen multiple Regressionsanalysen durchgeführt, ergänzt durch gerichtete Korrelationsanalysen. Die Prätest-Posttest-Analysen wurden mithilfe von t -Tests durchgeführt. Alle Analysen wurden in R (Version 4.02, RCore Team, 2014) durchgeführt. Die Effektstärke d wurde mithilfe von *Psychometrica* (Lenhard & Lenhard, 2017) ermittelt.

Ergebnisse

Reliabilitätsschätzungen und Interkorrelationen. Die internen Konsistenzen waren mit über $a = .70$ alle zufriedenstellend bis gut (Tab. 2). Die Interkorrelationen waren alle substantiell; jedoch fielen die Korrelationen innerhalb der

Tabelle 2. Deskriptive Werte der HAREBE-Einzelskalen und -Gesamtskalen

Var	N	M	SD	S	K	α
HA-KOMP	192	4.06	0.92	-0.39	-0.13	.88
RE-KOMP	192	3.71	1.03	-0.23	0.01	.90
BE-KOMP	193	4.68	0.77	-0.68	0.51	.79
HAREBE-KOMP	192	4.15	0.69	-0.21	0.40	.88
HA-LERN	192	3.85	0.96	-0.29	0.10	.79
RE-LERN	191	3.85	1.03	-0.32	0.06	.84
BE-LERN	192	3.86	1.09	-0.26	-0.50	.88
HAREBE-LERN	190	3.86	0.94	-0.21	-0.10	.93

Anmerkungen: Anzahl der Fälle (N), Mittelwert (M), Standardabweichung (SD), Schiefe (S), Exzess (K) und interne Konsistenz (Cronbachs α) der Skalen, korrespondierend zu den Faktoren erster und zweiter Ordnung.

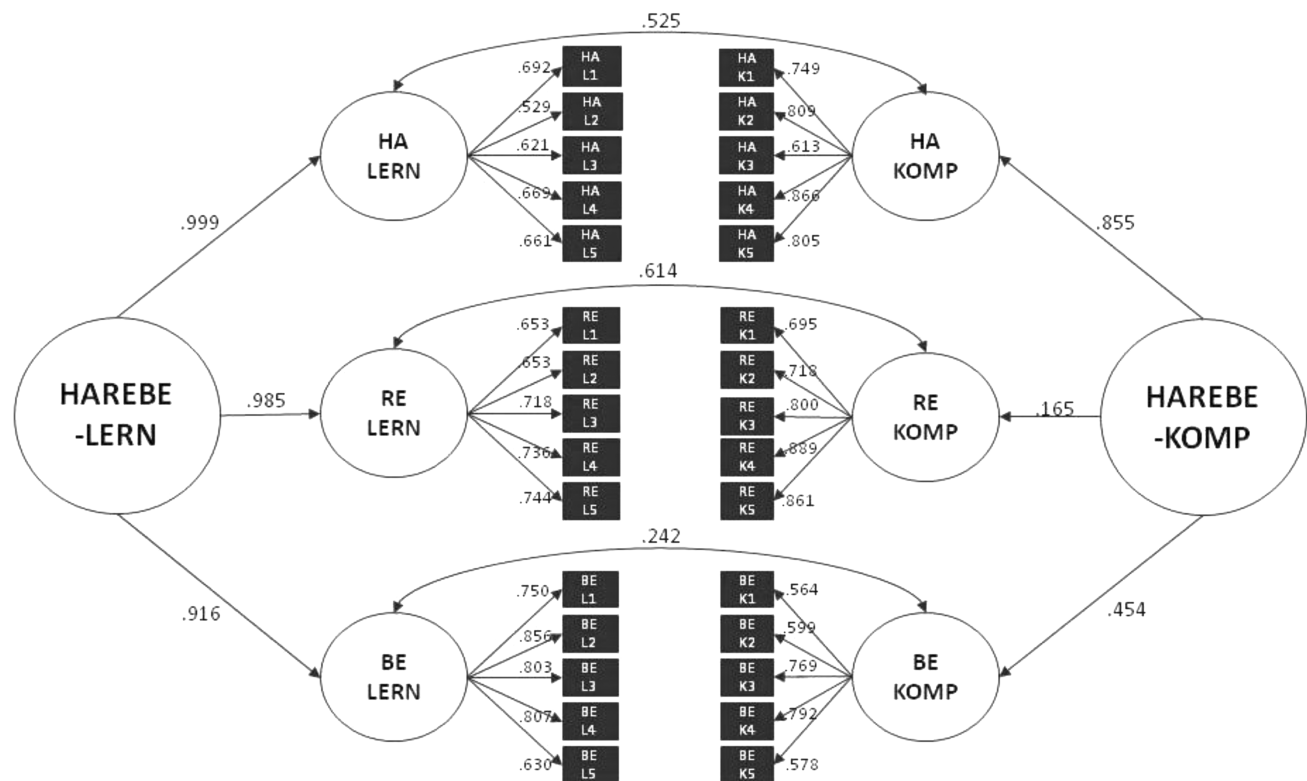


Abbildung 1. Strukturmodell der konfirmatorischen Faktorenanalyse. Die Faktorladungen und Korrelationskoeffizienten sind eingetragen.

Lernangebotsskalen ($r = .75 - r = .80$) deutlich höher aus als innerhalb der Kompetenzskalen ($r = .24 - r = .45$).

Faktorstruktur. Der Likelihood-Ratio-Test des postulierten hierarchischen Gesamtmodells war mit $\chi^2 = 536.954$ bei 396 Freiheitsgraden signifikant ($p < .001$). Das Modell zeigte einen akzeptablen bis guten Modellfit ($CFI = .942$, $RMSEA = .048$, $SRMR = .061$) mit den Informationskriterien $AIC = 15406.165$ und $BIC = 15629.480$. Die Faktorladungen sind in Abbildung 1 aufgeführt. Im ESM 3 findet sich außerdem eine Tabelle mit den Wortlauten der einzelnen Items, den Faktorladungen und den relevanten

Itemkennwerten. Das untersuchte Alternativmodell wies einen etwas schlechteren Modellfit auf ($CFI = .933$, $RMSEA = .052$, $SRMR = .062$) und schnitt auch bei den Informationskriterien schlechter ab ($AIC = 15434.367$, $BIC = 15647.973$).

Zusammenhänge der Kompetenzskalen mit personen- und gruppenspezifischen Merkmalen. Um Zusammenhänge der drei Kompetenzskalen mit den Personen- bzw. Gruppenmerkmalen *Fachwissen*, *Fachinteresse* und *studienrelevantes Praktikum* zu untersuchen, wurden multiple (logistische) Regressionen berechnet. Die Ergebnisse der Korre-

lationsanalysen werden berichtet, wenn sie deutliche Abweichungen aufweisen¹.

Die multiple Regression mit den Skalen HA-, RE- und BE-KOMP als unabhängige Variablen (UV) und dem personenbezogenen Merkmal *Fachwissen* als abhängige Variable (AV) war mit $F(3, 98) = 4.13, p < .01$ signifikant und zeigte mit $R^2 = .085$ eine geringe bis mittlere Varianzaufklärung nach Cohen (1988). Lediglich HA-KOMP war signifikant, $b = .403, t = 3.31, p < .01$.

Das Regressionsmodell mit der AV *Fachinteresse* war mit $F(3, 107) = 13.93, p < .001$ signifikant und zeigte eine hohe Varianzaufklärung, $R^2 = .261$. Lediglich RE-KOMP war in dem Modell signifikant, $b = .426, t = 4.78, p < .001$. Bei zusätzlich durchgeführten, gerichteten Korrelationsanalysen zeigten sich allerdings auch bei HA-KOMP ($r = .32, p < .001$) und BE-KOMP ($r = .26, p < .01$) moderate signifikante Zusammenhänge.

Das logistische Regressionsmodell mit der binären AV *Praktikum* war mit $\chi^2(173) = 218.28, p = .447$ nicht signifikant.

Zusammenhänge der Kompetenzskalen mit konstrukt-nahen und -fernen Variablen. Als konstrukt-nahe Variablen zur Validierung der Kompetenzskalen waren wie oben beschrieben die *Skala zur Erfassung der allgemeinen Selbstwirksamkeitserwartung* für HA-KOMP, die Skalen *Reflexionsfähigkeit* und *Theorie-Praxis-Reflexion* für RE-KOMP und mehrere Skalen zum beruflichen Wissen für BE-KOMP ausgewählt worden. Es wurden multiple Regressionsanalysen mit den drei Kompetenzskalen (HA-, RE-, BE-KOMP) als UV und den jeweiligen konstrukt-nahen Skalen als AV durchgeführt.

Das Regressionsmodell mit der AV *Selbstwirksamkeitserwartung* war mit $F(3, 127) = 9.63, p < .001$ signifikant und zeigte eine mittlere Varianzaufklärung, $R^2 = .166$. Lediglich HA-KOMP war signifikant, $b = .460, t = 4.06, p < .001$.

Das Regressionsmodell mit der abhängigen Variable *Reflexionskompetenz* war mit $F(3, 127) = 8.42, p < .001$ signifikant und zeigte eine mittlere Varianzaufklärung mit $R^2 = .146$. Dabei waren RE-KOMP mit $b = .341, t = 3.26, p < .01$ und BE-KOMP mit $b = .252, t = 2.69, p < .01$ signifikant. Ein ähnliches Muster zeigte sich auch bei den einzelnen Korrelationen, wobei HA-KOMP ($r = .23, p < .01$) eine geringe, aber signifikante Korrelation aufwies.

Das Regressionsmodell mit der AV *Theorie-Praxis-Reflexion* war mit $F(3, 126) = 9.85, p < .001$ signifikant und zeigte eine mittlere Varianzaufklärung mit $R^2 = .171$. Dabei waren RE-KOMP mit $b = .511, t = 5.25, p < .001$ und HA-

KOMP mit $b = -.233, t = -2.21, p < .05$ signifikant. Bei den Korrelationsanalysen war hingegen nur der Zusammenhang mit RE-KOMP ($r = .40, p < .001$) signifikant.

Die Regressionsmodelle mit den AV *Selbstwissen* ($F(3, 127) = 17.37, p < .001, R^2 = .274$), *Bedingungswissen* ($F(3, 127) = 15.79, p < .001, R^2 = .255$) sowie *Planungskompetenz* ($F(3, 127) = 13.08, p < .001, R^2 = .218$) zeigten signifikante Ergebnisse für die Gesamtmodelle. Dabei trug jeweils allein BE-KOMP bezogen auf das Selbstwissen mit $b = .470, t = 5.89, p < .001$, auf das Bedingungswissen mit $b = .462, t = 5.71, p < .001$ und die Planungskompetenz mit $b = .431, t = 5.00, p < .001$ signifikant zur Vorhersage bei. Bei den Korrelationsanalysen war das Muster ähnlich, aber nicht ganz so deutlich: Hier zeigte BE-KOMP zwar die stärksten Zusammenhänge mit $r = .47$ bis $r = .53, p < .001$, aber auch HA-KOMP ($r = .28$ bis $r = .29, p < .01$) und RE-KOMP ($r = .09, ns$, bis $r = .18, p < .05$) zeigten kleinere signifikante Zusammenhänge mit den Validierungsvariablen.

Prä-Post-Veränderung der Kompetenzeinschätzungen durch eine situierte Lehrveranstaltung. Die Analyse der Kompetenzeinschätzungen im Posttest (nach Abschluss der situierten Lehrveranstaltung) im Vergleich zum Prätest (zu Beginn der situierten Lehrveranstaltung) ergab, dass Studierende der o.g. Teilstichprobe sowohl HA-KOMP ($M_{\text{post}} = 4.29, SD_{\text{post}} = 0.80, M_{\text{pre}} = 3.97, SD_{\text{pre}} = 0.86$), $t(105) = 4.79, p < .001, d = 0.389$) als auch BE-KOMP ($M_{\text{post}} = 4.78, SD_{\text{post}} = 0.71, M_{\text{pre}} = 4.50, SD_{\text{pre}} = 0.77, t(105) = 3.74, p < .001, d = 0.345$) im Posttest signifikant höher einschätzten als im Prätest. Bei RE-KOMP ($M_{\text{post}} = 3.71, SD_{\text{post}} = 1.06, M_{\text{pre}} = 3.55, SD_{\text{pre}} = 1.08, t(105) = 1.62, p = .054, d = 0.164$) war das nicht der Fall. Im Kontrast dazu wurde zusätzlich die Veränderungssensitivität der Leistungsmotivation zwischen Prä- und Posttest untersucht. Die Studierenden schätzten ihre *Hoffnung auf Erfolg* im Posttest nicht signifikant höher ein als im Prätest ($M_{\text{post}} = 4.79, SD_{\text{post}} = 0.81, M_{\text{pre}} = 4.73, SD_{\text{pre}} = 0.82, t(105) = 0.75, p = .229, d = 0.065$).

Zusammenhänge der Lernangebotsskalen mit konstrukt-nahen Variablen. Zur Validierung von BE-LERN war die Skala *Berufstransfer* erhoben worden. Das multiple Regressionsmodell mit den Lernangebotsskalen (HA-, RE-, BE-LERN) als UV und der AV *Berufstransfer* war mit $F(3, 127) = 30.16, p < .001$ signifikant und zeigte eine hohe Varianzaufklärung mit $R^2 = .402$. Dabei war allein BE-LERN mit $b = .513, t = 4.52, p < .001$ signifikant. Bei den Korrelationsanalysen ging das Muster in eine ähnliche, aber nicht ganz so eindeutige Richtung: Die Korrelation

¹ Alle Korrelationen wurden mit *Pearsons Produkt-Moment-Korrelation* berechnet und zusätzlich mit *Spearman's Rho* überprüft. Es zeigten sich nur geringe Abweichungen in den Korrelationen von maximal .05; bis auf die Korrelation zwischen BE-KOMP und *Fachinteresse*, die bei $r = .26$ und $p = .47$ lag.

mit BE-LERN war am höchsten ($r = .64, p < .001$), allerdings waren auch die Korrelationen mit HA-LERN ($r = .51, p < .001$) und RE-LERN ($r = .52, p < .001$) substantiell.

Zuletzt wurde der Zusammenhang der drei Lernangebotskalen mit dem Fragebogen *InstEvaL Lehrveranstaltung* in einer multiplen Regression untersucht. Das Regressionsmodell war mit $F(3, 106) = 11.50, p < .001$ signifikant und zeigte eine mittlere bis hohe Varianzaufklärung, $R^2 = .224$. Dabei war allein BE-LERN mit $b = .326, t = 2.08, p < .05$ signifikant. Während bei den Korrelationsanalysen die Korrelation mit BE-LERN ($r = .47, p < .001$) weiterhin substantiell ausfiel, waren auch die Korrelationen mit den Skalen HA-LERN ($r = .46, p < .001$) und RE-LERN ($r = .39, p < .001$) signifikant.

Diskussion

Ziel der vorliegenden Arbeit ist es, ökonomisch einsetzbare Skalen zur Evaluation situierter Lehrveranstaltungen zur Verfügung zu stellen. Je drei Skalen zur studentischen Selbsteinschätzung relevanter Kompetenzkonstrukte (Handlungs-, Reflexions- und Berufswahlkompetenz) und zur Einschätzung des Lernangebots (Handlungs-, Reflexions- und Berufsorientierung) wurden neu entwickelt und validiert. Die Reliabilitätsschätzungen zeigen zufriedenstellende bis gute Werte. Die Ergebnisse zur Validität sprechen für die Anwendbarkeit der entwickelten Skalen. Die angenommene hierarchische Modellstruktur mit sechs latenten Faktoren erster Ordnung (entsprechend den sechs spezifischen Skalen) und zwei latenten, orthogonalen Faktoren zweiter Ordnung (den allgemeinen Kompetenz- und Lernangebotseinschätzungen), konnte bestätigt werden.

Bei der Skalenentwicklung wurde insbesondere auf folgende Aspekte geachtet: (1) Relevante Kompetenzaspekte situierter Lernens wurden ausgewählt. (2) Die Skalen Handlungs- und Reflexionskompetenz beziehen sich ausschließlich auf Kompetenzeinschätzungen und nicht auf Überzeugungen und Interessen. (3) Die Itemformulierungen sind nicht an bestimmte Fächer gebunden, sondern inhaltlich unspezifisch gehalten und damit in verschiedenen Fächern nutzbar. (4) Die Skalen sind ökonomisch einsetzbar.

Bezüglich der Faktorstruktur weist der Faktor RE-KOMP erster Ordnung eine vergleichsweise schwache Faktorladung auf dem Faktor HAREBE-KOMP zweiter Ordnung auf. Es scheint also deutlich zwischen den unterschiedlichen Facetten unterschieden zu werden; wobei insbesondere die eigene Fähigkeit zur Verknüpfung von Theorie und Praxis als wenig zusammenhängend mit den anderen Kompetenzen wahrgenommen wird. Zudem wird die eige-

ne Kompetenz zur Theorie-Praxis-Reflexion als eher gering eingeschätzt. In den Prätest-Posttest-Analysen verzeichnet die Reflexionskompetenz als einziges Maß keinen signifikanten Zuwachs. Möglicherweise erkennen Studierende erst beim Versuch der Theorie-Praxis-Reflexion deren Komplexität und stellen ein bislang unbekanntes Kompetenzdefizit fest. Möglich ist auch, dass sie das Lernangebot zwar wahrnehmen, die Kompetenzentwicklung in diesem Bereich jedoch nicht als zentrales persönliches Lernziel erachten (Angebot vs. Nutzung; Seidel, 2014). Möglicherweise benötigt die Kompetenzentwicklung der Theorie-Praxis-Reflexion intensivere und unterschiedliche Lerngelegenheiten über einen längeren Zeitraum. Zu Beginn des Studiums und in nur einer situierten Lehrveranstaltung sind Zuwächse daher nicht erkennbar.

Die hoch interkorrelierten Faktoren HA-LERN, RE-LERN und BE-LERN erster Ordnung zeigen hohe Faktorladungen auf dem zugehörigen Faktor HAREBE-LERN zweiter Ordnung. Studierende scheinen bei der Beurteilung einer Lehrveranstaltung eine stärker generalisierte Einschätzung abzugeben. Die Informationskriterien der getesteten Modelle deuten allerdings darauf hin, dass die postulierte, separate Modellierung der drei Lernangebotsfaktoren gegenüber einem generalisierten LERN-Faktor allein zu bevorzugen ist.

Bei der Validierung durch Zusammenhänge mit personen- und gruppenspezifischen Merkmalen ist u.a. das bildungspsychologische Fachwissen (selbstberichtete Klausurnote) herangezogen worden. Die Regressionsgewichte und Korrelationsmuster zeigten hier erwartungswidrig nur schwache Zusammenhänge mit den Kompetenzskalen. Wir betrachten die herangezogene Klausurnote der bildungswissenschaftlichen Vorlesung zwar als hinreichend objektives und inhaltsvalides Maß für das erworbene bildungspsychologische Fachwissen im Lehramt. Jedoch weist die Klausurnote nicht die Qualität eines psychometrisch überprüften und validierten Tests auf und hat studiengangsspezifische Schwerpunktsetzungen. In Zukunft wäre ein empirisch geprüfter und breit zusammengestellter Fachwissenstest für bildungswissenschaftliches Wissen zu verwenden (z.B. aus dem BilWiss-Projekt; Kunter et al., 2016). Die Klausurdurchschnittsnote ($M = 2.7, SD = 0.80$) könnte zudem darauf hindeuten, dass Studierenden das Fachwissen fehlte, um passende Wissens Elemente in praktischen Handlungssituationen abzurufen. Ferner könnte das deklarative Wissen zwar vorhanden, aber nur in bestimmten Kontexten und mit bestimmten Hinweisreizen abrufbar sein, die denen der Lernsituation ähneln. Es würde sich somit um ein Problem des „trägen Wissens“ (Gruber, Mandl & Renkl, 2000, S. 1) handeln.

Ebenso zeigte sich, dass das Fachinteresse v.a. mit Reflexionskompetenz zusammenhängt; diese könnte mit

intensiverer theoretisch-fachlicher Auseinandersetzung einhergehen und von höherem Fachinteresse profitieren. Hohes Fachinteresse könnte die wahrgenommene Relevanz theoretischen Wissens für Anwendungssituationen begünstigen. Im Gegensatz dazu könnte Handlungskompetenz auch dann erlebt werden, wenn man sich in Praxissituationen als selbstwirksam und handlungsfähig erlebt; unabhängig von den vermittelten theoretischen Fachinhalten. Auch Prozesse in Bezug auf die Berufswahlkompetenz könnten abgekoppelt vom Fachinteresse erfolgen; insbesondere, wenn eine hohe Theorie-Praxis-Diskrepanz wahrgenommen wird.

Die studienrelevante Praktikumserfahrung hatte bei der hier untersuchten Stichprobe keinen beobachtbaren Einfluss auf die spätere Kompetenzeinschätzung. Die Erfassung der Praktikumserfahrung könnte zu unspezifisch gewesen und zu früh im Studium erfolgt sein. Eine künftige Studie sollte Masterstudierende miteinbeziehen, da diese ihre Praktika vermutlich stärker hinsichtlich möglicher Berufsfelder und mit spezifischerem Handlungsbezug zum Fach auswählen.

Die Korrelationsmuster der Kompetenz- und Lernangebotsskalen mit den konstruktiven Skalen zeigen die erwartete inhaltliche Nähe. Zu den konstruktfernen Skalen sind die Regressionsgewichte und Korrelationen durchweg niedriger. Die Kompetenzskalen HA-KOMP und BE-KOMP zeigten sich im Prätest-Posttest-Vergleich veränderungssensitiv; es konnte eine Zunahme der selbst eingeschätzten Handlungs- und Berufswahlkompetenz im Zusammenhang mit einer situierten Lehrveranstaltung gezeigt werden. Zwar zeigte auch die Skala RE-KOMP eine Zunahme, diese war jedoch nicht signifikant. Erwartungskonform zeigten sich gleichzeitig keine signifikanten Veränderungen in der Leistungsmotivation.

Die HAREBE-Skalen wurden bei der Entwicklung und Pilotierung in situierten Lehrveranstaltungen verschiedener Fächer eingesetzt. Allerdings stammen die wesentlichen Validierungsergebnisse aus einem Bachelor-Lehramtsstudiengang. Um die Generalisierbarkeit der Ergebnisse zu untermauern, sind weitere Validierungsstudien in verschiedenen Fächern und Fachsemestern durchzuführen. Wünschenswert sind außerdem Querschnitts- (z. B. Bachelor- vs. Masterstudierende) oder Längsschnittstudien zur weiteren Untersuchung der studentischen Kompetenzentwicklung.

Ausblick

In einer Studienserie (Nowakowski, Verweken, Braun & Hannover, 2012), die den Einfluss unterschiedlicher Arten von Rückmeldungen auf die Studierendenorientierung von Dozierenden untersuchte, wurde festgestellt,

dass ergebnisorientierte (subjektive Kompetenzeinschätzungen) im Gegensatz zu prozessorientierten (Veranstaltungszufriedenheit) Studierendenrückmeldungen einen positiven Einfluss auf die anschließende Studierendenorientierung der Lehrenden hatten. In dieser Hinsicht können die Kompetenzskalen z. B. in Form einer Prä-Post-Veränderungsmessung wertvolle Wirkungshinweise für Lehrende geben. Dozierenden können nachvollziehen, welchen Lernzugewinn Studierende für sich selbst sehen und ihre Lehrschwerpunkte daran orientieren.

Darüber hinaus bieten die Skalen zur Handlungs-, Reflexions- und Berufsorientierung eine dringend benötigte Erweiterung von Lehrveranstaltungsevaluationen. Obwohl es in der Hochschullehre schon länger situierte Lehr-Lernformate gibt, wurden entscheidende Aspekte in standardisierte Lehrveranstaltungsevaluationen bisher nicht integriert. Die HAREBE-Skalen stellen hier eine innovative Ergänzung dar.

Die HAREBE-Skalen bieten darüber hinaus Anwendungsmöglichkeiten in der Forschung. Sie können als abhängige Variablen eingesetzt werden, um Interventionen für situierte Lehrveranstaltungen in der hochschuldidaktischen Forschung zu untersuchen – beispielsweise gerade solche Interventionen, die die Förderung von Theorie-Praxis-Reflexion zum Ziel haben.

Die Skalen sind inzwischen in das Evaluationssystem EvaSys (Electric Paper Evaluationssysteme, 2019) eingepflegt, welches deutschlandweit eingesetzt wird. Dadurch können Dozierende an Hochschulen die Skalen unkompliziert in ihr digitales Evaluationssystem einbinden und für eigene Lehrveranstaltungsevaluation nutzen.

Elektronische Supplemente (ESM)

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1026/0012-1924/a000271>

ESM 1. Validierungsvariablen.

ESM 2. Das kognitive Interview.

ESM 3. HAREBE Items.

Literatur

American Educational Research Association [AERA], American Psychological Association & National Council on Measurement in Education (2014). *Validity. Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association. Retrieved from <https://www.aera.net/Publications/Books/Standards-for-Educational-Psychological-Testing-2014-Edition>


- Baumert, J. & Kunter, M. (2013). Professionelle Kompetenz von Lehrkräften. In I. Gogolin, H. Kuper, H.-H. Krüger & J. Baumert (Hrsg.), *Stichwort: Zeitschrift für Erziehungswissenschaft* (S. 277–337). Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-658-00908-3_13
- Blömeke, S., Zlatkin-Troitschanskaia, O., Kuhn, C. & Fege, J. (2013). Modeling and Measuring Competencies in Higher Education (Professional and Vet Learning). In S. Blömeke, O. Zlatkin-Troitschanskaia, C. Kuhn & J. Fege (Eds.), *Modeling and Measuring Competencies in Higher Education: Tasks and Challenges* (pp. 1–10). Rotterdam: SensePublishers. https://doi.org/10.1007/978-94-6091-867-4_1
- Bosse, E. (2011). *Qualifizierung für interkulturelle Kommunikation: Trainingskonzeption und -evaluation*. IUDICIUM Verlag.
- Braun, E., Gusy, B., Leidner, B. & Hannover, B. (2008). Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp). *Diagnostica*, 54, 30–42. <https://doi.org/10.1026/0012-1924.54.1.30>
- Butz, B. (2008). Grundlegende Qualitätsmerkmale einer ganzheitlichen Berufsorientierung. In G.-E. Famulla, B. Butz, S. Deeken, U. Michaelis, V. Möhle & B. Schäfer (Hrsg.), *Berufsorientierung als Prozess. Persönlichkeit fördern, Schule entwickeln, Übergang sichern. Ergebnisse aus dem Programm „Schule – Wirtschaft/Arbeitsleben“* (S. 42–62). Baltmannsweiler: Schneider Verlag Hohengehren.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. <https://doi.org/10.1037/h0040957>
- Electric Paper Evaluationssysteme. (2019). *EvaSys*. Retrieved from <https://www.evasys.de/startseite.html>
- Erdfelder, E., Pesch, S. & Stahl, C. (2004). *Internetbasierte studentische Evaluation von Lehrveranstaltungen* (InstEval). Universität Mannheim.
- Goeze, A., Hetfleisch, P. & Schrader, J. (2013). Wirkungen des Lernens mit Videofällen bei Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 16(1), 79–113. <https://doi.org/10.1007/s11618-013-0352-x>
- Gollwitzer, M. & Schlotz, W. (2003). Das „Trierer Inventar zur Lehrveranstaltungsevaluation“ (TRIL): Entwicklung und erste testtheoretische Erprobungen (Psychologiedidaktik und Evaluation). In G. Krampen & H. Zayer (Hrsg.), *Neue Medien, Konzepte, Untersuchungsbefunde und Erfahrungen zur psychologischen Aus-, Fort- und Weiterbildung* (1. Aufl., S. 114–128). Berlin: Deutscher Psychologen Verlag.
- Gronostay, D. & Manzel, S. (2015). Entwicklung der professionellen Kompetenz von Studierenden im Lehramt Sozialwissenschaften – Erste Ergebnisse einer Quasi-Längsschnitterhebung zur Berufswahlmotivation. In G. Weißeno & C. Schelle (Hrsg.), *Empirische Forschung in gesellschaftswissenschaftlichen Fachdidaktiken: Ergebnisse und Perspektiven* (S. 155–165). Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-658-06191-3_11
- Gröschner, A., Schmitt, C. & Seidel, T. (2013). Veränderung subjektiver Kompetenzeinschätzungen von Lehramtsstudierenden im Praxissemester. *Zeitschrift für Pädagogische Psychologie*, 27(1–2), 77–86. <https://doi.org/10.1024/1010-0652/a000090>
- Gruber, H., Mandl, H. & Renkl, A. (2000). Was lernen wir in Schule und Hochschule: Träges Wissen? In H. Mandl & J. Gerstenmeier (Hrsg.), *Die Kluft zwischen Wissen und Handeln: empirische und theoretische Lösungsansätze* (S. 139–156). Göttingen: Hogrefe. <https://epub.uni-regensburg.de/26287/>
- Hascher, T. (2006). Veränderungen im Praktikum – Veränderungen durch das Praktikum. Eine empirische Untersuchung zur Wirkung von schulpraktischen Studien in der Lehrerbildung. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern* (S. 130–148). Weinheim und Basel: Beltz. <http://nbn-resolving.de/urn:nbn:de:0111-opus-73742>
- Hidi, S. & Harackiewicz, J. M. (2000). Motivating the Academically Unmotivated: A Critical Issue for the 21st Century. *Review of Educational Research*, 70, 151–179. <https://doi.org/10.3102/00346543070002151>
- Hu, L. & Bentler, P. M. (2009). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kiel, E. & Rost, F. (2002). *Einführung in die Wissensorganisation: grundlegende Probleme und Begriffe*. Würzburg: Ergon-Verlag.
- Klieme, E., Artelt, C. & Stanat, P. (2002). Fächerübergreifende Kompetenzen: Konzepte und Indikatoren. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. Aufl., S. 203–218). Weinheim: Beltz.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2. Aufl., S. 366). New York, NY: Guilford Press.
- König, J. & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung*. Berlin: Humboldt-Universität zu Berlin.
- Kunter, M., Baumert, J., Leutner, D., Terhart, E., Seidel, T., Dicke, T. et al. (2016). *Dokumentation der Erhebungsinstrumente der Projektphasen des BilWiss-Forschungsprogramms von 2009 bis 2016: Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung (BilWiss). Die Bedeutung des bildungswissenschaftlichen Hochschulwissens für den Berufseinstieg von Lehrkräften (BilWiss-Beruf)*. Frankfurt am Main: Johann Wolfgang von Goethe-Universität. Verfügbar unter: <http://publikationen.uni-frankfurt.de/frontdoor/index/index/docId/42800>
- Lang, J. W. B. & Fries, S. (2006). A Revised 10-Item Version of the Achievement Motives Scale. *European Journal of Psychological Assessment*, 22(3), 216–224. <https://doi.org/10.1027/1015-5759.22.3.216>
- Leiner, D. J. (2019). SoSci Survey. *SoSci Survey*. Retrieved from <https://www.socisurvey.de/>
- Lenhard, A. & Lenhard, W. (2017). *Psychometrica*. Verfügbar unter: <https://www.psychometrica.de/>
- Linninger, C. A. (2016). *Reflexion bei angehenden Lehrkräften. Bedeutung und Förderung im Professionalisierungsprozess*. Frankfurt am Main: Johann Wolfgang Goethe-Universität. <https://www.fachportal-paedagogik.de/literatur/vollanzeige.html?Fid=1117763#vollanzeige>
- Lipowski, K., Kaak, S. & Kracke, B. (2016). Individualisierung von schulischen Berufsorientierungsmaßnahmen – ein praxisorientiertes diagnostisches Verfahren zur Erfassung von Berufswahlkompetenz (Berufsorientierung im Lebenslauf – theoretische Standortbestimmung und empirische Analysen). In H. Faulstich-Wieland, S. Rahn & B. Scholand (Hrsg.), *Berufs- und Wirtschaftspädagogik – online*, 12, 1–16.
- Lipowski, K., Kaak, S., Kracke, B. & Holstein, J. (2015). *Handbuch Schulische Berufsorientierung. Praxisorientierte Unterstützung für den Übergang Schule – Beruf* (Materialien). Bad Berka: Thüringer Institut für Lehrerforschung, Lehrplanentwicklung und Medien (Thillm). https://www.schulportal-thueringen.de/get-datta/64eaf00a-415b-41de-95fd-9c63260c6956/Materialien_189_Heft.pdf
- Nowakowski, A., Vervecken, D., Braun, E. & Hannover, B. (2012). Was Hochschuldozierende aus Lehrevaluations-Rückmeldungen lernen können. *Zeitschrift für Erziehungswissenschaft*, 15(2), 253–271. <https://doi.org/10.1007/s11618-012-0275-y>

- Plake, B. S. & Wise, L. L. (2014). What Is the Role and Importance of the Revised AERA, APA, NCME Standards for Educational and Psychological Testing? *Educational Measurement: Issues and Practice*, 33(4), 4–12. <https://doi.org/10.1111/emip.12045>
- Porath, J. (2014). Individualisierung statt Geschlechterdifferenzierung – Lernaufgaben zur Förderung von Arbeits- und Berufsorientierung durch effektive Kommunikation. *Gruppendynamik und Organisationsberatung*, 45, 401–418. <https://doi.org/10.1007/s11612-014-0259-x>
- Rank, A., Gebauer, S., Hartinger, A. & Fölling-Albers, M. (2012). Situiertes Lernen in der Lehrerfortbildung. *Lehrerbildung auf dem Prüfstand*, 5(2), 180–199.
- RCore Team. (2014). *R: The R Project for Statistical Computing*. Retrieved from <https://www.r-project.org/>
- Reinders, H. (2016). *Service Learning – Theoretische Überlegungen und empirische Studien zu Lernen durch Engagement*. Weinheim und Basel: Beltz.
- Reinders, H., Hillesheim, S. & Sebald, J. (2015). *Service Learning an Universitäten: Skalendokumentation des Längsschnitts 2012–2014*. Universität Würzburg. Retrieved from <https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/11110>
- Revelle, W. (2020). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Retrieved from <https://CRAN.R-project.org/package=psych>
- Rindermann, H. & Amelang, M. (1994). *Das Heidelberger Inventar zur Lehrveranstaltungs-Evaluation: (HILVE). Handanweisung*. Heidelberg: Asanger.
- Roebers, S., Razum, O., Kutschmann, M. & Wagner, M. (2008). Entwicklung und Evaluation einer deutschen Version des Fragebogens Stroke Action Test (STAT) zur Ermittlung der Kenntnisse über das richtige Verhalten bei einzelnen Schlaganfallsymptomen. *Aktuelle Neurologie*, 35(7), 334–339. <https://doi.org/10.1055/s-2008-1067506>
- Rosseel, Y. (2012). *lavaan. latent variable analysis. The lavaan Project*. Retrieved from <http://lavaan.ugent.be/about.html>
- Rosseel, Y., Jorgensen, T. D., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V. et al. (2019). *lavaan: Latent Variable Analysis*. Retrieved from <https://CRAN.R-project.org/package=lavaan>
- RStudio Team. (2019). *RStudio*. Retrieved from <https://rstudio.com/products/team/>
- Schaper, N., Reis, O., Wildt, J., Horvath, E. & Bender, E. (2012). *Fachgutachten zur Kompetenzorientierung in Studium und Lehre. HRK-Fachgutachten. HRK nexus*. Verfügbar unter: https://www.hrk-nexus.de/fileadmin/redaktion/hrk-nexus/07-Downloads/07-02-Publikationen/fachgutachten_kompetenzorientierung.pdf
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research*, 8(2), 23–74.
- Schindler, G. (2004). Employability und Bachelor-Studiengänge – eine unpassende Verbindung. *Beiträge zur Hochschulforschung*, 4, 6–26.
- Schön, D. A. (1991). *The reflective practitioner: how professionals think in action*. Aldershot: Avebury.
- Schubarth, W., Speck, K., Seidel, A., Gottmann, C., Kamm, C. & Krohn, M. (2012). Das Praxissemester im Lehramt – ein Erfolgsmodell? Zur Wirksamkeit des Praxissemesters im Land Brandenburg. In W. Schubarth, K. Speck, A. Seidel, C. Gottmann, C. Kamm & M. Krohn (Hrsg.), *Studium nach Bologna: Praxisbezüge stärken?! Praktika als Brücke zwischen Hochschule und Arbeitsmarkt* (S. 137–169). Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-531-19122-5_8
- Schwarzer, R. & Jerusalem, M. (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: Freie Universität Berlin.
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60(6), 850–866.
- Sliwka, A. (2009). Reflexion: Das Bindeglied zwischen Service und Lernen. In K. Altenschmidt, J. Miller & W. Stark (Hrsg.), *Raus aus dem Elfenbeinturm? Entwicklungen in Service Learning und bürgerschaftlichem Engagement an deutschen Hochschulen* (S. 86–90). Weinheim und Basel: Beltz.
- Voss, T., Kunina-Habenicht, O., Hoehne, V. & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, 18(2), 187–223. <https://doi.org/10.1007/s11618-015-0626-6>
- Wagner, K., Klein, M., Klopp, E. & Stark, R. (2014). Instruktionale Unterstützung beim Lernen aus advokatorischen Fehlern in der Lehramtsausbildung. Effekte auf die Anwendung wissenschaftlichen Wissens. *Psychologie in Erziehung und Unterricht*, 61, 287–301. <https://doi.org/10.2378/peu2014.art23d>
- Wildt, J. (2013). Entwicklung und Potentiale der Hochschuldidaktik. In M. Heiner & J. Wildt (Hrsg.), *Professionalisierung der Lehre. Perspektiven formeller und informeller Entwicklung von Lehrkompetenz im Kontext der Hochschulbildung* (S. 27–57). Bielefeld: Bertelsmann.
- Wissenschaftsrat (2008). Empfehlungen zur Qualitätsverbesserung von Lehre und Studium. Verfügbar unter: https://www.wissenschaftsrat.de/download/archiv/8639-08.pdf?__blob=publicationFile&v=2
- Zumbach, J., Spinath, B., Schahn, J., Friedrich, M. & Kögel, M. (2007). Entwicklung einer Kurzskaala zur Lehrevaluation. In M. Krämer, K. Brusdeylins & S. Preiser (Hrsg.), *Psychologiedidaktik und Evaluation*, 1–10. Göttingen: Vandenhoeck & Ruprecht.

Förderung

Gefördert vom Land Baden-Württemberg, Ministerium für Wissenschaft, Forschung und Kunst, Programm „Wissenschaft lehren und lernen“ (WILLE), Laufzeit: 05/2016–12/2019, Umfang: 699.948,83 €
Open Access-Veröffentlichung ermöglicht durch die Universität Mannheim

ORCID

Anne-Sophie Waag
 <https://orcid.org/0000-0001-8031-5931>

Anne-Sophie Waag

Professur für Bildungspsychologie
Universität Mannheim
A5, 6
68131 Mannheim
waag@uni-mannheim.de