

**Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin**

Erfassung und Förderung studentischer
Forschungskompetenzen

Validierung eines Instrumentes zur Erfassung studentischer
Forschungskompetenzen sowie Gestaltung und Evaluation einer
forschungsorientierten Lernumgebung

Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

vorgelegt von

Franziska Böttcher-Oschmann, M.A.

Berlin, Juni 2019

Erstgutachterin: Prof. Dr. Felicitas Thiel, Freie Universität Berlin

Zweitgutachter: Prof. Dr. Rainer Watermann, Freie Universität Berlin

Datum der Disputation: 17.09.2019

Inhalt

Zusammenfassung	5
Abstract	7
1 Einleitung	9
2 Theoretischer Rahmen.....	15
2.1 Ableitung von Forschungskompetenzen aus zentralen, akademischen Kompetenzen	16
2.2 Modellierung und Erfassung von Forschungskompetenzen	23
2.3 Förderung von Forschungskompetenzen.....	31
2.4 Fragestellungen und Überblick über die empirischen Studien.....	40
Literatur I.....	43
3 Evaluating research-oriented teaching: a new instrument to assess university students’ research competences	55
3.1 Introduction	57
3.2 Methods	64
3.3 Results	68
3.4 Discussion	73
3.5 References	79
4 Validierung eines Fragenbogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen - Ein Instrument zur Evaluation forschungsorientierter Lehr- Lernarrangements	85
5 How do we prepare teacher training students for Evidence-based Practice? Promoting students’ research competences in Research-learning projects	125
5.1 Introduction	127
5.2 Methodology and methods	135
5.3 Results	140
5.4 Discussion	143
5.5 Conclusion.....	150
5.6 References	151

6 Gesamtdiskussion.....	157
6.1 Zusammenfassung und Diskussion der Ergebnisse.....	160
6.2 Grenzen und Stärken der Arbeit.....	170
6.3 Ausblick.....	173
Literatur II	176
Abbildungs- und Tabellenverzeichnis.....	183
Danksagung.....	185
Selbstständigkeitserklärung.....	187
Anteile an Studien mit Co-Autorenschaft	189

Zusammenfassung

Im Rahmen dieser Arbeit wird der Fragebogen zur Erfassung studentischer Forschungskompetenzen (F-Komp) via Selbsteinschätzungen als ein Instrument zur fach- bzw. disziplinübergreifenden Evaluation Forschungsorientierter Lehre (FoL) vorgestellt. Dieses Instrument basiert auf dem RMRK-W-Modell mit seinen Dimensionen *Recherche-, Methoden-, Reflexions- und Kommunikationskompetenzen* sowie *Fachliches Wissen*, dem die Annahme von fach- bzw. disziplinübergreifend gemeinsamen Phasen des Forschungsprozesses zugrunde liegt. Das Modell kann darüber hinaus als Grundlage für die Gestaltung von Lernarrangements zur Förderung studentischer Forschungskompetenzen dienen. Durch diese Förderung im Rahmen von Forschungsorientierter Lehre kann ein Beitrag zur Stärkung des Research-Teaching-Nexus geleistet werden.

Im theoretischen Rahmen dieser Arbeit werden studentische Forschungskompetenzen zuerst aus generischen, akademischen Kompetenzen abgeleitet, um anschließend deren domänenübergreifende Modellierung und Erfassung zu thematisieren. Danach folgt eine Darstellung von Merkmalen, welche die Förderung studentischer Forschungskompetenzen begünstigen sollten, und die Identifikation von Lehramtsstudierenden als eine besondere Zielgruppe für den Erwerb von Forschungskompetenzen für eine evidenzbasierte Praxis. Aus den theoretischen Überlegungen dieses Rahmens werden die beiden Forschungsfragen abgeleitet, welche in den drei empirischen Studien dieser Arbeit untersucht werden: (1) Können Forschungskompetenzen mit dem F-Komp valide erfasst werden? (2) Lassen sich Kompetenzen von Lehramtsstudierenden für eine evidenzbasierte Praxis in einer kompetenzmodellbasierten, forschungsorientierten Lernumgebung fördern?

In Studie 1 wurde anhand einer Stichprobe von 391 Studierenden in Bachelor-, Master- und Promotionsstudiengängen aus unterschiedlichen Disziplinen die faktorielle Validität des F-Komp mittels konfirmatorischer Faktorenanalysen überprüft. Es zeigte sich, dass sich die fünf theoretischen Dimensionen des F-Komp auch empirisch abbilden ließen und zusätzlich zwölf Sub-Dimensionen identifiziert wurden, welche die Facetten des RMRK-W-Modells widerspiegeln und zum Teil kombinierten. In Studie 2 wurden unterschiedliche Erwartungen formuliert hinsichtlich der Fragen, ob mit dem F-Komp (1) keine allgemeinen Lern- und Studienkompetenzen erfasst werden, sowie (2) inwieweit das Instrument valide im Hinblick auf die gewählte Methodik der Selbsteinschätzung misst. Diese Annahmen zur konvergenten und diskriminanten Validität des F-Komp wurden überprüft anhand von drei Teilstichproben mit 536 Masterstudierenden aus unterschiedlichen Fächern (Teilstichprobe 1),

98 Lehramtsstudierenden im Master (Teilstichprobe 2) sowie 78 Lehramtsstudierenden im Master (Teilstichprobe 3). Es zeigte sich, dass mit dem F-Komp keine Lern- oder Studienkompetenzen erfasst werden. Darüber hinaus zeigten sich bis zu mittlere Zusammenhänge zu einem Kompetenztest, der bildungswissenschaftliche Forschungskompetenzen erfasst und hohe Zusammenhänge zu forschungsbezogenen Selbstwirksamkeitserwartungen. In Studie 3 wurde für Lehramtsstudierende ein Lernforschungsprojekt als eine forschungsorientierte Lernumgebung zur Förderung von Forschungskompetenzen für eine evidenzbasierte Praxis konzipiert und evaluiert. Diese modellbasierte Lernumgebung wurde in einem quasi-experimentellen Prä-Post-Design mit einem Kompetenztest zur Erfassung bildungswissenschaftlicher Forschungskompetenzen sowie mit dem F-Komp unter Berücksichtigung des Response-Shift-Phänomens evaluiert. Neben 36 Studierenden der Lernforschungsprojekte, die eine eigene Forschungsfrage bearbeiteten, konnte ein Seminar mit 23 Studierenden als Kontrollgruppe hinzugenommen werden, welches die Anwendung forschungsbasierten Wissens für die Lehr-Lernpraxis thematisiert. Die Ergebnisse zeigten, dass sich die Studierenden der jeweiligen Seminargruppen vor allem in denjenigen Bereichen unterschieden, die auf das Durchlaufen des Forschungsprozesses abzielten. Abschließend werden die Ergebnisse der drei Studien integrierend vor dem Hintergrund möglicher Einflussfaktoren diskutiert, maßgebliche Grenzen der Arbeit werden aufgezeigt, sowie Implikationen für Forschung und Praxis dargestellt.

Schlagwörter: Forschungskompetenzen – Forschungsorientierte Lehre (FoL) — evidenzbasierte Praxis – confirmatorische Faktorenanalysen – quasi-experimentelles Prä-Post-Kontrollgruppen-Design – Response-Shift

Abstract

In the present thesis, the instrument to assess student *research competences* (R-Comp) via self-ratings is introduced as a tool for a cross-disciplinary evaluation of research-oriented teaching. This instrument is based on the RMRC-K-model comprising five dimensions: skills in reviewing the state of research, *methodological skills*, skills in reflecting on research findings, *communication skills*, and content *knowledge*. The model is founded upon the assumption that research shares common characteristics across various disciplines. Moreover, the model can serve as a basis for designing research-learning-projects to promote students' research competences. By promoting these competences in research-oriented teaching, the research-teaching-nexus can be strengthened.

In the theoretical framework of the thesis, research competences are first derived from essential academic competences. Then, the cross-disciplinary modelling and assessment of research competences are discussed. This is followed by characteristics that should promote the acquisition of student research competences as well as the identification of teacher training students as a special group to be prepared for an evidence-based practice. Based on these theoretical assumptions, two research questions are derived and investigated in three empirical studies: (1) Does the R-Comp validly measure research competences? (2) Can teacher training students' competences for evidence-based practice be promoted in a competence model-based, research-oriented learning environment?

In Study 1, the R-Comp's factorial structure was examined by confirmatory factor analyses with data from 391 university students in three groups, either enrolled in a Bachelor's degree, Master's degree, or a PhD program. The sample represented various academic disciplines. Results supported the hypothesized structure of the R-Comp for the five dimensions in accordance with the RMRC-K-model. Moreover, they provided evidence for a more detailed differentiation of all dimensions with twelve sub-dimensions reflecting theorized facets of the RMRC-K-model. In Study 2, different expectations were formulated regarding the questions of whether the R-Comp (1) does not measure competences for learning and studying and (2) to what extent the instrument measures validly with regard to the chosen method of self-assessment. These assumptions about convergent and discriminant validity of the R-Comp were analysed on the basis of three subsamples with 536 Master's students from different subjects (subsample 1), 98 teacher training students in Master's program (subsample 2), and 78 teacher training students in Master's program (subsample 3). Results indicated that the R-Comp did not measure competences for learning and studying. In addition,

there were up to medium correlations with a competence test measuring Educational Research Literacy and large correlations with research self-efficacy. In Study 3, a research-learning project to promote teacher training students' research competences for evidence-based practice was designed and evaluated. This model-based learning environment was evaluated in a quasi-experimental pre-post design including a competence test to assess Educational Research Literacy as well as the R-Comp, considering the response shift phenomenon. In addition to 36 students from research-learning projects, who investigated their own research questions, a seminar with 23 students could be added as a control group, which focused on the application of research-based knowledge for practice. Results indicated that students from the two seminar groups differed in those dimensions that focused on establishing research. In the last part of the thesis, key findings from the three empirical studies are discussed against the background of possible influencing factors. Moreover, significant limitations of the thesis are addressed and implications for research and practice are presented.

Keywords: Research Competences - Research-oriented Teaching – Evidence-based Practice – Confirmatory Factor Analyses – Quasi-Experimental Pre-Post-Control-Group-Design – Response-Shift

Kapitel 1

Einleitung

1 Einleitung

Die Verknüpfung von Forschung und Lehre, der sogenannte Research-Teaching-Nexus, erlangte im universitären Kontext in den vergangenen 20 Jahren verstärkt Aufmerksamkeit (Brew, 2006; Griffiths, 2004; Healey, 2005; Healey & Jenkins, 2009). Ausschlaggebend waren neben hochschulpolitischen Entwicklungen (z.B. Griffiths, 2004) empirische Hinweise dafür, dass die Qualität der Lehre an Universitäten nicht automatisch von der Forschungsproduktivität ihrer Mitglieder profitierte (Griffiths, 2004; Healey, 2005). Diese Hinweise stammten aus einer fächerübergreifenden Meta-Analyse von Hattie und Marsh (1996), in der kein empirischer Zusammenhang zwischen der Forschungsproduktivität und der Lehrqualität von ForscherInnen¹ mit Lehraufgaben gefunden wurde. Die Forderung wurde laut, dass der Zusammenhang zwischen diesen beiden Bereichen gestärkt werden müsse: „Universities need to set as a mission goal the improvement of the nexus between research and teaching“ (Hattie & Marsh, 1996, S. 533). Daraufhin wurden unterschiedlichste Konzepte erarbeitet, wie Forschung stärker in die universitäre Lehre integriert werden kann (Bundesassistentenkonferenz, 2009; Griffiths, 2004; Healey, 2005; Healey & Jenkins, 2009; L. Huber, 2014; Reinmann, 2016, 2017; Reinmann, Lübcke & Heudorfer, 2019).

Im Rahmen der zweiten Förderphase der Exzellenzinitiative von Bund und Ländern von 2012 bis 2017 bekam die Debatte um diese Verknüpfung eine neue Dynamik. Unter dem Begriff *Forschungsorientierte Lehre (FoL)* wurde die Notwendigkeit der Integration von Forschung in Lehre als ein zentraler Bestandteil der exzellenten universitären Lehre in Deutschland definiert. Dies geschah vor dem Hintergrund der Befürchtung, dass die gezielte Förderung von Spitzenforschung langfristig zu einer (weiteren) Entkopplung von Forschung und Lehre führen könnte. Im Rahmen der sogenannten Zukunftskonzepte wurde deshalb von den antragstellenden Universitäten gefordert, Programme zu entwickeln, um die in der Exzellenzinitiative geförderte Spitzenforschung in das reguläre Lehrangebot zu integrieren. Dabei sollen Studierende den aktuellen Stand der Forschung ihres Fachs, dessen Schwerpunkte und methodische Standards kennenlernen. Außerdem sollen sie die Möglichkeit bekommen, einzelne Phasen oder den gesamten Forschungsprozess selbstständig zu durchlaufen (Deutsche Forschungsgemeinschaft & Wissenschaftsrat, 2015).

¹ In der vorliegenden Arbeit wurde versucht, eine möglichst gendergerechte Sprache zu finden, indem geschlechtsunspezifische Begriffe wie „Studierende“ verwendet wurden. Damit sind grundsätzlich beide Geschlechter gemeint, wie „Studentinnen“ und „Studenten“. Sofern die Verwendung geschlechtsunspezifischer Begriffe nicht erfolgte, wurde das Binnen-I verwendet.

Als Grundlage für eine Implementation Forschungsorientierter Lehre in die universitären Curricula können Kompetenzmodelle genutzt werden: Sie bieten die Möglichkeit, Lehrpläne zu definieren, welche auf die intendierten Kompetenzen abgestimmt sind, sowie Lernumgebungen zu gestalten und entsprechende Instrumente zu entwickeln, um diese Kompetenzen als Outputs messbar zu machen (Erpenbeck & von Rosenstiel, 2007; Hetmanek, Engelmann, Opitz & Fischer, 2018; Klieme & Leutner, 2006). Diese Instrumente können sodann zur Evaluation der Wirksamkeit forschungsorientierter Lernumgebungen eingesetzt werden, um Implikationen für die Weiterentwicklung hochschuldidaktischer Formate zu ziehen. Da FoL-Programme jedoch meist nicht in einzelnen Fächern, sondern universitätsweit etabliert wurden (Deutsche Forschungsgemeinschaft & Wissenschaftsrat, 2015), bedarf es eines erhebungsökonomischen Evaluationsansatzes.

Diese Arbeit ist im Kontext eines interdisziplinären FoL-Programms entstanden. Ziel war es, ein valides Instrument zur universitätsweiten Erfassung von Forschungskompetenzen bereitzustellen, um die im Rahmen der Forschungsorientierten Lehre implementierten Lehr-Lernformate zu evaluieren. Die Basis dieses Evaluationsansatzes bildet ein neu entwickeltes, fach- und disziplinübergreifendes Kompetenzmodell zu Forschungskompetenzen (Thiel & Böttcher, 2014). Zu diesem Modell wurde ein Fragebogen zur *Erfassung* studentischer *Forschungskompetenzen* (F-Komp) entwickelt, der zur Evaluation Forschungsorientierter Lehre eingesetzt werden kann. Ziel dieser Arbeit ist zum einen die *Validierung* dieses Fragebogens. Darüber hinaus soll dieser Fragebogen eingesetzt werden, um eine neu konzipierte forschungsorientierte Lehrveranstaltung zur Förderung studentischer Forschungskompetenzen bei einer Gruppe von Lehramtsstudierenden hinsichtlich ihrer Wirksamkeit zu evaluieren.

Damit wurden die beiden Schwerpunkte dieser Arbeit identifiziert:

- Können Forschungskompetenzen mit dem neu entwickelten Fragebogen *valide erfasst* werden?
- Wie können Forschungskompetenzen bei Lehramtsstudierenden *gefördert* werden?

Diese Arbeit umfasst drei Abschnitte: Zuerst erfolgt ein theoretischer Rahmen derjenigen empirischen Studien, die im zweiten Teil dieser Arbeit vorgestellt werden. Anschließend erfolgt eine übergreifende Gesamtdiskussion der Studien. Im theoretischen Rahmen sollen fach- bzw. disziplinübergreifende Forschungskompetenzen zunächst definiert und deren Modellierung thematisiert werden, um anschließend Merkmale für deren Förderung

zu identifizieren. Auch wenn dieser Arbeit ein fach- bzw. disziplinübergreifender Ansatz für die Modellierung von Forschungskompetenzen zugrunde liegt, erfolgt eine Ausgestaltung von Lernumgebungen und deren Wirksamkeitsevaluation meist im Kontext von Fächern und Disziplinen (Hetmanek et al., 2018; Zlatkin-Troitschanskaia, Pant, Kuhn, Toepper & Lautenbach, 2016). Der theoretische Rahmen widmet sich daher auch der Frage, warum Forschungskompetenzen gerade für (angehende) Lehrkräfte relevant sind, die nach ihrem Abschluss als evidenzbasierte ExpertInnen arbeiten (J. Bauer & Prenzel, 2012; J. Bauer, Prenzel & Renkl, 2015; Borg, 2010; Davies, 1999; Hargreaves, 1996, 1997).

Die drei Studien dieser Arbeit befassen sich mit der validen Erfassung und Förderung studentischer Forschungskompetenzen. In Studie 1 *Evaluating research-oriented teaching: a new instrument to assess university students' research competences* (Kapitel 3) wird überprüft, ob sich die Dimensionen und Facetten des Kompetenzmodells (Thiel & Böttcher, 2014) auch empirisch durch den F-Komp abbilden lassen.

In Studie 2 *Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen - Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements* (Kapitel 4) werden zwei weitere Validierungsschritte des F-Komp vorgenommen: Einerseits ist es notwendig nachzuweisen, dass Forschungskompetenzen und keine allgemeinen Lern- und Studienkompetenzen erfasst werden. Zum anderen ist es notwendig, die Validität der Selbsteinschätzung von Kompetenzen zu überprüfen, die für den Fragebogen zur Erfassung studentischer Forschungskompetenzen aufgrund ihrer Möglichkeit ausgewählt wurde, Forschungskompetenzen fach- und disziplinübergreifend erfassen zu können.

In Studie 3 *How do we prepare teacher training students for Evidence-based Practice? Promoting students' research competences in Research-learning projects* (Kapitel 5) wird die Förderung studentischer Forschungskompetenzen fokussiert. Es wird eine forschungsorientierte Lehrveranstaltung auf Grundlage des neuen Kompetenzmodells entwickelt sowie der Kompetenzerwerb mit dem F-Komp evaluiert. Die Besonderheit dieser Studie ist, dass die Zielgruppe angehende Lehrkräfte sind, die, ähnlich wie Ärzte, für einen evidenzbasierten ExpertInnenberuf ausgebildet werden (J. Bauer et al., 2015). Daher liegt der Fokus der Lehrveranstaltung auf Kompetenzen für eine evidenzbasierte Praxis, welche sich aus Forschungskompetenzen ableiten lassen.

Im abschließenden Kapitel 6 werden die zentralen Befunde der einzelnen Studien diskutiert und in den theoretischen Kontext gesetzt. Darüber hinaus werden Grenzen und Stärken der Arbeit und ein Ausblick auf offene und weitere Forschungsfragen gegeben.



Kapitel 2

Theoretischer Rahmen

2 Theoretischer Rahmen

Für die empirische Erfassung und Förderung studentischer Forschungskompetenzen ist eine präzise Definition dieser Kompetenzen notwendig (Klieme & Hartig, 2008). Daher wird in den folgenden Kapiteln einerseits der Blick auf generische, akademische Kompetenzen gerichtet, um der Frage nachzugehen, wie sich Forschungskompetenzen aus diesen Kompetenzen ableiten lassen (Kapitel 2.1). Im Anschluss daran wird sowohl die Modellierung von Forschungskompetenzen anhand von fach- bzw. disziplinübergreifend gemeinsamen Phasen des Forschungsprozesses als auch deren Erfassung thematisiert (Kapitel 2.2). Nach der präzisen Definition dessen, was sich hinter Forschungskompetenzen verbirgt, werden Merkmale für die Gestaltung von Lernumgebungen zur gezielten Förderung dieser Kompetenzen sowie die Zielgruppe der Lehramtsstudierenden genauer betrachtet (Kapitel 2.3).

2.1 Ableitung von Forschungskompetenzen aus zentralen, akademischen Kompetenzen

Mit der Einführung des Qualifikationsrahmens für deutsche Hochschulabschlüsse (Kultusministerkonferenz, 2005) vollzog sich ein Wechsel von der Input- zur Outputorientierung im Hochschulsystem (Kultusministerkonferenz, 2005; Schaper, Reis, Wildt & Horvath, 2012). Während zuvor der Fokus vor allem auf den *Inhalten* eines Studiums lag (Inputorientierung), wurden nun angestrebte *Lernergebnisse* und die Beschreibung der damit einhergehenden Kompetenzen thematisiert (Outputorientierung). Sodann wurde es notwendig, den Input, also die Studieninhalte, auf die zu erzielenden Ergebnisse auszurichten. Damit rücken Kompetenzen von Studierenden als zentrale Outcomes universitärer Lehre in den Mittelpunkt.

2.1.1 Kompetenzen als studentische Outcomes

Entsprechend einer handlungstheoretischen Auffassung (Weinert, 2001) können Kompetenzen in kognitive und motivationale, volitionale und/oder soziale Outcomes ausdifferenziert werden (Nusche, 2008; Zlatkin-Troitschanskaia et al., 2016; Weinert, 2001). *Kognitive Outcomes* wiederum lassen sich in strukturelle und prozessbezogene Komponenten (in Anlehnung an Friedrich & Mandl, 1997 sowie Zlatkin-Troitschanskaia et al., 2016) bzw. in Wissen und Fertigkeiten (Nusche, 2008; Simonton, 2003) oder deklaratives und prozedurales Wissen (R. Mayer, 2003) unterteilen. Dies entspricht einer rein funktional-

pragmatischen Auffassung von Kompetenzen² (Klieme & Hartig, 2008; Klieme, Hartig & Rauch, 2008). Als strukturelle Komponenten lassen sich unterscheiden (Nusche, 2008; Zlatkin-Troitschanskaia et al., 2016): (a) *grundlegendes Wissen* zu sozialer Diversität, Politik, nationalen und internationalen Themen sowie (b) *fach- bzw. domänenspezifisches Wissen*, das z.B. im Fach Mathematik oder Geschichte erworben wird. Als prozessbezogene Komponenten lassen sich unterscheiden (Nusche, 2008; Zlatkin-Troitschanskaia et al., 2016): (c) *generische Fertigkeiten*, die über Disziplinen hinausgehen und transferierbar sind wie Problemlösen oder wissenschaftliches Denken und Argumentieren sowie (d) *fach- bzw. domänenspezifische Fertigkeiten*, wie Experimentieren in der Biologie, die nicht in andere Bereiche, wie z.B. Quellenarbeit in den Geschichtswissenschaften, transferierbar sind. Kognitive Kompetenzen als Ergebnis von Lern- und Entwicklungsprozessen (Erpenbeck & von Rosenstiel, 2007; Klieme et al., 2008) lassen sich jedoch abgrenzen von kontext- und domänenunabhängigen Fähigkeiten im Sinne fluider Intelligenz (Cattell, 1987; Weinert, 2001), die zeitstabil sind und weder Vorwissen noch Erfahrung benötigen (Klieme & Hartig, 2008; Weinert, 2001). Darüber hinaus lassen sich kognitive Kompetenzen abgrenzen von nicht-kognitiven Outcomes (Nusche, 2008; Zlatkin-Troitschanskaia et al., 2016), wie *motivationale, volitionale und/oder soziale Outcomes* (Weinert, 2001), welche Beliefs, volitionale Kontrollsysteme, Werte und die psychosoziale Entwicklung beispielsweise der Identität bzw. des Selbstvertrauens umfassen. Im Qualifikationsrahmen für deutsche Hochschulabschlüsse (Kultusministerkonferenz, 2005) liegt das Augenmerk neben fachlichen Kompetenzen insbesondere auf überfachlichen und damit generischen Kompetenzen.

2.1.2 Generische Kompetenzen universitärer Lehre

Universitäten sollen zwei zentralen Ausbildungsaufgaben nachkommen (Enders, 2010; Hochschulrahmengesetz, 1999; Wissenschaftsrat, 2006): Sie sollen einerseits für die Ausbildung des wissenschaftlichen Nachwuchses sorgen, indem sie Studierende für eine wissenschaftliche Karriere ausbilden. Andererseits sollen Universitäten aber auch eine praxis- und berufsorientierte Ausbildung ermöglichen, indem AbsolventInnen auf berufliche Tätigkeiten vorbereitet werden sollen, in denen *wissenschaftliche Erkenntnisse und Methoden* außerhalb der Forschung *angewendet* werden: Sie werden zu evidenzbasierten ExpertInnen. Damit nimmt neben der Vermittlung von fachlichen Kompetenzen insbesondere die

² Eine ausführlichere Unterscheidung des Kompetenzbegriffes nach Weinert (2001) sowie Klieme und KollegInnen (2008) wird in Kapitel 3.1 thematisiert.

Vermittlung fachübergreifender bzw. generischer Kompetenzen einen besonderen Stellenwert ein (Jaksztat, Schindler & Briedis, 2010; Kultusministerkonferenz, 2005; Wissenschaftsrat, 2008). Im Qualifikationsrahmen für Deutsche Hochschulabschlüsse (Kultusministerkonferenz, 2005) werden als generische Kompetenzen Methodenkompetenzen, soziale und kommunikative Kompetenzen ausdifferenziert. Soziale und kommunikative Kompetenzen beschreiben die Fähigkeiten, anderen das im Fach erworbene Wissen zu vermitteln und sich mit ihnen darüber auszutauschen sowie in einem Team arbeiten zu können und dort Verantwortung zu übernehmen. Methodenkompetenzen umfassen diejenigen Fähigkeiten, das im Fach erworbene Wissen anzuwenden und auf neue und unbekannte Situationen übertragen zu können. Über den Qualifikationsrahmen für Deutsche Hochschulabschlüsse hinaus wird unter Methodenkompetenzen auch die selbstorganisierte Gestaltung eigener Tätigkeiten und Aufgaben sowie eine kreative Herangehensweise beim Problemlösen verstanden (Erpenbeck & von Rosenstiel, 2007). Diese Problemlösekompetenz und die Selbstregulierung des eigenen Handelns und Lernens werden sowohl von der Kultusministerkonferenz (2005) als auch vom Wissenschaftsrat (2008) als zentrale überfachliche Kompetenzen eingeordnet, welche sowohl für den wissenschaftlichen Kontext als auch für den öffentlichen Sektor oder die Wirtschaft als wichtig erachtet werden.

Das *selbstregulierte Lernen* ist eine zentrale akademische Kompetenz, da AbsolventInnen während ihres Studiums gelernt haben sollten, sich umfangreiche Wissensbestände eigenständig zu erarbeiten und ihr eigenes Lernen autonom und sachangemessen zu steuern (Steuer et al., 2015). Selbstreguliertes Lernen besteht aus kognitiven und motivationalen Komponenten, welche sich wiederum unterteilen lassen in prozessbezogene und strukturelle Komponenten (Friedrich & Mandl, 1997). Zu *prozessbezogenen kognitiven Komponenten* zählen kognitive Strategien zur Verarbeitung von Informationen und zur Organisation des eigenen Lernens, metakognitive Strategien zur Kontrolle und Überwachung des eigenen Lernens sowie ressourcenbezogene Strategien, beispielsweise zur Regulation von Emotionen (Friedrich & Mandl, 1997; Steuer et al., 2015; Wild & Schiefele, 1994). *Strukturell kognitive Komponenten* bestehen aus Wissensbestandteilen, welche für den Erwerb neuen Wissens relevant sind. Dazu gehören Vorwissen (deklaratives Wissen), Handlungswissen, um die unter den prozessbezogenen kognitiven Komponenten aufgeführten Strategien einsetzen zu können (prozedurales Wissen), sowie Wissen über die Eignung verschiedener Strategien (konditionales Wissen; Friedrich & Mandl, 1997; Steuer et al., 2015; Wild & Schiefele, 1994). *Prozessbezogene motivationale*

Komponenten umfassen Strategien zur Bewältigung von Anforderungen, wie Kontrollüberzeugungen oder Anstrengungsattributionen, Volitionsstrategien zur Abschirmung der Motivation gegen konkurrierende Handlungen sowie Emotionen wie Freude, Langeweile oder Angst (Friedrich & Mandl, 1997). Zu strukturell motivationalen Komponenten zählen Bedürfnisse nach Kompetenz und Autonomie, eigene Interessen und Zielsetzungen oder Selbstwirksamkeitserwartungen (Deci & Ryan, 1993; Friedrich & Mandl, 1997). Darüber hinaus wird die selbstregulierte Steuerung des eigenen Handelns und wichtiger Entscheidungen nicht nur als Outcome, sondern auch als Voraussetzung für ein erfolgreiches Studium angesehen (Steuer et al., 2015). Da wissenschaftliches Wissen als unsicher bezeichnet werden kann (Hofer & Pintrich, 1997; Schommer, 1998; Trautwein & Lüdtke, 2007) und im wissenschaftlichen Kontext oftmals dynamische Probleme vorliegen, deren Ziele und Lösungswege nicht immer vordefiniert sind (Greiff, Wüstenberg & Funke, 2012; Leutner, Fleischer, Wirth, Greiff & Funke, 2012), ist die Selbstregulation von besonderer Relevanz.

Wissenschaftliches Denken stellt eine Spezialform des Problemlösens dar (Fischer et al., 2014; J. Mayer, 2007). Die zentralen Phasen des Problemlösens – *Problemidentifikation*, *Definition des Ist- und Sollzustandes*, *Planerstellung*, *Planausführung* und *Ergebnisbewertung* (Funke, 2003) – lassen sich auf das wissenschaftliche Denken und Argumentieren für den schulischen und hochschulischen Bereich übertragen (Fischer et al., 2014): (1) *Problemidentifikation*, ein wissenschaftliches oder praktisches Problem wird als solches erkannt. Es kann sich um Wissenslücken in verschiedenen Bereichen handeln, wie fehlendes Wissen über die Grundlagen von Phänomenen oder beobachtbare Fakten (Grundlagenforschung), fehlendes Wissen, um eine Zielsetzung in der Praxis zu erreichen (angewandte Forschung), oder fehlende Produkte oder Umsetzungsmöglichkeiten der aus Forschung und Praxis gewonnenen Erfahrungen (experimentelle Forschung; OECD, 2015; Stokes, 1997). Es wird festgestellt, dass die Lösung des Problems nicht unmittelbar erfolgen kann. Danach folgt die Definition des Ist-Zustandes durch die Recherche geeigneter Evidenz, was (2) in das *Stellen von Fragen* oder das Generieren von Hypothesen mündet, die als *möglicher Soll-Zustand* überprüft werden können. Anschließend wird ein Plan erstellt, zu dem (4) die *Konstruktion und Neugestaltung von Objekten* gehört, um auf Grundlage von bestehendem Wissen das Zielkonstrukt in der realen Welt erfassen zu können (z.B. in Form von Fragebögen, Prototypen von Maschinen, Instrumenten in Laboren). Im Zuge der Planausführung wird (5) *Evidenz generiert*, entweder auf induktivem oder deduktivem Weg.

Es folgen (6) die *Evaluation der Evidenz*, um zu prüfen, inwiefern die eigenen Annahmen gestützt werden, und (7) das *Ziehen von Schlussfolgerungen*, wozu wiederum verschiedene Evidenz integriert und kombiniert werden muss. Dies entspricht der Ergebnisbewertung. Am Ende steht (8) die *Kommunikation* und Überprüfung der Begründung und ihrer Ergebnisse. Beim wissenschaftlichen Denken und Argumentieren handelt es sich um einen *dynamischen Problemlöseprozess*, da beim wissenschaftlichen Denken die für die Lösung relevanten Informationen, in diesem Fall Evidenz, erst *während* des Prozesses generiert werden (Greiff et al., 2012).

Wissenschaftliches Denken wird zu den überfachlichen bzw. generischen Kompetenzen gezählt, die AbsolventInnen im Rahmen ihres Studiums erlangen sollen. Uneinigkeit besteht jedoch darüber, inwiefern wissenschaftliches Denken und Argumentieren domänenübergreifende oder domänenspezifische Kompetenzen darstellen (Fischer, Chinn, Engelmann & Osborne, 2018). Einige AutorInnen (Goldman, Ko, Greenleaf & Brown, 2018; Osborne, 2018; van Boxtel & van Drie, 2018) vertreten die Auffassung, dass wissenschaftliches Denken und Argumentieren überwiegend domänenspezifisch ist: Die erforderlichen Kompetenzen, um die Phasen des wissenschaftlichen Denkens und Argumentierens zu durchlaufen, werden in einem bestimmten Kontext erlangt und angewendet (z.B. in bestimmten Fächern oder Disziplinen). Diese Kompetenzen sind so fachspezifisch, dass sie sich nicht auf andere Kontexte übertragen lassen, z.B. die Phasen des wissenschaftlichen Denkens und Argumentierens in den Naturwissenschaften (Osborne, 2018). Allerdings gibt Renkl (2018) in einer Diskussion dieser fachspezifischen Ansätze zu bedenken, dass sich diese fachspezifisch ausbuchstabilten Prozesse zum Teil viel eher über Fächer hinweg gleichen als innerhalb eines Fachs. Übertragen auf ein Beispiel bedeutet das, dass Korrelationsanalysen sowohl in der Biologie als auch in den Erziehungswissenschaften im Rahmen der Datenanalyse angewendet werden; innerhalb der Erziehungswissenschaften werden allerdings nicht in allen Forschungsbereichen Korrelationsanalysen angewendet, z.B. nicht im Rahmen qualitativer Forschungsmethoden.

Hetmanek und KollegInnen (2018) gehen beim wissenschaftlichen Denken und Argumentieren von *cross-domain-applicable skills* aus, also von Fertigkeiten, die über verschiedene Domänen hinweg anwendbar sind. Dafür sollte die *Domäne* klar definiert sein. Als Domäne können unterschiedliche abstrakte oder konkrete Einheiten gelten: Disziplinen (z.B. Naturwissenschaften), Fächer (z.B. Biologie) oder spezifische Themen innerhalb eines Fachs (z.B. die Fruchtfliege als biologisches Objekt; vgl. Hetmanek et al., 2018). Als

domänenübergreifend werden Fähigkeiten dann bezeichnet, wenn sie in mehr als einer Domäne angewendet werden können. Zwar werden diese Fertigkeiten in einem bestimmten Fachkontext erworben (z.B. im Fach Biologie), allerdings können diese Fertigkeiten zur Lösung strukturell ähnlicher Probleme in anderen Domänen angewendet werden (Hetmanek et al., 2018). Dazu können ganz allgemein das Anwenden von Lernstrategien zählen oder, in Anlehnung an das zuvor dargestellte Beispiel, das Generieren von Evidenz mittels Korrelationsanalysen in unterschiedlichen Fächern. Diese *domänenübergreifenden* Kompetenzen lassen sich, so Hetmanek und KollegInnen (2018), abgrenzen von zeitstabilen, kontext- und domänenunabhängigen Formen des Problemlösens, die ohne Vorwissen und Erfahrung erfolgen können und Aspekte von fluider Intelligenz darstellen (Cattell, 1987; Klieme & Hartig, 2008; Weinert, 2001). Denn für einen Problemlöseprozess werden, ähnlich wie beim selbstregulierten Lernen, deklaratives, prozedurales und konditionales Wissen benötigt (Fischer et al., 2014; Leutner et al., 2012). Entsprechend dieser Auffassung lässt sich ein hierarchisches Modell der Anwendbarkeit von Fertigkeiten darstellen (Abbildung 2.1), in dem Hierarchie allerdings nicht gleichbedeutend mit Wichtigkeit ist (Hetmanek et al., 2018).

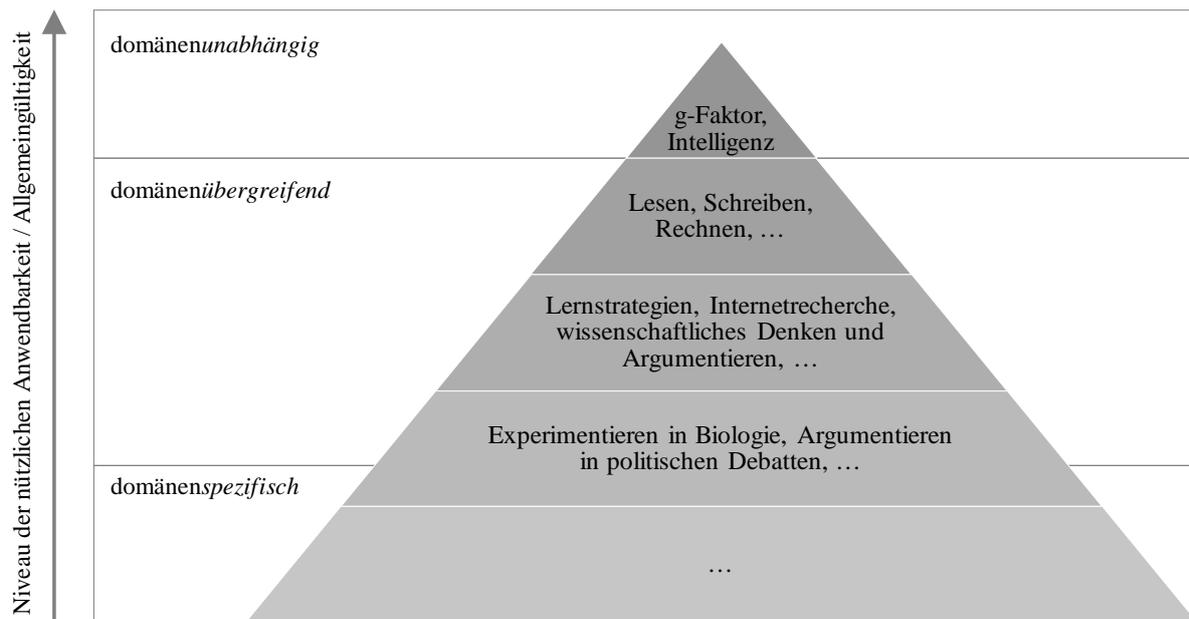


Abbildung 2.1 Übersetzung aus dem Englischen und Adaption³ der Grafik von Hetmanek et al. (2018, S. 210⁴) zur Hierarchie von kognitiven Fertigkeiten

Wissenschaftliches Denken und Argumentieren und die Phasen des Forschungsprozesses

Die dargestellten Phasen des wissenschaftlichen Denkens und Argumentierens zeigen große Überschneidungen zu den zentralen Phasen des Forschungsprozesses: Problemidentifikation und Fragenstellen als Recherche des Forschungsstandes; Hypothesengenerierung, Erfassung von Zielkonstrukten und Generierung von Evidenz als Anwendung von Forschungsmethoden; Evaluation von Evidenz und Ziehen von Schlussfolgerungen als Reflexion von Ergebnissen sowie die Kommunikation über Forschung (W. Bauer, Bleck-Neuhaus, Dombois & Wehrmann, 2013; Döring & Bortz, 2016; L. Huber, 2014; Schneider & Wildt, 2009; Shavelson & Towne, 2002). Für die AbsolventInnen von Universitäten, und zwar sowohl für den wissenschaftlichen Nachwuchs als auch für evidenzbasierte ExpertInnen, sind spezifische Kompetenzen für den Forschungsprozess notwendig, welche im Rahmen Forschungsorientierter Lehre erworben werden können.

³ Die Abbildung wurde insofern adaptiert, als dass „domänenunabhängig“ als Übersetzung für „general“ gewählt wurde, um die dargestellten Niveaus im Deutschen bestmöglich voneinander abzugrenzen.

⁴ Copyright 2018 © from “Cognitive Perspectives in Scientific Reasoning and Argumentation: Domain-Specific and Domain-General Interplay” by F. Fischer, C. A. Chinn, K. Engelmann, and J. Osborne. Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc This permission does not cover any third party copyrighted work which may appear in the material requested. Please check the figure caption or acknowledgements section of the book. Licence-Number 4647631426279.

Nachfolgend soll näher betrachtet werden, wie diese Forschungskompetenzen modelliert und gemessen werden können.

2.2 Modellierung und Erfassung von Forschungskompetenzen

Kompetenzmodelle können die Basis für die Erfassung von Kompetenzen bilden: Kompetenzen können darin theoriegeleitet beschrieben und systematisiert werden (Klieme & Leutner, 2006). Darüber hinaus kann anhand eines Kompetenzmodells die empirische Beobachtung von Kompetenzen durch die Operationalisierung in geeignete Messverfahren vorgenommen werden (Erpenbeck & von Rosenstiel, 2007; Klieme & Hartig, 2008). Für Forschungskompetenzen lassen sich neben fachlichen auch fach- bzw. disziplinübergreifende Merkmale für deren Modellierung und Erfassung identifizieren.

2.2.1 Modellierung von fach- bzw. disziplinunabhängigen Forschungskompetenzen

Forschungskompetenzen werden benötigt, um die spezifischen Anforderungen eines Forschungsprozesses absolvieren zu können. Diese Anforderungen sind fach- bzw. disziplinspezifisch: Die Natur-, Geistes- und Sozialwissenschaften und deren einzelne Fächer haben beispielsweise spezifische Forschungsmethoden (Döring & Bortz, 2016 für Sozial- und Humanwissenschaften; Beckert et al., 2015 für Chemie; Hartmann, Mohseni, Reckwitz, Rojek & Steckmann, 2012 für Geisteswissenschaften), eigene Fachzeitschriften (z.B. *Nature* in den Naturwissenschaften, *Learning and Instruction* für Bildungswissenschaften, *Historical Journal* für Geschichtswissenschaften) und Datenbanken (z.B. SciFinder für Naturwissenschaften, ERIC für Bildungswissenschaften, Web of Science / Arts & Humanities Citation Index für Geisteswissenschaften). So ist eine fach- bzw. disziplinspezifische Modellierung von Forschungskompetenzen zweifellos sinnvoll. Werden allerdings universitätsweite Programme wie Forschungsorientierte Lehre implementiert, so ist ein fach- bzw. disziplinübergreifender Ansatz der Modellierung förderlich, um auf dessen Basis eine erhebungsökonomische, universitätsweite Evaluation über Fächergrenzen hinweg durchführen zu können. Die Identifikation von gemeinsamen Aspekten des Forschungsprozesses über verschiedene Disziplinen hinweg bietet daher die Möglichkeit einer hochschulübergreifenden und damit auch einer fach- bzw. disziplinunabhängigen Betrachtung von Forschungskompetenzen.

Bei Forschung handelt es sich um einen systematischen, methodisch kontrollierten, nachvollziehbaren, zyklischen und zielgerichteten Erkenntnisprozess (Butts, 1991;

Mittelstraß, 1991; Shavelson & Towne, 2002). Dieser besteht in Anlehnung an die pragmatische Wissenschaftstheorie (Butts, 1991; Dewey, 1938; Mittelstraß, 1991) über Disziplingrenzen hinaus aus strukturell ähnlichen Phasen: (1) der Recherche des Forschungsstandes, da Forschung auf dem aktuellen Stand der Forschung basiert; (2) der Anwendung von Forschungsmethoden, da neues Wissen aufgrund der Beschreibung von Gegenständen der sozialen, natürlichen und kulturellen Umwelt generiert wird; (3) der Reflexion von Ergebnissen, da Schlussfolgerungen ausschließlich aufgrund dieser neu generierten und der bereits verfügbaren Evidenz gezogen werden; (4) der Kommunikation über diese Ergebnisse, da der Forschungsprozess festgehalten und transparent dokumentiert wird (Butts, 1991; Dewey, 1938; Mittelstraß, 1991). Werden Fächer und Disziplinen als Domäne definiert (vgl. Kapitel 2.1.2; Hetmanek et al., 2018), so können diese Phasen als *domänenübergreifend* aufgefasst werden. Dennoch sind diese Phasen spezifisch für den Kontext *Forschung* ausbuchstabiert, sodass die Phasen des Forschungsprozesses als Spezialfall des wissenschaftlichen Denkens und Argumentierens und damit ebenfalls als zu bewältigender, dynamischer Problemlöseprozess aufgefasst werden können. Abbildung 2.2 ist ausgehend von Hetmanek und KollegInnen (2018) so adaptiert worden, dass sich Forschungskompetenzen als *domänenübergreifende* Kompetenzen in diesem hierarchischen Modell verorten lassen.

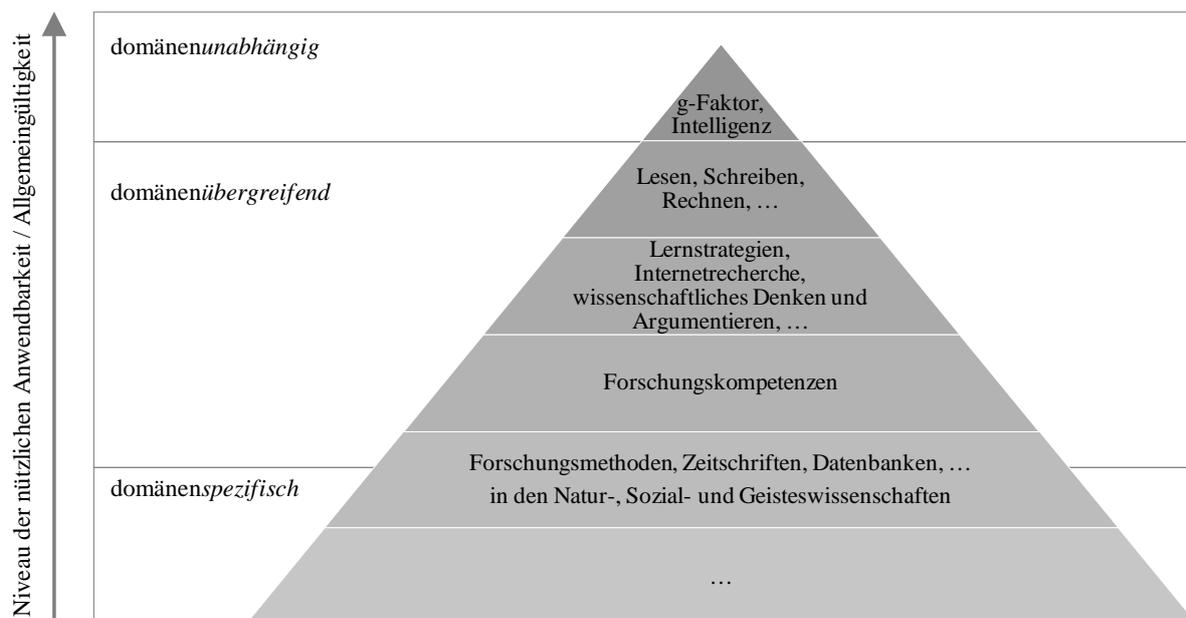


Abbildung 2.2 Domänenübergreifende Forschungscompetenzen als Spezialfall des wissenschaftlichen Denkens und Argumentierens (Adaption der übersetzten Abbildung von Hetmanek et al., 2018, S. 210⁵; vgl. Abbildung 2.1)

Bislang existieren einige Ansätze zur Modellierung von Forschungscompetenzen, in denen sich die vier aus der pragmatischen Wissenschaftstheorie abgeleiteten Phasen des Forschungsprozesses wiederfinden lassen. Diese Modelle lassen sich unterscheiden in Modelle zu Teilaspekten des Forschungsprozesses, fach- oder domänenspezifische Modelle und generische Modelle. Im Bereich der Modelle zu Teilaspekten des Forschungsprozesses sind beispielsweise das Kompetenzstrukturmodell zu *Information Literacy* (Trempler et al., 2015) oder zum wissenschaftlichen Schreiben (Winter-Hölzl, Wäschle, Wittwer, Watermann & Nückles, 2015; Winter-Hölzl, Watermann, Wittwer & Nückles, 2016) zu nennen. Im Bereich der fach- oder domänenspezifischen Modelle lassen sich dazu beispielhaft Kompetenzstrukturmodelle für Bildungswissenschaften (Groß Ophoff, Schladitz, Lohrmann & Wirtz, 2014; Groß Ophoff, Wolf, Schladitz & Wirtz, 2017), das Lehramt (Mandinach & Gummer, 2016), Sozialwissenschaften (Gess, Wessels & Blömeke, 2017; Gess, Geiger & Ziegler, 2018), Ingenieurwissenschaft (Meijers, van Overveld, C. W. A. M & Perrenet, 2005) oder Lebenswissenschaften (Holbrook & Devonshire, 2005; Valter & Akerlind, 2010) nennen.

⁵ Copyright 2018 © from “Cognitive Perspectives in Scientific Reasoning and Argumentation: Domain-Specific and Domain-General Interplay” by F. Fischer, C. A. Chinn, K. Engelmann, and J. Osborne. Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc This permission does not cover any third party copyrighted work which may appear in the material requested. Please check the figure caption or acknowledgements section of the book. Licence-Number 4647631426279.

Das RSD-Framework (Willison & O'Regan, 2007) repräsentiert ein fachübergreifendes Kompetenzniveauomodell. Den bislang publizierten Kompetenzmodellen fehlt es allerdings entweder an einer ganzheitlichen Betrachtung des Forschungsprozesses (Modelle zu Teilaspekten), an einer Übertragbarkeit auf andere Fächer und Disziplinen ohne weitere Adaption (fach- oder domänenspezifische Modelle), oder an theoretisch trennscharfen Dimensionen (RSD-Framework von Willison & O'Regan, 2007; vgl. Thiel & Böttcher, 2014). Dennoch lassen sich dort Hinweise für eine Operationalisierung der Phasen des Forschungsprozesses finden: Durchführung von Literaturrecherchen, Anwendung von Forschungsmethoden, Interpretation und Reflexion von Ergebnissen sowie Kommunikation über Forschung.

Durchführung von Literaturrecherchen

Am Anfang des Forschungsprozesses steht der theoretische oder praktische Rahmen oder ein konkretes Ausgangsproblem, auf dem sich der gesamte Forschungsprozess gründet (W. Bauer et al., 2013; Döring & Bortz, 2016; Shavelson & Towne, 2002). Es folgt die Phase des Recherchierens geeigneter Evidenz, bezeichnet als *Information Literacy* (Groß Ophoff et al., 2014, 2017; Trempler et al., 2015). Der Begriff *Information Literacy* verweist auf zwei Teilphasen des Forschungsprozesses: die Suche und Auswahl geeigneter Evidenz sowie die Bewertung und Nutzung dieser Evidenz (Trempler et al., 2015). In der ersten Phase wird aussagekräftige Evidenz, meist in Datenbanken oder Suchmaschinen, gesucht und ausgewählt. Dazu werden Schlagwörter oder Abstracts überflogen und weitere relevante Informationen wie die Plausibilität und Qualität der Evidenz berücksichtigt. Die zweite Phase umfasst die Bewertung der gefundenen Evidenz. Dazu sind sowohl eine tiefe Verarbeitung in Form von Lesen und Verstehen als auch eine systematische Beurteilung einzelner Aspekte von wissenschaftlichen Arbeiten notwendig. Darüber hinaus ist es während der Recherche oder im Anschluss daran notwendig, aus dem identifizierten Forschungsstand heraus Forschungsfragen zu identifizieren und zu generieren (Gess et al., 2017, 2018; Groß Ophoff et al., 2014, 2017; Mandinach & Gummer, 2016; Trempler et al., 2015; Willison & O'Regan, 2007).

Anwendung von Forschungsmethoden

Die Bearbeitung der Forschungsfrage erfordert zunächst ihre Operationalisierung in Teilfragen bzw. die Generierung von Hypothesen (Holbrook & Devonshire, 2005; Valter & Akerlind, 2010). Weiterhin werden methodengeleitet systematisch neue Erkenntnisse generiert (Meijers et al., 2005; Willison & O'Regan, 2007). Dies umfasst, Studiendesigns zu

entwickeln und umzusetzen (Gess et al., 2017, 2018; Holbrook & Devonshire, 2005; Valter & Akerlind, 2010), Daten zu erfassen und zu analysieren (Gess et al., 2017, 2018; Holbrook & Devonshire, 2005; Mandinach & Gummer, 2016; Valter & Akerlind, 2010; Willison & O'Regan, 2007).

Interpretation und Reflexion der Ergebnisse

Im Anschluss an die Erhebung und Analyse der Daten werden die erzielten Erkenntnisse kontextualisiert, indem die erzeugten Daten interpretiert und diskutiert (Gess et al., 2017, 2018; Groß Ophoff et al., 2014, 2017; Holbrook & Devonshire, 2005; Valter & Akerlind, 2010) sowie Schlussfolgerungen daraus gezogen werden. Diese Schlussfolgerungen können sich auf den Erkenntnisgewinn in Bezug zum Forschungsstand und auf den Forschungsprozess selbst (Mandinach & Gummer, 2016; Willison & O'Regan, 2007), auf die praktische Anwendbarkeit der Forschungsergebnisse (Groß Ophoff et al., 2014, 2017; Mandinach & Gummer, 2016; Meijers et al., 2005; Willison & O'Regan, 2007) oder auf gesellschaftliche oder ethische Konsequenzen (Meijers et al., 2005) beziehen.

Kommunikation über Forschung

Am Ende des Forschungsprozesses steht die Kommunikation der Ergebnisse mit der *Scientific Community* (W. Bauer et al., 2013; Döring & Bortz, 2016; Shavelson & Towne, 2002). Dazu zählt das Verfassen von Publikationen, das Halten von Präsentationen oder das Vorführen der Ergebnisse des Forschungsprozesses bzw. des generierten Wissens (Meijers et al., 2005). Für das Verfassen von Publikationen werden wissenschaftliche Schreibkompetenzen benötigt (Decker, Kaplan & Siebert-Ott, 2015; Winter-Hözl et al., 2015, 2016; Wolf, Zahner, Kostoris & Benjamin, 2014). Dabei nutzen verschiedene Disziplinen unterschiedliche Textsorten, für die jeweils spezifische Regeln der Erstellung gelten. Dennoch lassen sich verschiedene zyklische Prozesse des wissenschaftlichen Schreibens identifizieren (Winter-Hözl et al., 2015): *Planung* – relevante Informationen werden aus dem Langzeitgedächtnis erzeugt und organisiert und konkrete Ziele werden definiert; *Übersetzung* – das Inhaltswissen wird in geschriebene Sprache umgewandelt; *Überprüfung* – das Geschriebene wird überarbeitet, um die Qualität des Textes zu verbessern.

Da Kompetenzen laut funktional-pragmatischer Auffassung (vgl. Kapitel 2.1.1) jedoch nicht nur aus einer Fertigkeitkomponente, sondern auch aus einer Wissenskomponente bestehen, finden sich in einigen wenigen Modellen auch Wissensaspekte wieder.

Wissen über Forschung

Wissen bezieht sich in den entsprechenden Modellen auf das Wissen des jeweiligen Fachs oder der Domäne (Meijers et al., 2005), kann aber noch weiter ausdifferenziert werden in Wissen über den Forschungsprozess und seine einzelnen Phasen, z.B. Wissen über geeignete Recherchestrategien, Wissen zu Forschungsmethoden, methodologisches Wissen oder Wissen zu Textsorten und -bestandteilen (Gess et al., 2017, 2018; Winter-Hözl et al., 2015, 2016).

Die Vielzahl an fach- bzw. domänenspezifischen Modellen (Gess et al., 2017, 2018; Groß Ophoff et al., 2014, 2017; Holbrook & Devonshire, 2005; Mandinach & Gummer, 2016; Meijers et al., 2005; Valter & Akerlind, 2010) eignet sich nicht für eine fach- bzw. disziplinübergreifende Modellierung von Forschungskompetenzen. Dahingegen bieten generische Modelle, sofern sie trennscharfe Dimensionen aufweisen, die Möglichkeit, Kompetenzen fachübergreifend, entsprechend einer theoretischen Basis zu modellieren (vgl. Thiel & Böttcher, 2014).

Das RMRK-W-Modell

Das RMRK-W-Modell (Thiel & Böttcher, 2014) wurde mit dem Ziel entwickelt, die oben dargestellte Forschungslücke zu schließen. Es modelliert Kompetenzen in einem Strukturmodell für den Kontext Forschung entlang des Forschungsprozesses, allerdings fach- und disziplinübergreifend und mit trennscharfen Dimensionen: *Recherchekompetenzen*, *Methodenkompetenzen*, *Kommunikationskompetenzen*⁶, *Reflexionskompetenzen* sowie einer Dimension *Fachliches Wissen*. In seinen Facetten finden sich viele der oben dargestellten, aus anderen Modellen stammenden Kompetenzbereiche wieder (Abbildung 2.3). Eine ausführliche Beschreibung des Modells, seiner Dimensionen und Facetten wird in den Kapiteln 3 und 4 thematisiert.

⁶ Die im RMRK-W-Modell mit Methoden- und Kommunikationskompetenzen bezeichneten Dimensionen stellen insofern Spezialfälle der gleichnamigen Kompetenzen im Qualifikationsrahmen für deutsche Hochschulabschlüsse (Kultusministerkonferenz, 2005) dar, als dass sie explizit diejenigen Prozesse abbilden, die für das Absolvieren dieser *Forschungsphasen* notwendig sind.

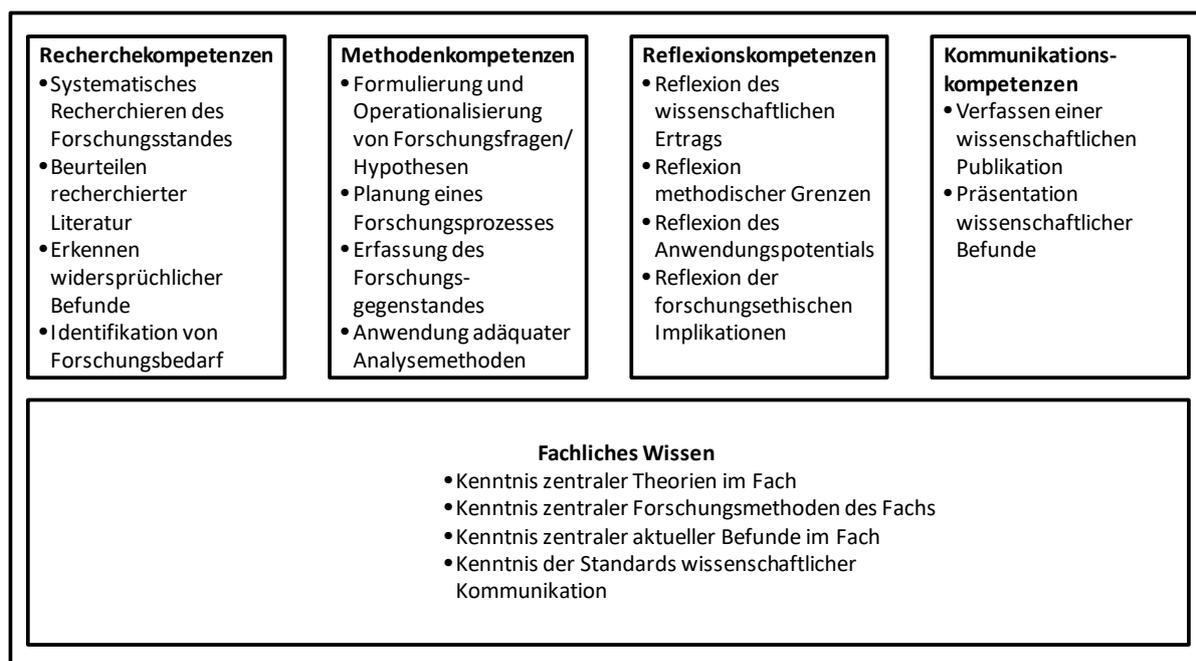


Abbildung 2.3 Das RMRK-W-Modell

Inwieweit Kompetenzmodelle jedoch nicht nur eine theoretische Konzeption, sondern auch ein empirisch messbares Konstrukt darstellen, muss empirisch überprüft werden (Erpenbeck & von Rosenstiel, 2007; Klieme & Hartig, 2008). Darüber hinaus muss für den Zweck der Evaluation von FoL-Programmen ein Messverfahren ausgewählt werden, das sich für die fach- bzw. disziplinübergreifende Erfassung von Forschungskompetenzen eignet.

2.2.2 Messverfahren zur fächerübergreifenden Erfassung von Forschungskompetenzen

Es existieren bereits verschiedene Verfahren zur Erfassung von Forschungskompetenzen. Allerdings eignen sich nicht alle Verfahren gleichermaßen für eine fach- bzw. disziplinübergreifende Erfassung von Forschungskompetenzen im Rahmen der universitätsweiten Evaluation Forschungsorientierter Lehre. Hier ist die Erhebungsökonomie ein wichtiges zu berücksichtigendes Kriterium. Ein Messverfahren ist ökonomisch, wenn wenig Zeit für dessen Durchführung benötigt wird, kein hoher Materialverbrauch anfällt, es bei mehreren Personen gleichzeitig einsetzbar ist sowie effizient und effektiv auszuwerten ist (Döring & Bortz, 2016). Für die fach- bzw. disziplinunabhängige Messung von Forschungskompetenzen ist es zudem sinnvoll, dass ein Verfahren mit möglichst wenig Adaption in unterschiedlichen Fächern oder Disziplinen eingesetzt werden kann.

Im Gegensatz zu qualitativen Ansätzen zur Erfassung von Forschungskompetenzen (z.B. Perrenet & Adan, 2010; Valter & Akerlind, 2010) sind standardisierte, quantitative Verfahren weitaus ökonomischer. Kompetenztests (z.B. Gess et al., 2017, 2018; Groß Ophoff et al., 2014, 2017) oder Selbsteinschätzungen (z.B. Dunn, Airola, Lo & Garrison, 2013; Willison, 2012) haben eine höhere Erhebungsökonomie, weil in jeder Situation, auch bei größeren Gruppen, derselbe Erhebungsablauf gewährleistet werden kann und sie effektiv und effizient ausgewertet werden können (Bach, 2013; Döring & Bortz, 2016; Frey, 2006). So gibt es beispielsweise einen festen Stamm an geschlossen formulierten Items oder Aufgaben, die es zu bearbeiten gilt bzw. geschulte TestleiterInnen, die für einen standardisierten Ablauf sorgen (Döring & Bortz, 2016). Kompetenztests bieten zwar einen objektiven Zugang zur Kompetenzmessung, indem sie Kompetenzen „von außen“ anhand proximaler Indikatoren erfassen (Bach, 2013; Erpenbeck & von Rosenstiel, 2007). Allerdings eignen sie sich weniger für eine fach- oder disziplinunabhängige Messung, da die Konstruktion von Aufgaben den Bezug auf eine Domäne voraussetzt (Zlatkin-Troitschanskaia et al., 2016). Daher wären diese Aufgaben ohne Adaption somit kaum auf andere Kontexte übertragbar. Bei Selbsteinschätzungen hingegen nimmt die untersuchte Person selbst eine Einschätzung ihrer Kompetenz aufgrund von selbstbezogenen Kognitionen vor (Bach, 2013). Standardisierte Selbsteinschätzungsskalen sind daher besonders erhebungsökonomisch (Bach, 2013). Damit ist eine Übertragbarkeit auf und Vergleichbarkeit mit anderen Kontexten prinzipiell möglich, sofern die Items möglichst allgemein formuliert sind (z.B. „Ich kenne die wichtigsten Theorien der sozialpädagogischen Forschung“ vs. „Ich kenne die wichtigsten Theorien in meinem Fach“).

Um dem Auftrag nach einem validen Instrument zur universitätsweiten Erfassung von Forschungskompetenzen nachzukommen, wurde der Fragebogen zur Erfassung studentischer Forschungskompetenzen (F-Komp) als Selbsteinschätzungsinstrument entwickelt. Der Fragebogen umfasst 32 Items in fünf Skalen entsprechend den Dimensionen des Modells: Recherche-, Methoden-, Reflexions- und Kommunikationskompetenzen sowie Fachliches Wissen⁷. Die Validierung eines neuen Instrumentes kann im Rahmen einer Inhalts-, Konstrukt- und Kriteriumsvalidierung vorgenommen werden (Bach, 2013; Döring & Bortz, 2016; Messick, 1995): Dabei können (a) durch eine theoretisch-argumentative Beurteilung eines Instrumentes durch ExpertInnen die inhaltliche Passung (Inhaltsvalidität), (b) die

⁷ Eine ausführliche Darstellung des Instrumentes sowie seiner Konstruktion erfolgt in Kapitel 3.1.

Dimensionalität des Messinstrumentes (faktorielle Validität im Rahmen der Konstruktvalidierung), (c) begründete Zusammenhänge zu anderen theoretischen Konstrukten (konvergente und diskriminante Validität im Rahmen der Konstruktvalidierung) und (d) die Übereinstimmung zu distalen Indikatoren (retrospektive, konkurrenente und prognostische Validität im Rahmen der Kriteriumsvalidierung) überprüft werden.

Nachdem die Frage der Modellierung und Erfassung von Forschungskompetenzen thematisiert wurde, kann nun deren Förderung genauer betrachtet werden (Klieme & Hartig, 2008).

2.3 Förderung von Forschungskompetenzen

Kompetenzmodelle können die Grundlage für die Förderung von Kompetenzen bilden, indem ausgehend von ihnen Kompetenzziele identifiziert und Curricula zur Förderung dieser Kompetenzen entwickelt werden (Hetmanek et al., 2018; Klieme & Hartig, 2008; Klieme & Leutner, 2006; Erpenbeck & von Rosenstiel, 2007). Darüber hinaus lassen sich gewisse Merkmale identifizieren, die den Erwerb von Forschungskompetenzen begünstigen.

2.3.1 Lernumgebungen zur Förderung des Erwerbs von Forschungskompetenzen

Lernumgebungen zum Erwerb von Kompetenzen sollten Merkmale berücksichtigen, die sich auf die Motivierung, auf kognitive Prozesse und auf die soziale Interaktion beziehen. Darüber hinaus ist eine hohe Authentizität und Realitätsnähe förderlich (Reinmann & Mandl, 2006). Eine Lernumgebung zur Förderung von Forschungskompetenzen sollte daher möglichst diejenigen Merkmale berücksichtigen, die diese Kompetenzen charakterisieren: Kompetenzen als kognitive Dispositionen, bestehend aus deklarativem und prozeduralem Wissen (R. Mayer, 2003) bzw. Wissen und Fertigkeiten (Nusche, 2008; Simonton, 2003), kennzeichnet, dass sie nicht direkt beobachtbar, sondern internal verfügbar sind. Darüber hinaus findet Forschung überwiegend in einem sozialen Kontext statt: Einerseits arbeiten ForscherInnen in Arbeitsgruppen mit Mitgliedern auf einer ähnlichen Erfahrungsstufe (z.B. KollegInnen in einem Forschungsprojekt), andererseits mit Mitgliedern, die in Bezug auf ihre Forschungserfahrung in einer hierarchischen Beziehung zueinanderstehen (z.B. ProfessorInnen und wissenschaftliche MitarbeiterInnen). In diesen sozialen Kontexten, z.B. im Rahmen von Forschungskolloquien, kann es dazu kommen, dass die ForscherInnen-Gruppe als *informationsverarbeitendes System* agiert (Wecker & Fischer, 2014): Tragen

Gruppenmitglieder in einer Diskussion etwas bei, so kann die gesamte Gruppe *als Ganzes* denken und das Wissen des einzelnen Gruppenmitglieds erweitern.

*Forschungsorientierte Lehre*⁸ (im Englischen bezeichnet als *Research-oriented Teaching*; Griffiths, 2004; Healey, 2005; Healey & Jenkins, 2009) eignet sich als kognitivistisch-konstruktivistische Lernumgebung (vgl. Reinmann & Mandl, 2006) dazu, authentische Situationen mit instruktionalen Mitteln zu unterstützen: Im Rahmen Forschungsorientierter Lehre bekommen Studierende die Gelegenheit, einzelne Phasen oder den gesamten Forschungsprozess selbstständig zu durchlaufen. Sie sollen befähigt werden, die Anforderungen des Forschungsprozesses bewältigen zu können, indem sie die dazu erforderlichen Routinen aufbauen. Zusätzlich wechseln sie auch immer wieder in eine rezeptive Rolle, indem sie gezielt Wissen zu den einzelnen Phasen des Forschungsprozesses erwerben. Lehrende vermitteln in diesen Phasen dieses Wissen gezielt, während sie sonst die Studierenden im Durchlaufen des Forschungsprozesses unterstützen und begleiten. Die Authentizität kann beispielsweise dadurch gewährleistet werden, dass Studierende entweder im Rahmen von Forschungspraktika direkt an Forschung partizipieren oder im Rahmen von Lernforschungsprojekten den gesamten Forschungsprozess durchlaufen.

Um *kognitive Verarbeitungsprozesse* von nicht direkt beobachtbaren Kompetenzen zu unterstützen, eignet sich das kognitive Modellieren. Damit kann eine mentale Repräsentation der zu erlernenden Zielstrategie aufgebaut werden, sodass internal verfügbare Strategien externalisiert werden können durch lautes Denken oder Prompting (Collins, Brown & Newman, 1989; Friedrich & Mandl, 1997; Lipowsky, 2015; Nückles & Wittwer, 2014; Seidel & Reiss, 2014). Inwieweit Lernende die Absicht haben, eine bestimmte Lernhandlung durchzuführen, kann abhängig von ihrer intrinsischen und extrinsischen *Motivation* sein (Schiefele & Schaffner, 2015). Es sollten daher Faktoren in eine Lernumgebung integriert werden, welche diese Lernmotivationsformen beeinflussen. Dies können die Bedürfnisse nach Kompetenzerleben, Autonomieerleben und sozialer Eingebundenheit sein (Deci & Ryan, 1993). Auf Ebene der *sozialen Interaktion* können Elemente integriert werden, die durch das Gefühl der sozialen Eingebundenheit das Aufrechterhalten von Lernmotivation begünstigen (Wecker & Fischer, 2014), aber auch kognitive Verarbeitungsprozesse unterstützen: Gruppenarbeit und die Betreuung von Studierenden. Damit kann einerseits die Interaktion

⁸ Zum Konzept der Forschungsorientierten Lehre können weitere Ansätze hinzugezählt werden: *Forschungsorientiertes Lernen* (L. Huber, 2014), *Forschen üben* (Reinmann, 2017), *Lernen für Forschung* (Reinmann, 2016; Reinmann et al., 2019).

zwischen ForscherInnen auf einer ähnlichen Erfahrungsstufe simuliert werden (als Interaktion zwischen Lernenden in Gruppenarbeit) und andererseits die Interaktion zwischen ForscherInnen mit unterschiedlichen Erfahrungsstufen (als Interaktion zwischen Lehrenden und Lernenden in der Betreuungssituation). Im Rahmen von Gruppenarbeit haben die Gruppenmitglieder die Möglichkeit, durch lernförderliche soziale Aktivitäten, wie beispielsweise das Erklären und die damit verbundene Externalisierung eigener Gedanken, selbst zu lernen und das eigene Verständnis zu fördern (Wecker & Fischer, 2014). Bei unterschiedlichen Sichtweisen von Gruppenmitgliedern kann es darüber hinaus zu (sozio-)kognitiven Konflikten kommen, die wiederum durch die Interaktion in der Gruppe, z.B. durch zusätzliche Erklärungen und Informationen, aufgelöst werden können (Wecker & Fischer, 2014). So kann die Gruppe als informationsverarbeitendes System simuliert werden. Die Betreuung von Studierenden kann insofern zum Kompetenzerwerb beitragen, als dass durch die Interaktion mit „erfahrenen Gemeinschaftsmitgliedern“ (Reinmann & Mandl, 2006, S. 648) einerseits der studentische Kompetenzerwerb begünstigt wird; andererseits können sich Studierende durch diese Interaktion mit Denkmustern, Einstellungen und normativen Standards der *Community of Practice* auseinandersetzen (Nückles & Wittwer, 2014; Reinmann & Mandl, 2006).

Der Cognitive-Apprenticeship-Ansatz (Collins et al., 1989) kann die Förderung von Kompetenzen im Allgemeinen und von Forschungskompetenzen im Speziellen auf allen angesprochenen Ebenen begünstigen. Kernpunkte dieses Ansatzes sind neben dem *Modeling* zur Unterstützung kognitiver Verarbeitungsprozesse das *Coaching* und *Scaffolding* (Collins et al., 1989; Lipowsky, 2015; Seidel & Reiss, 2014). Die Funktion der drei Elemente soll anhand konkreter Beispiele des Forschungsprozesses dargestellt werden: Beim *Modeling* können (meta-)kognitive Prozesse für Studierende verfügbar gemacht werden, indem die Such- und Auswahlprozesse beim Recherchieren von Literatur durch lautes Denken der Lehrenden zugänglich werden. Wenn die Studierenden dann eine eigene Literaturrecherche durchführen, werden sie durch die Lehrenden mit *Coaching* unterstützt. Stoßen die Lernenden dabei an ihre Grenzen, können die Lehrenden durch *Scaffolding* Hilfestellungen geben. So können gezielte Hinweise (*Prompts*) zu Suchstrategien, Rückmeldungen zu den Auswahlkriterien oder schematische Darstellungen zur Verwendung von Suchoperatoren eingesetzt werden. Darüber hinaus bietet der *Cognitive-Apprenticeship*-Ansatz die Möglichkeit, dass Studierende dazu angeregt werden, ihr neu erworbenes Wissen in Worte zu fassen (*Artikulation*), ihr Wissen mit dem Wissen der Lehrenden zu vergleichen und ggf. zu optimieren (*Reflexion*) und den

Forschungsprozess schließlich ohne Unterstützung selbstständig zu durchlaufen (*Exploration*), nachdem die Lehrenden sich komplett zurückgezogen haben (*Fading*).

Wird im Rahmen von Forschungsorientierter Lehre der Cognitive-Apprenticeship-Ansatz eingesetzt, können neben kognitiven auch motivationale und soziale Komponenten gefördert werden: Kompetenzerleben kann dadurch gewährleistet werden, dass Studierende befähigt werden, selbstständig den Forschungsprozess zu durchlaufen und Hinweise (*Prompts*) von den Lehrenden erhalten, wenn sie an ihre Grenzen kommen. Autonomieerleben kann dadurch erreicht werden, indem Studierende gewisse Freiheitsgrade im Forschungsprozess während des *Coachings* und *Fadings* erhalten. Die soziale Eingebundenheit kann dadurch ermöglicht werden, dass Studierende in Gruppen den Forschungsprozess durchlaufen und dabei von Lehrenden unterstützt werden. In diesem sozialen Kontext haben Studierende die Möglichkeit, durch die Interaktion mit Lehrenden oder erfahrenen Gruppenmitgliedern in einem authentischen Handlungsfeld, hier Forschung, in die *Scientific Community* einsozialisiert zu werden.

Die in Kompetenzmodellen definierten Facetten lassen sich im Rahmen von Lernumgebungen zur Förderung dieser Kompetenzen in Aufgaben umformulieren, die es zu absolvieren gilt (Hetmanek et al., 2018). Werden diese Aufgaben möglichst nah am Forschungsalltag gestaltet, so kann die Authentizität der Lernumgebung erhöht werden. Nachfolgend sollen anhand von Beispielen aus publizierten Studien solche Aufgaben vorgestellt werden, die den einzelnen Schritten des Forschungsprozesses entsprechen und damit zur Förderung von Kompetenzen beitragen können. Im Bereich der *Durchführung von Literaturrecherchen* sollen Studierende Recherchestrategien anwenden, indem sie geeignete Datenbanken für die Literaturrecherche kennen, dort recherchieren können und die gefundene Literatur anhand von Qualitätskriterien beurteilen (Ghali et al., 2000; Green, Schmitt-Wilson, Versland, Kelting-Gibson & Nollmeyer, 2016; Groß Ophoff, Schladitz, Leuders, Leuders & Wirtz, 2015; Thoren et al., in Druck; Valter & Akerlind, 2010; Wenglein, J. Bauer, Heininger & Prenzel, 2015). Im Bereich der *Anwendung von Forschungsmethoden* sollen Studierende eigene Fragestellungen oder Hypothesen formulieren (Ghali et al., 2000; Holbrook & Devonshire, 2005; Kippers, Poortman, Schildkamp & Visscher, 2018), bestehende Instrumente zur Generierung von Evidenz kennen und verwenden können (Thoren et al., in Druck), eigene Studiendesigns entwickeln sowie Daten erheben und analysieren (Green et al., 2016; Holbrook & Devonshire, 2005; Kippers et al., 2018; Reeves & Honig, 2015; Valter & Akerlind, 2010; van der Scheer, Glas & Visscher, 2017). Im Bereich der *Interpretation und*

Reflexion der Ergebnisse sollen Studierende Daten interpretieren und mitunter auch Schlussfolgerungen für die Praxis ziehen (Ghali et al., 2000; Green et al., 2016; Holbrook & Devonshire, 2005; Kippers et al., 2018; Reeves & Honig, 2015; Reeves & Chiang, 2017; Valter & Akerlind, 2010; Wenglein et al., 2015). Im Bereich der *Kommunikation über Forschung* sollen Studierende einerseits Schreib- oder Argumentationskompetenzen anwenden, indem sie Forschungsberichte oder fiktive Förderanträge verfassen und dabei u.a. Formatierungsvorgaben und Zitationsstile anwenden (Groß Ophoff et al., 2015; Valter & Akerlind, 2010; Wenglein et al., 2015). Andererseits sollen Studierende ihre Forschungsergebnisse präsentieren (Valter & Akerlind, 2010). Darüber hinaus wird in einigen Formaten gezielt *Wissen über Forschung* vermittelt, welches zur Ausführung der erforderlichen Aufgaben benötigt wird, z.B. Wissen zu Forschungsdesigns (Green et al., 2016; Groß Ophoff et al., 2015) oder fach- bzw. disziplinspezifisches Wissen (Holbrook & Devonshire, 2005; Valter & Akerlind, 2010).

Die Gestaltung der Lernumgebungen in den zuvor genannten Studien zielt auf kognitive Aktivierung, Motivierung und Kooperation im Lernprozess ab. Zu diesem Zweck kommen ganz unterschiedliche Elemente zum Einsatz: Gruppenarbeit (Ghali et al., 2000; Green et al., 2016; Kippers et al., 2018; Reeves & Honig, 2015; Reeves & Chiang, 2017; Thoren et al., in Druck; Valter & Akerlind, 2010; Wenglein et al., 2015), Feedback, Coaching und Scaffolding durch ExpertInnen (Groß Ophoff et al., 2015; Kippers et al., 2018; Reeves & Honig, 2015; Reeves & Chiang, 2017; van der Scheer et al., 2017; Wenglein et al., 2015), Vorlagen und Übungsmaterialien für einzelne Phasen des Forschungsprozesses wie Excel-Tabellen oder fiktive Datensätze (Green et al., 2016; Reeves & Honig, 2015; Reeves & Chiang, 2017) oder die Anknüpfung an authentische Probleme (Ghali et al., 2000; Reeves & Honig, 2015) durch Vignetten (Wenglein et al., 2015) oder Online-Lernumgebungen (Holbrook & Devonshire, 2005).

Durch die Orientierung an konkreten Aufgaben entsprechend der einzelnen Schritte des Forschungsprozesses, kann der Kompetenzerwerb in den intendierten Bereichen gefördert werden. So konnten Groß Ophoff und KollegInnen (2015) nach einem Semester Zuwächse von kleinen bis großen Effekten der durch einen Kompetenztest erfassten *Information Literacy* und des *Evidenzbasierten Schlussfolgerns* finden. Im Laufe dieses Semesters besuchten Studierende eines Studiengangs zu frühkindlicher Bildung einen Kurs zu Forschungsmethoden. In diesem Kurs führten sie unter anderem Literaturrecherchen durch und fertigten eine Seminararbeit entsprechend wissenschaftlicher Standards an, zu der sie

Feedback erhalten haben. Kippers und KollegInnen (2018) untersuchten die Effekte einer einjährigen Intervention bei Lehrkräften und Schulleitungen, in welcher diese die Fähigkeit weiterentwickeln sollten, Daten zu erheben, zu analysieren und zu interpretieren. Der Effekt dieser Intervention wurde mit einem standardisierten Fragebogen mit halb-offenem Antwortformat zur Erfassung dieser Fähigkeiten untersucht. Zum Ende der Intervention zeigte sich ein Zuwachs dieser Fähigkeiten mit einem großen Effekt.

Anknüpfend an die dargestellten Merkmale zur Gestaltung von Lernumgebungen zur Förderung von Forschungskompetenzen und in Anlehnung an die zuvor berichteten Studien, wurde eine Intervention für eine spezielle Zielgruppe entwickelt. Bei dieser Zielgruppe handelt es sich um Lehramtsstudierende, die an Universitäten für eine evidenzbasierte Praxis ausgebildet werden.

2.3.2 Forschungskompetenzen für eine evidenzbasierte Praxis

Studierende von professionsbezogenen Wissenschaften werden im Gegensatz zu Studierenden von Grundlagenwissenschaften auf ganz konkrete berufliche Anforderungen vorbereitet: So werden Studierende in der Medizin (Ghali et al., 2000; Sackett, Rosenberg, Gray, Haynes & Richardson, 1996) oder im Lehramt (Green et al., 2016; Kippers et al., 2018; Reeves & Honig, 2015; Reeves & Chiang, 2017; Thoren et al., in Druck; van der Scheer et al., 2017) explizit auf ihre Rolle als evidenzbasierte ExpertInnen vorbereitet. Während die Vorbereitung auf eine *evidenzbasierte Praxis* im Bereich der Medizin schon seit längerem in die Ausbildung der Studierenden integriert ist (Sackett et al., 1996), wird dies im Lehramt erst in jüngster Zeit thematisiert (J. Bauer & Prenzel, 2012; J. Bauer et al., 2015; Borg, 2010; Davies, 1999; Hargreaves, 1996, 1997; Trempler et al., 2015). So wird die Vorbereitung auf eine evidenzbasierte Praxis mittlerweile sowohl national (Kultusministerkonferenz, 2004) als auch international (J. Bauer & Prenzel, 2012; European Commission, 2013; Hamilton et al., 2009) als ein wichtiger Bestandteil des Lehralltags gesehen.

Die vielfach für Lehrkräftehandeln aufgestellte Annahme (Terhart, 2002), dass eine Wirkungskette Lehrkräftebildung – Lehrkräftehandeln – Lernerfahrung von SchülerInnen – SchülerInnen-Outcomes besteht, wurde auch implizit der evidenzbasierten Praxis zugrunde gelegt: Evidenzbasierte Praxis sollte einen positiven Einfluss auf SchülerInnen-Outcomes haben (J. Bauer & Prenzel, 2012; Trempler et al., 2015). Diese Wirkungskette ist jedoch aufgrund ihrer Komplexität schwer zu überprüfen (König, 2015), sodass bislang dazu wenig Evidenz vorliegt (z.B. Marsh, 2012). Nichtsdestotrotz treffen Lehrkräfte in ihrer beruflichen

Praxis immer häufiger auf Ergebnisse aus (inter-)nationalen Schulleistungsstudien oder Vergleichsarbeiten (z.B. PISA, TIMMS bzw. VERA); die Evaluation von Unterrichtsqualität und eine evidenzbasierte Schulentwicklung nehmen einen größer werdenden Stellenwert in der beruflichen Praxis ein (Humpert, Hauser & Nagl, 2006). Lehrkräfte sollen daher aus dieser Evidenz Maßnahmen zur Entwicklung des eigenen Unterrichts ableiten können und das implizite und intuitive Handeln, welches auf praktischer Erfahrung beruht (Ophardt & Thiel, 2013), *ergänzen* (J. Bauer et al., 2015).

Evidenz kann für verschiedene Phasen des Unterrichtens relevant sein, die hier exemplarisch in einer „überzeichneten Reinform“ (Trempler et al., 2015, S. 147) dargestellt werden sollen. So kann beispielsweise evidenzbasiert *geplant* werden (Trempler et al., 2015): Für einen bestimmten Inhaltsbereich wird eine passende Unterrichtsmethode gesucht. Dazu wird der Inhalt der Unterrichtsstunde, die Zielgruppe und weitere relevante Faktoren von der Lehrkraft analysiert und entsprechende Evidenz gesucht, die möglichst vergleichbar mit diesen Faktoren ist. Solche Evidenz ist insbesondere in Meta-Analysen zu finden (Davies, 1999; in der konkreten Umsetzung z.B. Petty, 2014). Die entsprechende Unterrichtsmethode wird dann in die Planung des Unterrichts integriert. Es kann aber auch evidenzbasiert *reflektiert* werden, im Sinne einer *reflection-on-action* (Schön, 1983): So kann das Unterrichtsgeschehen im Nachhinein mental durchdacht, verschiedene alternative Verläufe simuliert und diese Optionen kriterial beurteilt werden (Ophardt & Thiel, 2013). Im Rahmen dieses Durchdenkens und Simulierens kann Evidenz zu verschiedenen Unterrichtsmethoden verwendet werden, um eine Entscheidung zu treffen, welche Methode in einer künftig vergleichbaren Unterrichtssituation eingesetzt wird.

Letztlich kann Evidenz mit allen Bereichen des professionellen Wissens von Lehrkräften (Shulman, 1987; in der Übersetzung von Baumert & Kunter, 2006) kombiniert werden und ggf. zu einer gegenseitigen Ergänzung führen (Mandinach & Gummer, 2016): So werden sowohl Fachwissen (Content Knowledge) als auch fachdidaktisches Wissen (Pedagogical Content Knowledge) benötigt, da Lehrkräfte einerseits Evidenz innerhalb ihres Fachs verorten, andererseits die aus der Evidenz generierten Maßnahmen in didaktische Schritte umsetzen können sollten. Kenntnisse über die Psychologie der Lernenden (Knowledge of Learners and Their Characteristics) sind insofern notwendig, als dass Evidenz zu nicht-kognitiven Merkmalen wie Motivierungsqualität die zu ergreifenden Maßnahmen unterstützen können. Organisationswissen (Knowledge of Educational Contexts) kann helfen zu verstehen, wie sich die Evidenz in den Kontext der eigenen Schule oder des regionalen

Schulkontextes einordnen lässt. Erziehungsphilosophisches, bildungstheoretisches und bildungshistorisches Wissen (Knowledge of Educational Ends, Purposes and Values) kann helfen zu verstehen, wie Bildungsstandards zustande kommen und wie Evidenz zu SchülerInnen-Leistungen diese Standards erfüllen. Allgemeines pädagogisches Wissen (General Pedagogical Knowledge) und Wissen über das Fachcurriculum (Curriculum Knowledge) werden benötigt, um die generierte Evidenz in Bezug zum Fachcurriculum und den formulierten Kompetenzziele zu setzen.

Da die evidenzbasierte Praxis und deren einzelne Teilschritte in der Medizin bereits ausbuchstabiert und in die medizinische Praxis integriert wurden (Sackett et al., 1996), ist das medizinische Konzept auf das Lehramt übertragen worden (Hargreaves, 1996, 1997). Lehrkräfte im Lehralltag sollen in der Lage sein, (a) benötigte Informationen für ein praktisches Problem in Fragen umzuwandeln, (b) bestehende Evidenz zu suchen, um diese Fragen zu beantworten, (c) die gefundene Evidenz hinsichtlich ihrer Gültigkeit und Nützlichkeit zu bewerten, (d) die Evidenz für die Lösung des Praxisproblems anzuwenden und (e) das Ergebnis dieser Anwendung von Evidenz zu evaluieren (Hargreaves, 1996, 1997). Damit angehende Lehrkräfte auf diese evidenzbasierte Praxis schon während der Lehramtsausbildung vorbereitet werden können, wurden in Deutschland bereits im Jahr 2004 die Standards für die Lehrerbildung (Kultusministerkonferenz, 2004) im Bereich *Innovieren* um das evidenzbasierte Entscheiden erweitert. Evidenzbasiertes Entscheiden kann sowohl aus der *Nutzung* als auch aus dem *Generieren* von Evidenz bestehen.

Die *Nutzung* von Evidenz (Borg, 2010; Davies, 1999) umfasst die „reflektierte Nutzung der besten verfügbaren und empirisch gesicherten Evidenz für das Treffen und die Begründung von beruflichen Entscheidungen“ (Trempler et al., 2015, S. 145). Dazu gehört es, (Forschungs-)Fragen zu stellen, Informationen auszuwählen, und diese gefundene Evidenz hinsichtlich ihrer Qualität zu bewerten (Groß Ophoff et al., 2014, 2017; Trempler et al., 2015). Weiterhin ist es notwendig, mit der gefundenen Evidenz, die in ihren typischen Repräsentationsformen dargestellt wird (z.B. Tabellen und Abbildungen), umgehen zu können, sie zu interpretieren und zu beurteilen (Groß Ophoff et al., 2014, 2017), um Schlussfolgerungen daraus ableiten zu können.

Das *Generieren* von Evidenz (Borg, 2010; Davies, 1999) beinhaltet darüber hinaus den Aspekt der Anwendung basaler Forschungsmethoden. Dazu ist die systematische Bearbeitung einer gestellten Frage mit Forschungsmethoden nötig, um Schlussfolgerungen für die Praxis zu ziehen (Borg, 2010; Davies, 1999). Mandinach und Gummer (2016) haben das Generieren

von Evidenz in den fünf Schritten ihres Modells zu *Data Literacy for Teachers* (DLFT) verdeutlicht: (1) Probleme identifizieren und Fragen stellen – Benennung eines praktischen Problems vor dem Kontext der SchülerInnen einer Klasse und weiterer schulischer Bedingungen; (2) Daten nutzen – Erhebung und Analyse von benötigten Daten aus verschiedenen Datenquellen; dieser Schritt beinhaltet explizit das *Generieren* von Evidenz, sofern diese nicht bereits aus anderen Quellen verfügbar ist, und die Anwendung *basaler* Forschungsmethoden; (3) Daten in Informationen transferieren – basale Statistiken nutzen können, um Evidenz zu interpretieren und zusammenzufassen, um Schlussfolgerungen daraus zu ziehen; (4) Informationen in Entscheidungen transferieren – den Unterricht aufgrund der Schlussfolgerungen und mit den dazu entwickelten Maßnahmen anpassen; (5) Outcomes evaluieren – die Auswirkungen der Maßnahmen mit dem Ausgangszustand vergleichen, prüfen, ob sich Veränderungen ergeben haben und gegebenenfalls den Zyklus von vorne beginnen.

Das Generieren von Evidenz kann als Spezialform von Forschungskompetenzen gesehen werden. Die Phasen und die dazu benötigten Kompetenzen entsprechen den in Kapitel 2.2.1 dargestellten Phasen des Forschungsprozesses: Durchführung von Literaturrecherchen, Anwendung von Forschungsmethoden, Interpretation und Reflexion der Befunde. Das heißt, dass die für eine evidenzbasierte Praxis benötigten Kompetenzen zur Generierung von Evidenz zwar prinzipiell in verschiedenen Domänen anwendbar sind, dass sie aber für das evidenzbasierte Entscheiden im spezifischen Kontext der Domäne Lehramt angewendet werden. In der Systematisierung der Abbildung 2.2 wären diese Kompetenzen damit auf der Grenze zwischen *domänenspezifischen* und *domänenübergreifenden* Kompetenzen einzuordnen.

Die Förderung von Kompetenzen für eine evidenzbasierte Praxis sollte nach Ansicht von Mandinach und Gummer (2016) am besten *während* der Lehramtsausbildung in Lernumgebungen ermöglicht werden, in denen sowohl die Nutzung und das Generieren von Evidenz als auch das Unterrichten ermöglicht wird. Diese Kombination soll bewirken, dass Lehramtsstudierende so früh wie möglich die Ergänzung ihres professionellen Wissens durch Evidenz in der praktischen Umwelt erfahren können (Mandinach & Gummer, 2016). Übertragen auf den Kontext dieser Arbeit bedeutet das, dass Studierende eine forschungsorientierte Lehrveranstaltung besuchen, die im schulischen Kontext verortet ist. Dazu eignen sich Lehrveranstaltungen in Form von Lernforschungsprojekten, die während einer längeren Praxisphase an den Schulen begleitend an der Universität stattfinden.

Studierende haben dann sowohl die Möglichkeit, Evidenz durch das Durchlaufen des Forschungsprozesses zu generieren als auch Schlussfolgerungen für ihre eigene, erste Unterrichtspraxis zu ziehen.

2.4 Fragestellungen und Überblick über die empirischen Studien

Wie gezeigt wurde, stellen Kompetenzmodelle einerseits die Grundlage für die Erfassung von Kompetenzen dar (Erpenbeck & von Rosenstiel, 2007; Klieme & Hartig, 2008). Andererseits unterstützen sie die Definition von Kompetenzzielen und dienen als Orientierung bei der Entwicklung von Lernumgebungen zur Förderung dieser Kompetenzen (Erpenbeck & von Rosenstiel, 2007; Hetmanek, Engelmann, Opitz & Fischer, 2018; Klieme & Leutner, 2006). Den Outcome dieser Lernumgebungen können Forschungskompetenzen darstellen, die als eine Spezialform des domänenübergreifenden wissenschaftlichen Denkens und Argumentierens identifiziert wurden. Ausgehend von der pragmatischen Wissenschaftstheorie (Butts, 1991; Dewey, 1938; Mittelstraß, 1991), nach der sich gemeinsame Merkmale von Forschung identifizieren lassen, wurde das RMRK-W-Modell (Thiel & Böttcher, 2014) mit den Dimensionen *Recherche-, Methoden-, Reflexions- und Kommunikationskompetenzen* sowie *Fachliches Wissen* entwickelt. Dieses Modell bildete die Grundlage für die fach- bzw. disziplinübergreifende Erfassung dieser Kompetenzen via Selbsteinschätzung mit dem Fragebogen zur Erfassung studentischer Forschungskompetenzen. Im Rahmen der vorliegenden Arbeit soll dieser Fragebogen validiert werden. Daher ergibt sich die erste Forschungsfrage:

- *Forschungsfrage 1:* Können Forschungskompetenzen mit dem F-Komp valide erfasst werden (Studie 1 und 2)?

Aus dem RMRK-W-Modell (Thiel & Böttcher, 2014) lassen sich Kompetenzziele ableiten, auf deren Grundlage forschungsorientierte Lernumgebungen zum Erwerb von Forschungskompetenzen gestaltet und evaluiert werden können. Als eine wichtige Zielgruppe für forschungsorientierte Lehre wurden Lehrkräfte identifiziert, die im Studium u.a. für eine evidenzbasierte Praxis ausgebildet werden sollen. Die zweite Forschungsfrage bezieht sich auf die Gestaltung von Lernumgebungen zur Förderung von Forschungskompetenzen für eine evidenzbasierte Praxis:

- *Forschungsfrage 2:* Lassen sich Kompetenzen von Lehramtsstudierenden für eine evidenzbasierte Praxis in einer kompetenzmodellbasierten, forschungsorientierten Lernumgebung fördern (Studie 3)?

Zur Überprüfung der ersten Forschungsfrage wurden die Studien 1 und 2 durchgeführt. Während die erste Studie sich vor allem auf die Überprüfung der faktoriellen Validität bezieht, werden in der zweiten Studie die konvergente und diskriminante Validität überprüft. Die zweite Forschungsfrage wird in Studie 3 behandelt.

Studie 1: Erfassung von Forschungskompetenzen – faktorielle Validität

In Studie 1 (Evaluating research-oriented teaching: a new instrument to assess university students' research competences) wird der neu entwickelte *Fragebogen zur Erfassung studentischer Forschungskompetenzen* ausführlich vorgestellt. Mit diesem Instrument ist es möglich, Forschungskompetenzen als *Outcome* forschungsorientierter Lernumgebungen *fach- bzw. disziplinübergreifend* zu erfassen. Nach einer theoretischen Herleitung des RMRK-W-Modells sowie des zugrundeliegenden Kompetenzbegriffs werden die Konstruktion des Fragebogens sowie der bis dato durchgeführte Schritt der *Inhaltsvalidierung* berichtet. Anhand einer Stichprobe von 391 Studierenden in Bachelor-, Master- und Promotionsstudiengängen aus unterschiedlichen Disziplinen wurde die faktorielle Validität des Instrumentes mit Hilfe konfirmatorischer Faktorenanalysen sowie die Reliabilität auf manifester und latenter Ebene überprüft.

Studie 2: Erfassung von Forschungskompetenzen – konvergente und diskriminante Validität

In Studie 2 (Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen – Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements) werden weitere Schritte zur Konstruktvalidierung vorgenommen. Es werden verschiedene Erwartungen formuliert hinsichtlich der Fragen, ob mit dem *Fragebogen zur Erfassung studentischer Forschungskompetenzen* (1) keine allgemeinen Lern- und Studienkompetenzen erfasst werden sowie (2) inwieweit das Instrument valide im Hinblick auf die gewählte Methodik der Selbsteinschätzung misst. Zur ersten Frage werden Zusammenhänge untersucht zu Konstrukten, die mit den in Kapitel 2.1 dargestellten akademischen Kompetenzen in Verbindung stehen: zu (a) epistemologischen Überzeugungen, da wissenschaftliches Wissen, welches im Laufe des Forschungsprozesses generiert wird, als unsicher bezeichnet werden kann (Hofer & Pintrich, 1997; Schommer, 1998; Trautwein & Lüdtke, 2007), weswegen die Fähigkeit zum selbstregulierten Lernen notwendig ist, zu (b) ausgewählten, prozessbezogenen kognitiven Lernstrategien, da insbesondere diese Strategien relevant sind für die Bewältigung der Anforderungen des komplexen und selbstorganisierten Forschungsprozesses (Willison & O'Regan, 2007) sowie zu (c) domänenunabhängigen, kognitiven Fähigkeiten, da diese

zeitstabile Fähigkeiten zum Problemlösen darstellen, für die im Gegensatz zu akademisch vermittelten Kompetenzen kein Vorwissen benötigt wird (Klieme & Hartig, 2008; Weinert, 2001). Dabei wurden die Analysen zu (a) und (b) auf latenter, sowie die Analysen zu (c) auf manifester Ebene durchgeführt. Zur zweiten Frage werden die Zusammenhänge (d) auf manifester Ebene zu einem Kompetenztest sowie (e) auf latenter Ebene zu einem Instrument untersucht, das forschungsbezogene Selbstwirksamkeitserwartungen ebenfalls über die Methodik der Selbsteinschätzungen erfasst. Damit wurde die *konvergente* und die *diskriminante Validität* des Instrumentes anhand von drei Teilstichproben mit 536 Masterstudierenden aus unterschiedlichen Fächern (Teilstichprobe 1), 98 Lehramtsstudierenden im Master (Teilstichprobe 2) sowie 78 Lehramtsstudierenden im Master (Teilstichprobe 3) überprüft.

Studie 3: Förderung von Forschungskompetenzen

In Studie 3 (How do we prepare teacher training students for Evidence-based Practice? Promoting students' research competences in Research-learning projects) wird eine forschungsorientierte Lernumgebung konzipiert und evaluiert. Diese Lernumgebung ist speziell auf die Zielgruppe von angehenden Lehrkräften zugeschnitten, da Lehrkräfte im Rahmen ihres Studiums für eine evidenzbasierte Praxis ausgebildet werden sollen (Kultusministerkonferenz, 2004), indem sie Kompetenzen zur *Nutzung* und zur *Generierung* von Evidenz erwerben (Davies, 1999; Kapitel 2.3.2). Als Grundlage der Gestaltung der Lernumgebung werden zwei Modelle gewählt: einerseits das RMRK-W-Modell, welches Forschungskompetenzen differenziert und fach- bzw. disziplinübergreifend beschreibt und insbesondere den Aspekt der *Generierung von Evidenz* (Davies, 1999) umfasst, andererseits das Modell zur bildungswissenschaftlichen Forschungskompetenz (Groß Ophoff et al., 2014, 2017), welches speziell auf die Zielgruppe von BildungswissenschaftlerInnen ausgerichtet ist und ausschließlich den Aspekt der *Nutzung von Evidenz* (Davies, 1999) fokussiert. Anhand dieser Modelle werden Kompetenzziele definiert und ein forschungsorientiertes Curriculum entwickelt. Studierende erhalten im Rahmen von Lernforschungsprojekten die Gelegenheit, den gesamten Forschungsprozess selbstständig zu durchlaufen. Sie werden von den Lehrenden insbesondere mit *Modeling* und *Coaching* unterstützt. Die Wirksamkeit dieser Lehrveranstaltung wurde in einem quasi-experimentellen Prä-Post-Kontrollgruppendesign anhand einer Stichprobe von 36 Lehramtsstudierenden in der Experimentalgruppe und 23 Studierenden in der Kontrollgruppe überprüft. Es wurde eine Kombination der standardisierten Erhebungsverfahren *Kompetenzselbsteinschätzung* und *Kompetenztest* unter

der Berücksichtigung des Response-Shift-Phänomens (Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999) verwendet, welches in längsschnittlichen Studien auftritt und eine Veränderung der Kriterien für die Selbsteinschätzung eines internalen Konstruktes beschreibt.

Literatur I

- Bach, A. (2013). *Kompetenzentwicklung im Schulpraktikum. Ausmaß und zeitliche Stabilität von Lerneffekten hochschulischer Praxisphasen* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 87, 1. Aufl.). Münster: Waxmann.
- Bauer, J. & Prenzel, M. (2012). Science education. European teacher training reforms. *Science*, 336(6089), 1642–1643.
- Bauer, J., Prenzel, M. & Renkl, A. (2015). Evidenzbasierte Praxis - im Lehrerberuf?! Einführung in den Thementeil. *Unterrichtswissenschaft*, 43(3), 188–192.
- Bauer, W., Bleck-Neuhaus, J., Dombois, R. & Wehrtmann, I. (2013). *Forschungsprojekte entwickeln - von der Idee bis zur Publikation* (1. Aufl.). Baden-Baden: Nomos.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Beckert, R., Fanghänel, E., Habicher, W. d., Knölker, H.-J., Metz, P. & Schwetlick, K. (2015). *Organikum. Organisch-chemisches Grundpraktikum* (24., vollständig überarbeitete und aktualisierte Auflage). Weinheim: Wiley-VCH Verlag.
- Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429.
- Brew, A. (2006). *Research and teaching. Beyond the divide* (Universities into the 21st century). Houndmills, Großbritannien: Palgrave Macmillan.
- Bundesassistentenkonferenz (2009). *Forschendes Lernen - wissenschaftliches Prüfen. Ergebnisse der Arbeit des Ausschusses für Hochschuldidaktik* (Schriften der Bundesassistentenkonferenz, Bd. 5, Neuaufl. nach der 2. Aufl. 1970). Bielefeld: UVW, Webler.
- Butts, R. E. (1991). Methodology, metaphysics and the pragmatic unity of science. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 23–28). Berlin: W. de Gruyter.
- Cattell, R. B. (1987). *Intelligence. Its structure, growth and action*. Amsterdam: Elsevier Science.

- Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship. Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Hrsg.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (S. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108–121.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238.
- Decker, L., Kaplan, I. & Siebert-Ott, G. (2015). *Modellierung und Erfassung akademischer Textkompetenzen (AkaTex Working Papers 3)* (2. korrigierte Auflage). Siegen, Köln. Zugriff am 06.08.2018. Verfügbar unter https://www.uni-siegen.de/phil/akatex/publikationen_und_vortraege/download/ako-workingpaper-neu.pdf
- Deutsche Forschungsgemeinschaft & Wissenschaftsrat (2015). *Bericht der Gemeinsamen Kommission zur Exzellenzinitiative an die Gemeinsame Wissenschaftskonferenz*. Bonn. Zugriff am 03.06.2019. Verfügbar unter <http://www.gwk-bonn.de/fileadmin/Papers/DFG-WR-Bericht-Juni2015.pdf>
- Dewey, J. (1938). *Unity of science as a social problem*. Chicago, IL: University of Chicago Press.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (Springer-Lehrbuch, 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Berlin: Springer.
- Dunn, K. E., Airola, D. T., Lo, W.-J. & Garrison, M. (2013). What teachers think about what they can do with data. Development and validation of the data driven decision-making efficacy and anxiety inventory. *Contemporary Educational Psychology*, 38(1), 87–98.
- Enders, J. (2010). Hochschulen und Fachhochschulen. In D. Simon, A. Knie & S. Hornbostel (Hrsg.), *Handbuch Wissenschaftspolitik* (S. 443–456). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Erpenbeck, J. & von Rosenstiel, L. (2007). Einführung. In J. Erpenbeck & L. von Rosenstiel (Hrsg.), *Handbuch Kompetenzmessung. Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (2., überarbeitete und erweiterte Auflage, S. XVII–XLV). Stuttgart: Schäffer-Poeschel.
- European Commission (2013). *Supporting teacher competence development for better learning outcomes*.

- Fischer, F., Chinn, C. A., Engelmann, K. F. & Osborne, J. (Hrsg.). (2018). *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge*. New York, NY: Routledge.
- Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R. et al. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, 2(3), 28–45.
- Frey, A. (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften – eine erste Standortbestimmung zu bereits publizierten Instrumenten. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (Zeitschrift für Pädagogik Beiheft, Bd. 51, S. 30–46). Weinheim: Beltz.
- Friedrich, H. F. & Mandl, H. (1997). Analyse und Förderung selbstgesteuerten Lernens. In F. E. Weinert & H. Mandl (Hrsg.), *Psychologie der Erwachsenenbildung* (S. 237–293). Göttingen: Hogrefe.
- Funke, J. (2003). *Problemlösendes Denken* (Kohlhammer Standards Psychologie, 1. Aufl.). Stuttgart: Kohlhammer.
- Gess, C., Geiger, C. & Ziegler, M. (2018). Social-scientific research competency. Validation of Test Score Interpretations for Evaluative Purposes in Higher Education. *European Journal of Psychological Assessment*, 1–14.
- Gess, C., Wessels, I. & Blömeke, S. (2017). Domain-specificity of research competencies in the social sciences. Evidence from differential item functioning. *Journal for Educational Research Online*, 9(2), 11–36.
- Ghali, W. A., Saitz, R., Eskew, A. H., Gupta, M., Quan, H. & Hershman, W. Y. (2000). Successful teaching in evidence-based medicine. *Medical Education*, 34(1), 18–22.
- Goldman, S. R., Ko, M.-L. M., Greenleaf, C. & Brown, W. (2018). Domain-specificity in the practices of explanation, modeling, and argument in the sciences. In F. Fischer, C. A. Chinn, K. F. Engelmann & J. Osborne (Hrsg.), *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge* (S. 121–141). New York, NY: Routledge.
- Green, J. L., Schmitt-Wilson, S., Versland, T. M., Kelting-Gibson, L. & Nollmeyer, G. E. (2016). Teachers and data literacy: A blueprint for professional development to foster data driven decision making. *Journal of Continuing Education and Professional Development*, 3(1), 14–32.

- Greiff, S., Wüstenberg, S. & Funke, J. (2012). Dynamic problem solving. A new assessment perspective. *Applied Psychological Measurement*, 36(3), 189–213.
- Griffiths, R. (2004). Knowledge production and the research–teaching nexus: The case of the built environment disciplines. *Studies in Higher Education*, 29(6), 709–726.
- Groß Ophoff, J., Schladitz, S., Leuders, J., Leuders, T. & Wirtz, M. A. (2015). Assessing the development of educational research literacy: The effect of courses on research methods in studies of educational science. *Peabody Journal of Education*, 90(4), 560–573.
- Groß Ophoff, J., Schladitz, S., Lohrmann, K. & Wirtz, M. (2014). Evidenzorientierung in bildungswissenschaftlichen Studiengängen. In K. Drossel, R. Strietholt & W. Bos (Hrsg.), *Empirische Bildungsforschung und evidenzbasierte Reformen im Bildungswesen* (S. 251–275). Münster: Waxmann.
- Groß Ophoff, J., Wolf, R., Schladitz, S. & Wirtz, M. (2017). Assessment of educational research literacy in higher education: Construct validation of the factorial structure of an assessment instrument comparing different treatments of omitted responses. *Journal for Educational Research Online/Journal für Bildungsforschung Online*, 9(2), 37–68.
- Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A. & Wayman, J. C. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Zugriff am 03.06.2019. Verfügbar unter https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm_pg_092909.pdf
- Hargreaves, D. H. (1996). *Teaching as a research-based profession: Possibilities and prospects*. Teacher Training Agency Annual Lecture, London.
- Hargreaves, D. H. (1997). In defence of research for evidence-based teaching: a rejoinder to Martyn Hammersley. *British Educational Research Journal*, 23(4), 405–419.
- Hartmann, D., Mohseni, A., Reckwitz, E., Rojek, T. & Steckmann, U. (Hrsg.). (2012). *Methoden der Geisteswissenschaften. Eine Selbstverständigung* (Erste Auflage). Weilerswist: Velbrück Wissenschaft.
- Hattie, J. & Marsh, H. W. (1996). The relationship between research and teaching: A meta-analysis. *Review of Educational Research*, 66(4), 507–542.
- Healey, M. (2005). Linking research and teaching to benefit student learning. *Journal of Geography in Higher Education*, 29(2), 183–201.

- Healey, M. & Jenkins, A. (2009). *Developing undergraduate research and inquiry*. Heslington, York, England: Higher Education Academy.
- Hetmanek, A., Engelmann, K. F., Opitz, A. & Fischer, F. (2018). Beyond intelligence and domain knowledge. Scientific reasoning and argumentation as a set of cross-domain skills. In F. Fischer, C. A. Chinn, K. F. Engelmann & J. Osborne (Hrsg.), *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge* (203-226). New York, NY: Routledge.
- Hochschulrahmengesetz in der Fassung der Bekanntmachung vom 19. Januar 1999 (BGBl. I S. 18), das zuletzt durch Artikel 6 Absatz 2 des Gesetzes vom 23. Mai 2017 (BGBl. I S. 1228) geändert worden ist (1999). Zugriff am 03.06.2019. Verfügbar unter <https://www.gesetze-im-internet.de/hrg/BJNR001850976.html#BJNR001850976BJNG000102310>
- Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140.
- Holbrook, N. J. & Devonshire, E. (2005). Simulating scientific thinking online: an example of research-led teaching. *Higher Education Research & Development*, 24(3), 201–213.
- Huber, L. (2014). Forschungsbasiertes, Forschungsorientiertes, Forschendes Lernen: Alles dasselbe? Ein Plädoyer für eine Verständigung über Begriffe und Unterscheidungen im Feld forschungsnahen Lehrens und Lernens. *Das Hochschulwesen*, 62(1+2), 32–39.
- Humpert, W., Hauser, B. & Nagl, W. (2006). Was (zukünftige) Lehrpersonen über wissenschaftliche Methoden und Statistik wissen sollen und wollen. *Beiträge zur Lehrerinnen-und Lehrerbildung*, 24(2), 231–244.
- Jaksztat, S., Schindler, N. & Briedis, K. (2010). *Wissenschaftliche Karrieren. Beschäftigungsbedingungen, berufliche Orientierungen und Kompetenzen des wissenschaftlichen Nachwuchses* (HIS: Forum Hochschule Nr. 14).
- Kippers, W. B., Poortman, C. L., Schildkamp, K. & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation*, 56, 21–31.
- Klieme, E. & Hartig, J. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (Zeitschrift für Erziehungswissenschaft Sonderheft 8 | 2007, S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Klieme, E., Hartig, J. & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 3–22). Göttingen: Hogrefe & Huber Publishers.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876–903.
- König, J. (2015). Stand der Forschung zu wissenschaftsbezogenen Kompetenzen und weiterführende Fragen. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik, Beiheft, Bd. 61, S. 226–235). Weinheim: Beltz Juventa.
- Kultusministerkonferenz (2004). *Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004.*
- Kultusministerkonferenz (2005). *Qualifikationsrahmen für Deutsche Hochschulabschlüsse. Im Zusammenwirken von Hochschulrektorenkonferenz, Kultusministerkonferenz und Bundesministerium für Bildung und Forschung erarbeitet und von der Kultusministerkonferenz am 21.04.2005 beschlossen.*
- Leutner, D., Fleischer, J., Wirth, J., Greiff, S. & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien. Untersuchungen zur Dimensionalität. *Psychologische Rundschau*, 63(1), 34–42.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (Springer-Lehrbuch, Bd. 103, S. 69–105). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Mandinach, E. B. & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366–376.
- Marsh, J. A. (2012). Interventions promoting educators' use of data. Research insights and gaps. *Teachers College Record*, 114(11), 1–48.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biomedizinischen Forschung: Ein Handbuch für Lehramtsstudenten und Doktoranden* (S. 177–186). Berlin: Springer.

- Mayer, R. E. (2003). What causes individual differences in cognitive performance? In R. J. Sternberg & E. Grigorenko (Hrsg.), *Perspectives on the psychology of abilities, competencies, and expertise* (S. 263–273). New York, NY: Cambridge University Press.
- Meijers, A. W. M., van Overveld, C. W. A. M & Perrenet, J. C. (2005). *Criteria for academic bachelor's and master's curricula*. Delft, Niederlande: TU Delft.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741.
- Mittelstraß, J. (1991). Einheit und Transdisziplinarität. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 12–22). Berlin: W. de Gruyter.
- Nückles, M. & Wittwer, J. (2014). Lernen und Wissenserwerb. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie. Mit Online-Materialien* (Psychologie 2014, 6., vollständig bearb. Aufl., S. 225–252). Weinheim: Julius Beltz.
- Nusche, D. (2008). *Assessment of learning outcomes in higher education. A comparative review of selected practices* (OECD Education Working Papers, Bd. 15). Frankreich: OECD publishing.
- OECD (2015). *Frascati manual 2015. Guidelines for collecting and reporting data on research and experimental development* (The measurement of scientific, technological and innovation activities). Paris: OECD.
- Ophardt, D. & Thiel, F. (2013). *Klassenmanagement. Ein Handbuch für Studium und Praxis* (Schulpädagogik). Stuttgart: Kohlhammer.
- Osborne, J. (2018). Styles of scientific reasoning: What can we learn from looking at the product, not the process, of scientific reasoning? In F. Fischer, C. A. Chinn, K. F. Engelmann & J. Osborne (Hrsg.), *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge* (S. 162–186). New York, NY: Routledge.
- Perrenet, J. & Adan, I. (2010). The academic merits of modelling in higher mathematics education: A case study. *Mathematics Education Research Journal*, 22(2), 121–140.
- Petty, G. (2014). *Evidence-based teaching. A practical approach* (Second edition). Oxford: Oxford University Press.

- Reeves, T. D. & Chiang, J.-L. (2017). Building pre-service teacher capacity to use external assessment data: An intervention study. *The Teacher Educator*, 52(2), 155–172.
- Reeves, T. D. & Honig, S. L. (2015). A classroom data literacy intervention for pre-service teachers. *Teaching and Teacher Education*, 50, 90–101.
- Reinmann, G. (2016). Gestaltung akademischer Lehre: semantische Klärungen und theoretische Impulse zwischen Problem- und Forschungsorientierung. *Zeitschrift für Hochschulentwicklung*, 11(5), 225–244.
- Reinmann, G. (2017). Prüfungen und Forschendes Lernen. In H. A. Mieg & J. Lehmann (Hrsg.), *Forschendes Lernen. Wie die Lehre in Universität und Fachhochschule erneuert werden kann* (S. 115–128). Frankfurt: Campus Verlag.
- Reinmann, G., Lübcke, E. & Heudorfer, A. (Hrsg.). (2019). *Forschendes Lernen in der Studieneingangsphase. Empirische Befunde, Fallbeispiele und individuelle Perspektiven*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Reinmann, G. & Mandl, H. (2006). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie. Ein Lehrbuch* (Anwendung Psychologie, 5., vollst. überarb. Aufl., S. 613–658). Weinheim: Beltz PVU.
- Renkl, A. (2018). Scientific reasoning and Argumentation: Is there an over-emphasis on discipline specificity? In F. Fischer, C. A. Chinn, K. F. Engelmann & J. Osborne (Hrsg.), *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge* (S. 194–200). New York, NY: Routledge.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *BMJ*, 312(7023), 71–72.
- Schaper, N., Reis, O., Wildt, J. & Horvath, E. (Hrsg.). (August 2012). *Fachgutachten zur Kompetenzorientierung in Studium und Lehre* (HRK Projekt nexus).
- Schiefele, U. & Schaffner, E. (2015). Motivation. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (Springer-Lehrbuch, S. 153–175). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schneider, R. & Wildt, J. (2009). Forschendes Lernen und Kompetenzentwicklung. In L. Huber (Hrsg.), *Forschendes Lernen im Studium. Aktuelle Konzepte und Erfahrungen* (Motivierendes Lehren und Lernen in Hochschulen, Bd. 10, S. 53–68). Bielefeld: UVW, Webler.
- Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology*, 68(4), 551–562.

- Schön, D. A. (1983). *The reflective practitioner. How professionals think in action*. New York, NY: Basic Books.
- Schwartz, C. E. & Sprangers, M. A. G. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11), 1531–1548.
- Schwartz, C. E. & Sprangers, M. A. G. (2010). Guidelines for improving the stringency of response shift research using the thentest. *Quality of Life Research*, 19(4), 455–464.
- Seidel, T. & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie. Mit Online-Materialien* (Psychologie 2014, 6., vollständig bearb. Aufl., S. 253–275). Weinheim: Julius Beltz.
- Shavelson, R. J. & Towne, L. (2002). *Scientific Research in Education*. Washington, D.C.: National Academy Press.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23.
- Simonton, D. K. (2003). Expertise, competence, and creative ability. In R. J. Sternberg & E. Grigorenko (Hrsg.), *Perspectives on the psychology of abilities, competencies, and expertise* (S. 213–240). New York, NY: Cambridge University Press.
- Sprangers, M. A. G. & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research. A theoretical model. *Social Science & Medicine*, 48(11), 1507–1515.
- Steuer, G., Engelschalk, T., Jöstl, G., Roth, A., Wimmer, B., Schmilz, B. et al. (2015). Kompetenzen zum selbstregulierten Lernen im Studium. Ergebnisse der Befragung von Expert(inn)en aus vier Studienbereichen. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik, Beiheft, Bd. 61, S. 203–225). Weinheim: Beltz Juventa.
- Stokes, D. E. (1997). *Pasteur's quadrant. Basic science and technological innovation*. Washington, D.C.: Brookings Institution Press.
- Terhart, E. (2002). *Standards für die Lehrerbildung: eine Expertise für die Kultusministerkonferenz*. Institut für Schulpädagogik und Allgemeine Didaktik.
- Thiel, F. & Böttcher, F. (2014). Modellierung fächerübergreifender Forschungskompetenzen - Das RMKR-W-Modell als Grundlage der Planung und Evaluation von Formaten forschungsorientierter Lehre. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre* (I 2.10, S. 109–124). Berlin: Raabe.

- Thoren, K., Wißmann, J., Harks, M., Wenger, M., Kinder, A. & Hannover, B. (in Druck). Förderung von Datennutzungskompetenzen in der Lehramtsausbildung. Konzeption und Evaluation dreier Seminare. In I. Gogolin, B. Hannover & A. Scheunpflug (Hrsg.), *Evidenzbasierung in der Lehramtsausbildung* (Edition der Zeitschrift für Erziehungswissenschaft).
- Trautwein, U. & Lüdtke, O. (2007). Epistemological beliefs, school achievement, and college major: A large-scale longitudinal study on the impact of certainty beliefs. *Contemporary Educational Psychology*, 32(3), 348–366.
- Trempler, K., Hetmanek, A., Wecker, C., Kiesewetter, J., Wermelt, M., Fischer, F. et al. (2015). Nutzung von Evidenz im Bildungsbereich. Validierung eines Instruments zur Erfassung von Kompetenzen der Informationsauswahl und Bewertung von Studien. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik, Beiheft, Bd. 61, S. 144–166). Weinheim: Beltz Juventa.
- Valter, K. & Akerlind, G. (2010). Introducing students to ways of thinking and acting like a researcher: A case study of research-led education in the sciences. *International Journal of Teaching and Learning in Higher Education*, 22(1), 89–97.
- Van Boxtel, C. & van Drie, J. (2018). Historical reasoning: The interplay of domain-specific and domain-general aspects. In F. Fischer, C. A. Chinn, K. F. Engelmann & J. Osborne (Hrsg.), *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge* (S. 142–161). New York, NY: Routledge.
- Van der Scheer, E. A., Glas, C. A.W. & Visscher, A. J. (2017). Changes in teachers' instructional skills during an intensive data-based decision making intervention. *Teaching and Teacher Education*, 65, 171–182.
- Wecker, C. & Fischer, F. (2014). Lernen in Gruppen. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie. Mit Online-Materialien* (Psychologie 2014, 6., vollständig bearb. Aufl., S. 277–296). Weinheim: Julius Beltz.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Seattle, WA: Hogrefe & Huber.
- Wenglein, S., Bauer, J., Heininger, S. & Prenzel, M. (2015). Kompetenz angehender Lehrkräfte zum Argumentieren mit Evidenz. Erhöht ein Training von Heuristiken die Argumentationsqualität? *Unterrichtswissenschaft*, 43(3), 209–224.

- Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für differentielle und diagnostische Psychologie*, 15(4), 185–200.
- Willison, J. (2012). When academics integrate research skill development in the curriculum. *Higher Education Research & Development*, 31(6), 905–919.
- Willison, J. & O'Regan, K. (2007). Commonly known, commonly not known, totally unknown: A framework for students becoming researchers. *Higher Education Research & Development*, 26(4), 393–409.
- Winter-Hölzl, A., Wäschle, K., Wittwer, J., Watermann, R. & Nückles, M. (2015). Entwicklung und Validierung eines Tests zur Erfassung des Genrewissens Studierender und Promovierender der Bildungswissenschaften. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik, Beiheft, Bd. 61, S. 185–202). Weinheim: Beltz Juventa.
- Winter-Hölzl, A., Watermann, R., Wittwer, J. & Nückles, M. (2016). Warum schreiben Promovierende bessere Abstracts als Studierende? Genrewissen schlägt Textverständnis und Forschungskompetenz. *Unterrichtswissenschaft*, 44(1), 7–24.
- Wissenschaftsrat (2006). *Empfehlungen zur künftigen Rolle der Universitäten im Wissenschaftssystem*. Berlin. Zugriff am 03.06.2019. Verfügbar unter <https://www.wissenschaftsrat.de/download/archiv/7067-06.pdf>
- Wissenschaftsrat (2008). *Empfehlungen des Wissenschaftsrates zur Qualitätsverbesserung von Lehre und Studium*. Berlin. Zugriff am 03.06.2019. Verfügbar unter <https://www.wissenschaftsrat.de/download/archiv/8639-08.pdf>
- Wolf, R., Zahner, D., Kostoris, F. & Benjamin, R. (2014). *A case study of an international performance-based assessment of critical thinking skills*. Council for Aid to Education.
- Zlatkin-Troitschanskaia, O., Pant, H. A., Kuhn, C., Toepper, M. & Lautenbach, C. (2016). *Messung akademisch vermittelter Kompetenzen von Studierenden und Hochschulabsolventen. Ein Überblick zum nationalen und internationalen Forschungsstand* (Edition ZfE, Band 1). Wiesbaden: Springer VS.

Kapitel 3

Evaluating research-oriented teaching: a new instrument to assess university students' research competences

Anmerkung: Dies ist die Version der Autorin von einer Arbeit, die in *Higher Education* veröffentlicht wurde (© by Springer Science+Business Media Dordrecht 2017). The final publication is available at <http://link.springer.com>

Böttcher, F. & Thiel, F. (2018). Evaluating research-oriented teaching: A new instrument to assess university students' research competences. *Higher Education*, 75 (1), 91–110. <https://doi.org/10.1007/s10734-017-0128-y>

3 Evaluating research-oriented teaching: a new instrument to assess university students' research competences

Abstract

Several concepts have been developed to implement research-oriented teaching in higher education in the last fifteen years. The definition of research competences, however, has received minor attention so far. Some approaches to modelling research competences describe these competences along the research process but either focus on a specific academic discipline and / or specific facets, or offer no empirical verification for cross-disciplinary approaches. We have developed a new approach to model research competences across various academic disciplines in form of a competence model – the **RMRC-K-model**. This model comprises five dimensions: skills in reviewing the state of research, methodological skills, skills in reflecting on research findings, communication skills, and content knowledge. In this study we present an instrument (*R-Comp*) developed on the basis of this cross-disciplinary competence model. The factorial structure of the *R-Comp* was examined with data from 391 university students in three groups, either enrolled in a Bachelor's degree, Master's degree, or a PhD program. The sample represented various academic disciplines. Results from confirmatory factor analyses supported the hypothesized structure of the *R-Comp* for the five dimensions in accordance with the RMRC-K-model. Additionally, results provided evidence for a more detailed differentiation of all dimensions with sub-dimensions reflecting theorized facets of the model. Both, Cronbach's α and construct reliability H indicated overall good reliability. Despite limitations to this study, there is some first evidence of a valid and reliable assessment of student research competences.

Keywords: research competences, research-oriented teaching, competence model, confirmatory factor analysis

3.1 Introduction

The purpose of this study is to introduce a reliable and valid instrument for the evaluation of *research-oriented teaching* in higher education in various academic disciplines. For this purpose, we examine the theoretically proposed factor structure of a newly developed instrument designed to assess university students' acquisition of **research competences** – *R-Comp* – on the basis of a cross-disciplinary competence model: RMRC-K-model (Thiel & Böttcher, 2014).

In order to strengthen the research-teaching nexus in higher education (Brew, 2006; Griffiths, 2004; Healey, 2005), many universities have established initiatives (Albrecht et al., 2003; BAK, 2009; Brew, 2006, Huber, 2009; Kossek, 2009; Tremp, 2005), which are generally targeted at linking research and university teaching. Mick Healey (2005) has categorized four different ways for that research-teaching linkage: a) research-tutored teaching, b) research-led teaching, c) research-based teaching, and d) research-oriented teaching. In case of a) research-tutored teaching, students are provided with the opportunity to write and discuss research papers in specific courses or tutorials; in b) research-led teaching, current research is integrated into university lectures; in c) research-based teaching, students carry out own research projects; in d) research-oriented teaching, students participate in current research projects in order to acquire methodological competences. University students benefit from those approaches to strengthen the research-teaching nexus in various ways, for example, by gaining a better understanding of their discipline and by becoming inspired by research (Visser-Wijnveen et al., 2015), as well as by feeling more motivated to engage in research activities if those teaching arrangements are perceived as valuable (Smyth et al., 2016). In our opinion and in contrast to Healey's (2005) categorization, the term *research-oriented teaching* generally comprises all of those approaches that provide students with the opportunity to develop their research competences by experiencing research practice in university teaching.

Competence models should form the basis for the development of research-oriented teaching arrangements as well as for the development of instruments to evaluate the acquisition of research competences. In practice, however, competences are rarely described systematically.

There are only a few approaches to modelling research competences so far: some are subject-specific, for example in the educational sciences (Educational Research

Literacy – ERL; Groß Ophoff et al., 2014), biosciences (Valter & Akerlind, 2010) or engineering (Academic Competences Quality Assurance – ACQA; Meijers et al., 2005). They describe research competences in their subject area oriented to the logic of the research process, partly defining competence levels (Meijers et al. 2005). Other approaches focus on specific aspects of the research process, for instance *scientific writing* (Blömeke & Zlatkin-Troitschanskaia, 2013), or *scientific literacy* (Chang et al., 2011). The subject-specific approaches as well as the approaches which focus on specific aspects cannot be applied for a cross-disciplinary approach to modelling research competences, because their concepts cannot readily be transferred to other disciplines. There are generic competence approaches (Ginns et al., 2009; Marsh et al., 2002) which, for example, include problem-solving or analytical thinking. These competences are relevant for aspects of studying or graduating in general, but, however, are not specific to the area of *research*. Furthermore, there is one cross-disciplinary model (Research Skill Development Framework – RSD; Willison & O'Regan, 2007), which describes research competences along the research process, too, but additionally, at different levels. However, this approach is lacking distinct competence dimensions and offers no empirical verification of its validity. Hence, it is also not directly applicable for the assessment of students' research competences. We see the need for the development of a new model of students' research competences, which describes them across various academic disciplines and which offers distinct competence dimensions. In the following, we present our approach to model research competences across various academic disciplines. We chose a systematization and operationalization along the research process as well, and, additionally, based it on the pragmatic theory of science (Butts, 1991; Dewey, 1938; Mittelstraß, 1991).

3.1.1 A new Approach to Modelling Research Competences

Our new approach to modelling research competences is founded upon a definition of research as “a systematic, methodologically controlled, and transparent process with the aim of generating new knowledge” (own translation, Thiel & Böttcher, 2014, p. 118). We combine several aspects in the conceptualization of our model: first, it has a cross-disciplinary orientation as it is based on the pragmatic theory of science (Butts, 1991; Dewey, 1938; Mittelstraß, 1991). Second, it is systematized and operationalized along the research process in accordance with the previously presented approaches. Third, it includes dimensions of skills as well as a knowledge dimension, because it reflects theoretical considerations about the concept of competence (Klieme et al., 2008; Mayer, 2003; Simonton, 2003; Weinert, 2001). In the following, we discuss each aspect in greater detail.

The pragmatic theory of science states that the unity of science lies in the unity of academic research practice (Mittelstraß, 1991) and that “the sciences are unified [...] by the method” (Butts, 1991, p. 33). The central argument of this pragmatic point of view is the cross-disciplinary concept of method: “Method concerns the way and means by which claims to validity in science are justified” (Mittelstraß, 1991, p. 19). According to Butts (1991) research practice across various academic disciplines has in common that the aim of research is generating new knowledge on the basis of empirical data. Furthermore, research practice has a “descriptive nature of all scientific work” (p. 33) which means that the process of research should be recorded and described.

In accordance with this theoretical approach, the research process shares common characteristics across various disciplines – even if the research methodology is different. First, every research project has to be based on the current state of research in the subject area. Second, new insights should be generated by using accepted and systematically controllable methods. Third, results and methods should be published/ presented to the scientific community according to the social nature of research knowledge (Bauer et al., 2013). This unifying research practice results in the definition of research (p.3) as a systematic, methodologically controlled and transparent process with the aim of generating new knowledge.

Besides the method, the academic researchers' attitude of “will to inquire, to examine, to discriminate, to draw conclusions only on the basis of evidence after taking pains to gather all available evidence” (Dewey, 1938, p. 31 in Butts, 1991, p. 34f.) is also a unifying characteristic of research practice. This attitude leads to a reflective perspective on research practice. Hence, reflecting on research is also an important task during the research process and should be included in the definition of research. Therefore, research can be understood as a systematic, methodologically controlled, and transparent process, which needs to be reflected on, and which is aimed at generating new knowledge.

We described common characteristics of research practice so far. In the following, we sum up these characteristics in a set of skills required to carry out research across various academic disciplines:

- systematically reviewing previous relevant findings regarding a topic, evaluating the quality of previous relevant publications, identifying research needs (*skills in reviewing the state of research*),

- investigating the research question in a methodologically controlled and systematic manner, systematically operationalizing relevant aspects of the object of research, and systematically analyzing collected information (*methodological skills*),
- reflecting on limitations, as well as theoretical, ethical, and practical implications (*skills in reflecting on research findings*),
- presenting (oral and written) research findings according to professional standards, in a way that makes the methodological approach transparent for the members of the scientific community (*communication skills*).

Additionally, besides this set of skills we conceptualized a dimension of *content knowledge* which includes knowledge of central theoretical constructs, methods, and disciplinary standards for presenting research findings. We integrated this dimension in our model because knowledge is an essential requirement for the acquisition of research competences. With this approach we are in line with the ACQA-method (Meijers et al., 2005), in which a knowledge dimension was conceptualized as well.

Finally, the conceptualization of our model draws on theoretical considerations about the two most prevalent concepts of competence in Germany (Klieme et al., 2008; Weinert, 2001). The first approach to discuss the concept of competence is one of the most popular in Germany: the concept of competence as action competence by Weinert (2001), according to whom competence is defined as “cognitive, motivational, and social prerequisites necessary and/or available for successful learning and action” (p. 51). According to this concept, competence consists of cognitive as well as motivational and social dispositions. The cognitive dispositions are represented by *abilities* and *skills* (Weinert, 2001). Weinert emphasizes that “basic mental abilities” (Weinert, 2001, p. 57), which are innate, have to be distinguished from skills, which are acquired by learning. Nevertheless, both, abilities and skills as cognitive dispositions are more or less stable over time and make action possible. In contrast to that, the motivational and social dispositions may vary over time and influence the direction as well as the dynamics of action (Birkelbach, 2005). Weinert (2002) himself emphasizes the need to separately assess cognitive and motivational aspects of competence in order to be able to examine reciprocal influences.

The second significant approach focuses on cognitive dispositions: In the functional-pragmatic concept of competence (Klieme et al., 2008), competences are defined “as context-specific *cognitive* dispositions that are acquired by learning and needed to successfully cope with certain situations or tasks in specific domains” [emphasis in the original] (Klieme et al., 2008, p. 9). According to this conception, cognitive dispositions do explicitly not include motivational or volitional aspects. Additionally, as results of learning processes, these cognitive dispositions can be differentiated from basic mental abilities (Weinert, 2001) such as intelligence. These dispositions are applicable in specific situations and are thereby context-specific. The specificity of domains implies that different competences need to be applied in mathematics, reading, or research practice. Hence, competences have to be modelled separately for each domain.

Cognitive dispositions have been further elaborated on by Mayer (2003) and Simonton (2003). Mayer (2003) characterizes competence as “specialized knowledge” (p. 265) consisting of *declarative* and *procedural* knowledge. Declarative knowledge is represented by knowledge about “facts and concepts” (p. 268), whereas procedural knowledge is represented by knowledge of how to do something. However, Simonton (2003) declares competences as a composite of “any acquired skill or knowledge that constitutes an essential component for performance or achievement in a given domain” (p. 230). In this regard, competence consists of *skills* and *knowledge*. Nevertheless, in both approaches, cognitive dispositions are conceptualized as including a *knowledge* part and as including a disposition that leads to performance.

Our approach to modelling student research competences draws on aspects of Klieme et al.’s (2008) conceptualization of competence as well as Simonton’s (2003) conceptualization of cognitive dispositions. Taking the functional-pragmatic concept of competence by Klieme et al. (2008) into account, we define *research* competences as cognitive dispositions, which are (a) acquired by learning in university courses, (b) applied in the context of carrying out the research process, and (c) needed to successfully cope with tasks arising from research-oriented teaching arrangements. Such tasks, for example, are participating in a current research project, where students apply research methods or present their findings to a research colloquium. Our concept of cognitive dispositions is in line with Simonton’s (2003) conceptualization of competences as skills *and* knowledge.

Based on our theoretical considerations, we present our model's dimensions in the following section. For each dimension, we further derived corresponding facets for skills as well as for *content knowledge*, which are presented in the following section, too.

3.1.2 RMRC-K-Model and Construction of a new Instrument

On the basis of the three essential aspects - the pragmatic theory of science, the modelling along the research process, and the conceptualization of competences as skills and knowledge – we derived a structural competence model: the **RMRC-K-model** (Thiel & Böttcher, 2014; Figure 3.1) with four dimensions of skills and one dimension of content knowledge, which are further broken down into facets as follows:

Skills in reviewing the state of research:

- systematically reviewing the state of research,
- evaluating relevant literature,
- identifying contradictory findings,
- identifying research need.

Methodological skills:

- formulating and operationalizing research questions / hypotheses,
- planning the research process,
- selecting appropriate research methods,
- applying adequate methods.

Skills in reflecting on research findings:

- reflecting on implications of research results,
- reflecting on methodological limitations,
- reflecting on practical implications,
- reflecting on ethical implications.

Communication skills:

- writing academic publications,
- presenting research findings.

Content knowledge of:

- central / key theories,
- central research methods,
- previous findings,
- standards of communication in academic research.

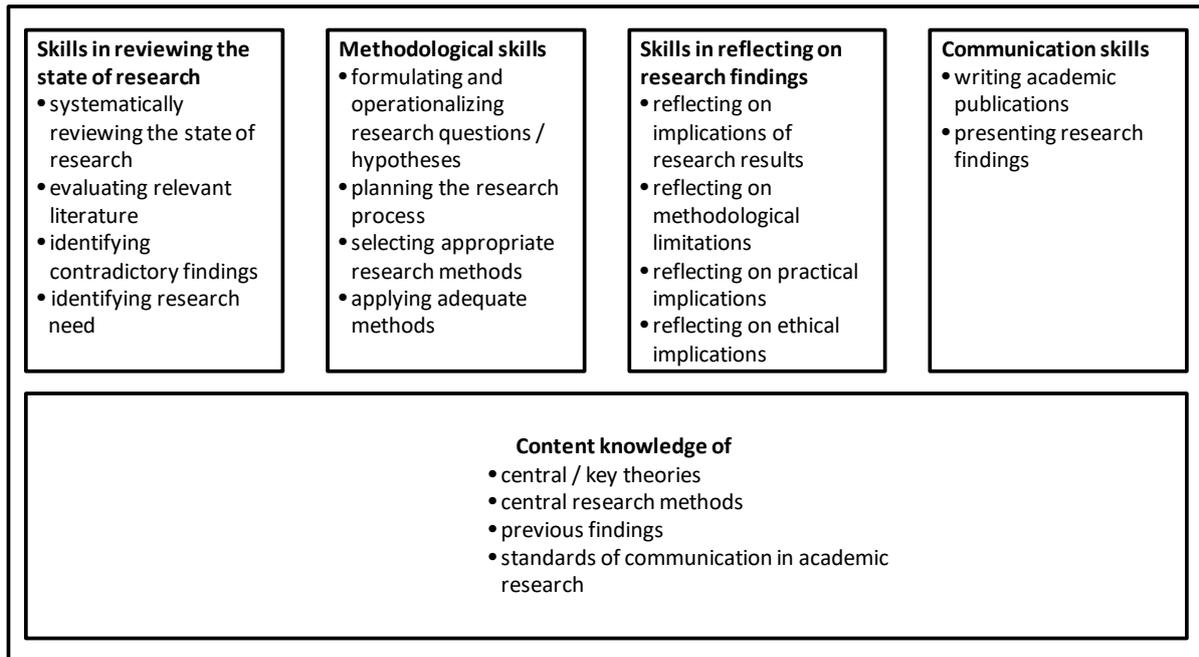


Figure 3.1 The RMRC-K-model

On the basis of the RMRC-K-model, we developed items for an instrument to assess student **research competences** via self-ratings: the *R-Comp*. In item generation, the cross-disciplinary aspect of assessing research competences had to be carefully met. These various academic disciplines have specific research terminologies. We alternated those specific research terms, for example “database”, “materials”, and “sources” as options for the research object, to phrase the items in the most cross-disciplinary way. Furthermore, some items were negatively phrased to avoid an acquiescence bias (Diekmann, 1995, p. 413).

The first version consisted in total of 44 items across the five aforementioned dimensions: skills in reviewing the state of research (six items), methodological skills (11 items), skills in reflecting on research findings (eight items), communication skills (four items), and content knowledge (15 items).

An expert validation (Langer, 2000) was then conducted featuring 12 experts from the social sciences, natural sciences, and humanities. They evaluated the phrasing and relevance of each item for their academic discipline by selecting either the response “true for my discipline” or “not true for my discipline”. The experts had the opportunity to propose an alternative phrasing for inappropriate items, and moreover, to propose further competence dimensions. No further dimensions were proposed indicating a successful conceptualization of dimensions in the RMRC-K-model. Few critical comments were given, which were addressed by adapting item phrasing, deleting items, and developing new items. The resulting version consisted of 43 items.

In order to provide a short and economic instrument, in a second step we further selected 32 out of 43 items on the basis of recommendations by Bühner (2011). Accordingly, items were retained which, according to their psychometric properties, and content, best represented the corresponding dimensions.

3.1.3 The Current Study

The purpose of this study is to examine whether the theoretically proposed structure of the *R-Comp* can be empirically confirmed. In order to evaluate the factorial validity, we chose a confirmatory rather than an exploratory approach. Our primary aim is to verify the *R-Comp*'s a priori proposed structure with five scales representing the five dimensions. This structure is a result of our considerations of the RMRC-K-model. Only if the evaluation of the confirmatory approach was not successful, we still had the option for exploratory factor analyses.

In the following section we present the study's sample and methodological considerations of our study. After presenting the results of our analyses, we will discuss these, and furthermore, the limitations of our study.

3.2 Methods

3.2.1 Participants and Data Collection

In order to evaluate the cross-disciplinary approach of the RMRC-K-model, we recruited participants from various academic disciplines at a German university. We also aimed to collect data in three different stages of participants' study programs to examine differences between these stages. Data collection was carried out in lectures on a voluntary basis.

Data were collected from 391 students (69.1% female, 24.2% male) in 12 university courses of a German university representing different academic disciplines (Table 3.1): 80.8% social sciences, 12.0% agriculture sciences, 7.2% natural sciences. The majority of students were in a Master's degree program (68.5%), 27.4% in a Bachelor's degree program⁹, and 4.1% in PhD programs.

Data were collected either in a paper-pencil format (97%) or an online format (3%). The online survey was added, because one sub-sample did not have regular courses any more. Prior to the survey, participants were informed about the study's aims, the voluntary nature of the survey, and their anonymity was guaranteed.

Table 3.1 Overview of sub-samples

Subject Area	Discipline	Degree Program	Data Collection Format	<i>n</i>
Social Sciences				316
	Business Informatics	Master	Paper-Pencil	30
	Educational Science	Bachelor	Paper-Pencil	71
	Educational Science	Master	Paper-Pencil	57
	Latin American Studies	PhD	Paper-Pencil	5
	Psychology	Master	Paper-Pencil	55
	Teacher Training Program	Master	Paper-Pencil	98
Agricultural Sciences				47
	Biomedical Science	PhD	Online	11
	Veterinary Medicine	Bachelor ¹	Paper-Pencil	36
Natural Sciences				28
	Biology	Master	Paper-Pencil	5
	Chemistry	Master	Paper-Pencil	17
	Physics	Master	Paper-Pencil	6
		Bachelor		107
		Master		268
		PhD		16

Notes: ¹Participants in this sub-sample were in a state examination program. To improve international comparability, they were classified as students in a Bachelor's degree program.

⁹ This sub-sample includes participants from *Veterinary Medicine*. This study program is a state examination program, licensed and organized by national government agencies. It is separated in two sequences (first and second state examination). Participants in our sub-sample were in the first sequence, which is comparable to the level of a Bachelor's degree program. Therefore, we classified these participants as Bachelor students to gain a better understanding of our sub-sample in international contexts.

3.2.2 Measures

The newly developed *R-Comp*, based on the RMRC-K-model, was administered to participants, who rated both their skills in conducting research and their knowledge about research on a 5-point response scale, ranging from “1 – strongly disagree” to “5 – strongly agree”.

The *R-Comp* consists of 32 items, reflecting the five theoretical dimensions as follows: skills in reviewing the state of research (four items; e.g., “*I am able to systematically review the state of research regarding a specific topic.*”), methodological skills (eight items; e.g., “*I am able to decide, which data / sources / materials I need to address my research question.*”), skills in reflecting on research findings (six items; e.g., “*I am able to critically reflect on methodological limitations of my own research findings*”), communication skills (five items; e.g., “*I am able to write a publication in accordance with the standards of my discipline.*”), and content knowledge (nine items; e.g., “*I have a sound knowledge of the main research methods in my discipline.*”). German items were translated into English for this article and are reported in Table 3.4 in the results section. Responses to negatively phrased items were reversed prior to the analyses.

Participants also indicated their sex, academic discipline and number of semesters.

3.2.3 Analyses

Confirmatory factor analyses were conducted using *Mplus* (Version 6). A robust estimator (MLR) was used as recommended for small samples, incomplete and non-normal data (Satorra & Bentler, 2001; Yuan & Bentler, 2000). The percentage of missing cells of data was 3.44%. The highest amount of missing data per variable was 6.6% for the variable sex.

Model fit evaluation was carried out on the basis of a combination of fit indices indicating approximate fit. Values for exact model fit are also presented. Exact model fit occurs when the chi-square(χ^2)-associated p-value is non-significant. Nevertheless, we did not take these values into consideration when evaluating model fit, since models with small sample size, non-normal data or highly complex models are more likely to be rejected using this approach (Schermelleh-Engel et al., 2003). Evaluation of approximate model fit was carried out on the basis of the following fit indices and their corresponding cut-off values: χ^2/df values indicating acceptable fit if value ranges between 2 and 3, and good fit if value is under 2 (Schermelleh-Engel et al., 2003); the comparative fit index (CFI) with a cut-off value close to .95, the root mean square error of approximation (RMSEA) with a cut-off value close

to .06 for samples $N > 250$, and the standardized root mean square residual (SRMR) with a cut-off value close to .08 (Hu & Bentler, 1999).

To compare models, we inspected AIC and BIC values, whereby smaller AIC and BIC values indicated the best fitting model (Geiser, 2011; Schermelleh-Engel et al., 2003).

The evaluation of the model parameters focused on the standardized factor loadings. According to Kline (2005), loadings above .20 are said to be substantial loadings on their factor. Additionally, high factor loadings are indicators for convergent validity (Kline, 2005).

In order to examine the psychometric properties of the newly developed instrument, we evaluated internal consistency of all manifest scales (dimensions) inspecting Cronbach's α . Additionally, as we examined the latent factor structure of the *R-Comp* and α might be a misestimator in latent modelling (Brown, 2006), we also calculated the construct reliability H (Hancock & Mueller, 2001), which indicates how well each latent factor is measured by its corresponding indicators. H -values range between 0 and 1 with values above .70 considered desirable. H -values indicate how much variance in the latent construct is accounted for by its indicators (Hancock & Mueller, 2001). Since Cronbach's α is also a widely used value, we present both indicators.

Models tested

Two plausible models were evaluated to confirm the proposed structure of the *R-Comp*. First, a *5-factor model* was tested according to the five hypothesized dimensions (*skills in reviewing the state of research, methodological skills, skills in reflecting on research findings, communication skills, and content knowledge*) as latent factors¹⁰. All indicators were specified as single loadings on their corresponding dimension.

Second, a *5-factor model with sub-dimensions* was tested. Based on theoretical considerations, all indicators were assigned to sub-dimensions. These sub-dimensions group similar aspects of each skill and knowledge dimension together, reflecting and partly combining the theorized facets of the model¹¹. Hence, a hierarchical model with five

¹⁰ The skill dimensions as well as the knowledge dimension were specified as factors at the same hierarchical level in the empirical model. In pre-studies (Böttcher & Thiel, 2016), we examined whether the skills and knowledge dimensions should be modelled separately or in a combined model. Our results indicated that a combined model in line with Simonton's (2013) conceptualization of competence was the more appropriate solution.

¹¹ In the empirical model, some theorized facets had to be combined on the basis of theoretical considerations about the research process. Otherwise, some sub-dimensions would have consisted of single indicators.

dimensions on the first level and overall twelve sub-dimensions on a second level was specified. Each sub-dimension consisted of two to four indicators specified as single loadings. Schematic illustrations of both tested models including labels for the sub-dimensions in the *5-factor model with sub-dimensions* are presented in Figure 3.2.



Figure 3.2 Schematic illustration of the five-factor model and the five-factor model with sub-dimensions

3.3 Results

The *5-factor model* showed only one indicator for approximate model fit, the SRMR value, which was below the cut-off value (Table 3.2).

Table 3.2 Model fit for the tested models

Model	χ^2	df	p	χ^2/df	CFI	RMSEA	SRMR	SCF	AIC	BIC
5-factor model	1235.92	454	< .001	2.72	.892	.066	.055	1.126	29747.95	30168.63
5-factor model with sub-dimensions	784.78	442	< .001	1.78	.953	.045	.038	1.116	29255.95	29724.26
Higher-order 5-factor model with sub-dimensions	800.77	447	< .001	1.79	.951	.045	.040	1.117	29264.69	29713.16
Bi-factor model	1076.90	432	< .001	2.49	.911	.062	.042	1.099	29583.46	30091.46

Notes: CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = Standardized Root Mean Square Residual; SCF = Scaling Correction Factor; AIC = Akaike information criterion; BIC = Bayesian information criterion.

The *5-factor model with sub-dimensions* showed approximate fit. Both AIC and BIC values indicated that this model fitted better than the *5-factor model*. Inter-factor correlations ranged between .76 and .87 (Table 3.3). According to recommendations by Brown (2006), inter-factor correlations close to .80 or above indicate that the model should be respecified as a higher-order factor model. Therefore, we tested a third model: the *higher-order 5-factor model with sub-dimensions*, consisting of a higher-order factor that could be termed “*research competences*” on the first level, the four skills and the knowledge dimensions on the second level, and the twelve sub-dimensions on the third level. A schematic illustration of this model is shown in Figure 3.3.

Table 3.3 Latent inter-factor correlations

Dimensions of research competences	1.	2.	3.	4.	5.
1. Skills in reviewing the state of research	-				
2. Methodological skills	.87	-			
3. Skills in reflecting on research findings	.82	.87	-		
4. Communication skills	.82	.82	.82	-	
5. Content knowledge	.85	.81	.76	.85	-

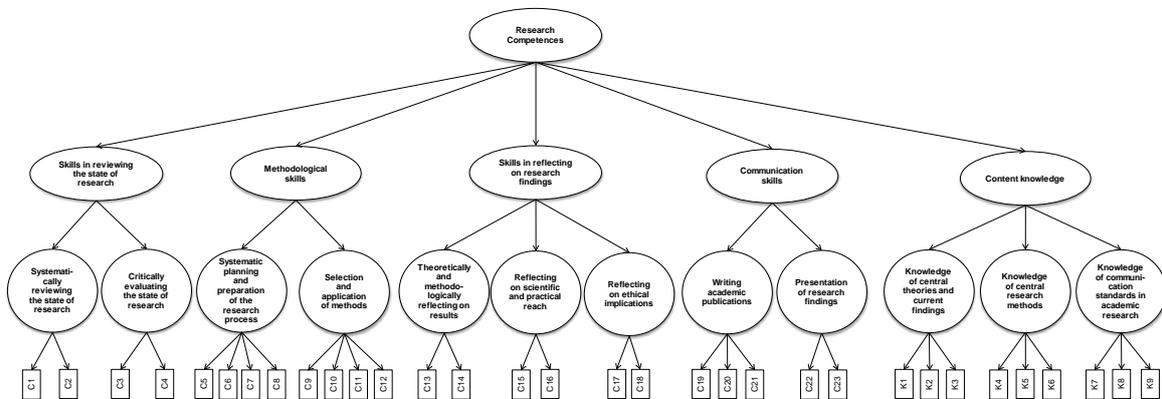


Figure 3.3 Schematic illustration of the higher-order five-factor model with sub-dimensions

The *higher-order 5-factor model with sub-dimensions* showed approximate fit (Table 3.2), which was almost as good as the *5-factor model with sub-dimensions*. However, the AIC values indicated that this more restricted model did not fit the data better than the *5-factor model with sub-dimensions*. Only the BIC values indicated that the higher order solution fitted the data better. Taking into consideration the combination of fit indices indicating approximate fit as well as the AIC value, the *5-factor model with sub-dimensions* should be preferred over the higher-order solution. Since our decision to favor one model and

reject the other was so close, two plausible bi-factor models¹² were then tested. Those models are an alternative way to model higher-order factor structures (Reise et al., 2010).

Hence, a fourth model was specified as an alternative way to model the *higher-order 5-factor model with sub-dimensions*. In this model, the five dimensions and the above identified sub-dimensions were considered in a bi-factorial structure. This *bi-factor model with sub-dimensions* consisted of a general factor “research competences”, five uncorrelated specific factors with sub-dimensions representing the four skills dimensions and the knowledge dimension as well as their corresponding facets. However, this model could not be computed due to convergence problems: the model was too complex (Brown, 2006) as it consisted of the largest number of freely estimated parameters compared to the above mentioned models. Hence, it cannot be considered an alternative to the higher-order model. For that reason, we specified a fifth model consisting of a general factor “research competences” and five uncorrelated specific factors according to the five hypothesized dimensions – the *bi-factor model*. This model represents an alternative way to model the five-factorial structure in the *5-factor model*. As the number of freely estimated parameters was smaller than in the *bi-factor model with sub-dimensions*, no convergence problems occurred. A schematic illustration of both bi-factor models is shown in Figure 3.4.

¹² We thank an anonymous reviewer for his / her suggestion of testing bi-factor models.

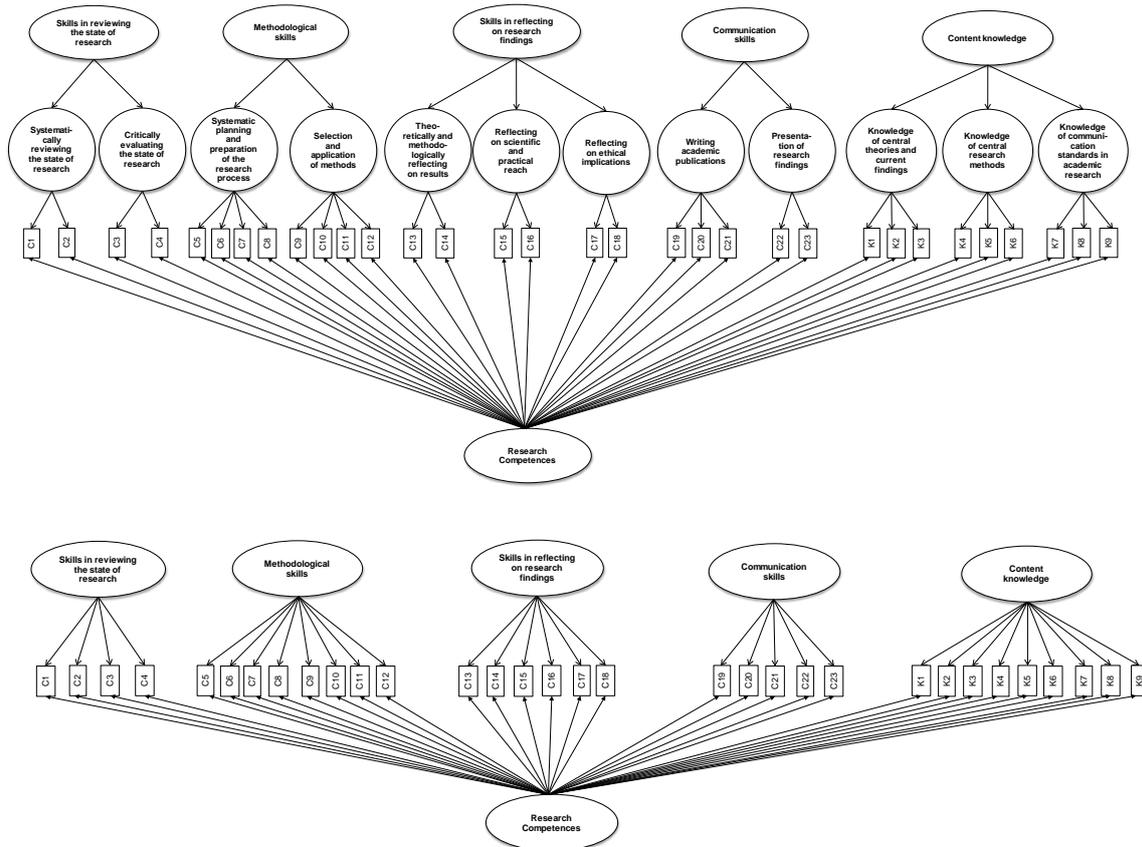


Figure 3.4 Schematic illustration of the bi-factor model with sub-dimensions and the bi-factor-model

The *bi-factor model* did not fit the data better than the *higher-order 5-factor model with sub-dimensions* or the *5-factor model with sub-dimensions* as indicated by the fit indices for approximate model fit as well as by the AIC and BIC values (Table 3.2). Thus, after taking into account all results we decided in favor of the *5-factor model with sub-dimensions*.

For the *5-factor model with sub-dimensions*, all observed variables (indicators) had significant factor loadings close to .50 or above. Indicators of the negatively phrased items showed the lowest factor loadings. Factor loadings from the latent sub-dimensions to the latent dimensions were significant with values above .77. Manifest means for the indicators vary between 2.37 and 3.67, and for the sub-dimensions between 2.80 and 3.58. All standardized factor loadings, as well as manifest means, are presented in Table 3.4.

The construct reliability values ranged between .77 and .89 for the sub-dimensions, and between .93 and .99 for latent factors (dimensions) indicating good to excellent construct reliability. Over 77% of the variability in the latent sub-dimensions was accounted for by the

composite of indicators, and over 93% was accounted for by the composite of latent factors. Cronbach's α for all sub-scales (sub-dimensions) ranged from .73 to .87, and for all scales (dimensions) from .87 to .92, indicating good internal consistency. All values for reliability are presented in Table 3.4.

Table 3.4 Factor loadings, means, standard deviations, values for construct reliability, and values for Cronbach's α of the dimensions and sub-dimensions

Dimensions and sub-dimensions	λ	$M (SD)$
Skills in reviewing the state of research ($H = .99$; $\alpha = .87$)		
<i>Systematically reviewing the state of research</i> ($H = .90$; $\alpha = .87$)	.91	3.58 (1.10)
I know how and where to target a search of the state of research regarding a specific topic.	.83	3.67 (1.16)
I am able to systematically review the state of research regarding a specific topic.	.93	3.48 (1.19)
<i>Critically evaluating the state of research</i> ($H = 0.76$; $\alpha = .73$)	.99	2.93 (0.95)
Based on the state of research, I am able to identify gaps/unaddressed questions for further research.	.83	3.05 (1.13)
I can evaluate the methodological quality of researched findings well.	.70	2.81 (0.97)
Methodological skills ($H = .95$; $\alpha = .88$)		
<i>Systematic planning and preparation of the research process</i> ($H = .80$; $\alpha = .75$)	.96	3.31 (0.82)
I find it difficult to formulate specific research questions / hypotheses. ^(a)	.55	3.35 (1.06)
I am able to decide, which data / sources / materials I need to address my research question.	.77	3.53 (1.00)
I am able to plan a research process.	.81	3.17 (1.18)
I find it difficult to operationalize each step of the research process. ^(a)	.48	3.18 (1.08)
<i>Selection and application of methods</i> ($H = .88$; $\alpha = .87$)	.93	2.88 (0.89)
I find it easy to decide, which methods I need to use to examine a specific research topic.	.80	2.87 (1.00)
I am good at judging which method is inappropriate to answer a specific research question.	.82	3.07 (1.00)
I can apply different research methods appropriate to my research question.	.85	3.06 (1.03)
I can confidently apply even complex methods to analyze data / sources / materials.	.73	2.51 (1.11)
Skills in reflecting on research findings ($H = .97$; $\alpha = .92$)		
<i>Theoretically and methodologically reflecting on results</i> ($H = .88$; $\alpha = .86$)	.99	3.35 (1.04)
I am able to adequately interpret my own research findings by relating them to key theories in the subject area.	.91	3.26 (1.07)
I am able to critically reflect on methodological limitations of my own research findings.	.84	3.43 (1.14)
<i>Reflecting on scientific and practical reach</i> ($H = .87$; $\alpha = .87$)	.98	3.18 (1.05)
I am able to reflect on the implications of my own research findings on my discipline.	.90	3.17 (1.10)
I am able to discuss my research findings with regard to their potential applications.	.86	3.20 (1.14)
<i>Reflecting on ethical implications</i> ($H = .88$; $\alpha = .86$)	.77	3.37 (1.07)
I am able to critically reflect on the social / ethical implications of my research.	.91	3.29 (1.14)
I am able to take a stand on social / ethical issues of research in my discipline.	.82	3.46 (1.15)
Communication skills ($H = .93$; $\alpha = .89$)		
<i>Writing academic publications</i> ($H = .89$; $\alpha = .85$)	.95	3.17 (1.17)
I can write up research findings in accordance with the current conventions in my discipline.	.88	3.31 (1.27)
I am able to write a publication in accordance with the standards of my discipline.	.90	2.94 (1.37)
I find it difficult to write a report that meets the standards of academic writing. ^(a)	.67	3.27 (1.33)
<i>Presentation of research findings</i> ($H = .85$; $\alpha = .85$)	.86	2.80 (1.23)

I am able to prepare research findings for a presentation at a research colloquium.	.87	3.04 (1.33)
I am able to present my research at a scientific meeting in accordance with current standards in my discipline.	.85	2.54 (1.27)
Content knowledge ($H = .97$; $\alpha = .88$)		
<i>Knowledge of central theories and current findings</i> ($H = .77$; $\alpha = .74$)	.83	3.00 (0.79)
I have a good overview of the main (current) research findings in my discipline.	.72	2.84 (0.93)
I am informed about the main (current) theories in my discipline.	.79	3.15 (0.92)
I am informed about the history of theory / paradigm shifts in my discipline.	.63	3.00 (1.05)
<i>Knowledge of central research methods</i> ($H = .88$; $\alpha = .87$)	.80	3.01 (0.91)
I have a sound knowledge of the main research methods in my discipline.	.80	3.33 (1.01)
I would describe my methodological knowledge as sophisticated and comprehensive.	.86	2.81 (1.02)
I am very familiar with different research methods in my subject area.	.84	3.07 (1.02)
<i>Knowledge of communication standards in academic research</i> ($H = .83$; $\alpha = .80$)	.89	2.71 (1.06)
I am informed about the most important national and international academic publication outlets in my discipline.	.72	2.71 (1.19)
I am informed about the standards for academic publications that apply in my discipline.	.86	3.08 (1.32)
I am informed about the standards that apply to the presentation of research findings at congresses and meetings in my subject area.	.70	2.37 (1.25)

Notes: **Bold** print indicates the model's dimensions. *Italic* print indicates the model's sub-dimensions. Normal print refers to the model's indicators. λ = standardized factor loadings; M = manifest means; SD = standard deviations; H = construct reliability; α = Cronbach's alpha; ^(a) Recoded items for analyses.

3.4 Discussion

The study's main purpose was to validate the factorial structure of the *R-Comp*, an instrument designed to assess university students' research competences on the basis of a cross-disciplinary competence model. The results of the confirmatory factor analyses supported the hypothesized structure of the *R-Comp* for the five dimensions in accordance with the RMRC-K-model. According to this model, students' research competences comprise *skills in reviewing the state of research*, *methodological skills*, *skills in reflecting on research findings*, *communication skills*, and *content knowledge*. Moreover, the results provided evidence for a more detailed differentiation of all dimensions with a total of twelve sub-dimensions. These sub-dimensions partly combine and reflect the facets specified in the model.

For the dimension *skills in reviewing the state of research* with four theoretical facets, the data only showed two sub-dimensions. These sub-dimensions represent two different phases of reviewing the state of research: the phase of searching and reviewing the state of research on the one hand, and critically evaluating the state of research on the other hand, including the evaluation of prior research and the identification of future research questions.

For the four theoretical facets of the *methodological skills* dimension, the data showed two sub-dimensions in the empirical model representing two phases: The first phase includes the planning and preparation of the research process, such as formulating and operationalizing research questions or hypotheses. The second phase includes the selection and application of methods appropriate to the research questions or hypotheses.

For the dimension *skills in reflecting on research findings* with four theoretical facets, the data showed three sub-dimensions representing three objects of reflection: reflecting (a) results within the process of research itself considering both, the study's theoretical and methodological approaches, (b) practical and theoretical implications for the subject area, and (c) ethical and social implications.

For the dimension *communication skills*, the two theoretical facets could be empirically confirmed: (a) the writing of academic publications, and (b) the presentation of research findings.

For the four theoretical facets of the *content knowledge* dimension, the data showed three empirical sub-dimensions: (a) the knowledge of central theories and current findings, (b) the knowledge of central research methods, and (c) the knowledge of communication standards in academic research.

As indicated by our findings of sub-dimensions, students in our sample seemed to perceive particular phases or aspects of the skill and knowledge dimensions. The perception of separate facets or aspects might be related to different levels of expertise (Simonton, 2003) or "levels of competency" (Mayer, 2003, p. 265). Students in our study featured three different levels of expertise. Students at the beginning of their studies can be seen as novices (Simonton, 2003) who might not have had so many opportunities to carry out the particular phases of the research process. In university courses of a Bachelor's degree program, students are introduced to single requirements of the research process in different exercises or small projects. Particular skills are encouraged step by step, for example first, searching for literature in the library and second, evaluating prior research. Students in a Master's degree program might have had more opportunities to gain practical experience by carrying out particular phases of the research process.

In contrast, students in a PhD program already carry out the entire research process as junior researchers. They are likely to be the most experienced group in our sample with the highest degree of proficiency (Simonton, 2003) or the highest level of competence (Mayer,

2003). It could be assumed that the most experienced students perceive the phases or aspects of the research process as homogenous: for example, PhD students might critically evaluate prior research *during* the process of searching and reviewing prior research. In this case, a model without sub-dimensions should represent the data from this sub-sample better. However, due to the under-representation of PhD students in our sample, we were not able to test this hypothesis.

The overall high inter-factor correlations between the dimensions in our model revealed that research competences were closely interrelated in our study. Since high inter-factor correlations are indicators of poor discriminant validity (Kline, 2005), we tested a higher-order solution. We favored the *5-factor model with sub-dimensions* due to minor advantages over the *higher-order 5-factor model with sub-dimensions*. Nevertheless, the overall good fit of the higher-order solution might also indicate that this model is more suited to particular sub-groups of students, depending on their level of expertise.

A higher-order solution would imply that skills and knowledge are linked on a higher-order level by a factor that could be termed “research competences”, which influences the whole research process. Students with a high level of research competences would also show a high level of skills and knowledge. This could be true for students in our sample with a high level of expertise. In contrast, students with a lower level of expertise might perceive the skills and knowledge required to carry out research as separate dimensions. Students who are more experienced might perceive conducting research as a coherent process in which the single aspects of the research process are linked together. It is conceivable that these students might plan and prepare the methodological approach *while* reviewing the state of research.

As an alternative way to model higher-order structures (Reise et al., 2010), two bi-factor solutions were also tested. A bi-factor solution would imply that there is a general factor “research competences”, which influences all indicators, as well as specific factors, which account for the variance that cannot be accounted for by the general factor. These specific factors would represent the skills and knowledge dimensions. Research competences could then be accounted for by an overall competence to conduct the research process. Additionally, there would be skills and knowledge which are needed for the research process and which could be separately developed.

Finally, we assume that either a higher-order solution or a bi-factor solution might be true for students with a higher level of expertise. For students with a lower level of expertise,

we assume a factor solution consisting of sub-dimensions to be the best fitting solution to model students' research competences.

The standardized factor loadings showed substantial loadings on their corresponding factor. Overall high factor loadings serve as indicators of convergent validity (Kline, 2005) indicating that the dimensions are well conceptualized by their indicators. The reliability of the *R-Comp* was evaluated in the manifest as well as in the latent approaches for the *5-factor model with sub-dimensions*. Both approaches showed good to very good reliability indicating that *R-Comp* can be considered an instrument which precisely measures the construct of research competences.

Limitations and Future Directions

Some limitations to our study need to be considered. The sub-samples in this study were too small to examine group invariance across academic disciplines. These analyses would be important in order to verify the cross-disciplinary aspect of our approach to model research competences. Hence, sampling needs to be extended to the, in our study, under-represented academic disciplines such as the natural sciences, and the, as of yet, not represented disciplines such as the humanities.

Moreover, further analyses need to be conducted focusing on the models' structure depending on the stages of study programs and thus the students' level of expertise. We assumed that some of the models we had to reject due to model fit, could be taken into consideration for students at a certain stage of their study program. Due to the under-representation of Bachelor and especially PhD students in our sample, we were unable to examine group invariance, to identify differences in model structure between the stages of study programs, and to examine correlations between the stages of study programs and scale scores. These analyses would be important to provide further evidence of validity¹³. However, before group invariance has not been demonstrated, analyses of differences between groups are premature (Stein et al., 2006). Therefore, in the future more participants need to be recruited from Bachelor's degree programs and PhD programs.

Some of our identified sub-dimensions only consist of two indicators per factor. According to Bollen and Davis (2009), who state that each latent variable should be identified by at least two indicators (the *2+ emitted path rule*), we met the necessary condition of model

¹³ Thanks to one of our anonymous reviewers for rising this point.

identification with latent variables. Due to some methodologists' recommendations of a minimum of three indicators per factor (e.g., Kline, 2005), in the future, more items could be developed to serve as additional indicators for these latent variables.

In order to avoid an acquiescence bias, some items were negatively phrased. This can lead to problems with internal consistency or factor structures¹⁴ (Barnette, 2000), which may be a reason for the lower factor loadings of negatively phrased indicators compared to other variables in our model. There is one way to deal with this problem: testing the aforementioned models based only on positively worded items¹⁵. On the basis of our present data, this would mean to test the aforementioned models without those negatively phrased items. However, this deletion would lead to a loss of insight, as those items represent essential facets of our theoretical model. Another approach would be to test our models based on *positively re-phrased* items. In this approach, there would be no loss of insight. Consequently, it can be examined, whether including only positive items would result in an improvement in model fit. As in this approach, there is the need for new data, we are not able to perform these analyses at the moment. Hence, in the future, a study with an experimental setting with two conditions would need to be conducted to test the suitability of each version. Thereby, a *R-Comp*'s version including also negatively phrased items could be tested against a *R-Comp*'s version with positively re-phrased items representing the same theoretical facets.

The *R-Comp* is an instrument for assessing student research competences via self-evaluation. To complement this subjective approach, we are planning to additionally develop a questionnaire for university lecturers. For the evaluation of research-oriented teaching arrangements it would be fruitful to compare the lecturer's intended acquisition of research competences in these research-oriented teaching arrangements to students' self-evaluation. In addition, performance tests could be integrated in the assessment of students' research competences to serve as more objective approaches of assessment.

In our conceptualization of research competences, we have not operationalized motivational, volitional, and personal orientations as recommended by Weinert (2001) in his conceptualization of action competence. Weinert (2002) proposed a separate assessment of cognitive and personal dispositions in order to evaluate their reciprocal influences. Future

¹⁴ We are very thankful that one of our anonymous reviewers has drawn attention to this important limitation of our study.

¹⁵ We appreciate our anonymous reviewer for providing this idea.

studies could be complemented by the assessment of motivational or volitional dispositions such as orientations or interest in addition to assessing cognitive dispositions. This would allow for an evaluation of these reciprocal influences from an action competence perspective.

Last but not least, we reported results of confirmatory factor analyses for the German version of the *R-Comp* in this study. In order to do so, we translated items into English for this publication. An evaluation of the factor structure of the English version is needed to be able to use the instrument in future international investigations.

Despite of the aforementioned limitations, this study can be viewed as a first step towards establishing a reliable and valid instrument for the assessment of student research competences across various academic disciplines based on the RMRC-K-model.

Implications

Our study introduces an instrument for assessing cross-disciplinary research competences of students. We were able to validate the factorial structure of the instrument on the basis of a heterogeneous sample of students at different stages of study programs representing various academic disciplines. We offer a model for students' research competences with distinct, empirically confirmed competence dimensions, which overcomes the restrictions of previous approaches to modelling research competences as outlined in the theory section. Our approach to modelling research competences has a cross-disciplinary focus as it is based on the pragmatic theory of science (Butts, 1991; Dewey, 1938; Mittelstraß, 1991), and it is systemized along the research process. Moreover, the modelling of research competences is based on the functional-pragmatic concept of competence (Klieme et al., 2008) with competences conceptualized as skills and knowledge (Simonton, 2003). The results of our study provide some evidence in favor of the hypothesis that there *is* a unifying research practice. Across various academic disciplines research can be described as a process, which is systematic, methodologically controlled, and transparent, which needs to be reflected upon, and which includes discipline-specific content knowledge.

Furthermore, our approach to modelling students' research competences is applicable to diverse contexts and purposes in higher education. On the one hand, the theoretically founded RMRC-K-model provides a basis for implementing research-oriented teaching arrangements. It spells out relevant aspects of the research process in the form of the skills and knowledge dimensions. In research-oriented teaching arrangements, students' acquisition of research competences can be supported by focusing either on some of these dimensions, or on

the research process as a whole. On the other hand, the *R-Comp* can be applied to evaluate these research-oriented teaching arrangements. On the basis of such evaluation results, the acquisition of research competences can be monitored and specifically tailored to students' needs.

Finally, our approach to modelling research competences has implications for initiatives of research-oriented teaching in higher education. Research-oriented teaching approaches enable students to improve their research competences, for example by providing them with the opportunity to participate in current research projects. With the current study, we contribute a new theory-based approach in higher education research for modelling and assessing university students' acquisition of research competences. Students, as junior researchers and as future professionals, benefit from developing research competences during their study programs in any discipline. These competences are useful in a variety of professional contexts, in which the "will to inquire, to examine, to discriminate, to draw conclusions only on the basis of evidence" (Dewey, 1938, p. 31 in Butts, 1991, p. 34f.) is a prerequisite for decision making.

3.5 References

- Albrecht, L., Arends, H., Burg, A., Jubin, B., Langer, C., Lehmann, J. et al. *FL² Forschendes Lernen – Lehrende Forschung*. Zugriff am 06.11.2013. Verfügbar unter <https://incom.org/action/open-file/131430>
- BAK (2009). *Forschendes Lernen - wissenschaftliches Prüfen. Ergebnisse der Arbeit des Ausschusses für Hochschuldidaktik* (Schriften der Bundesassistentenkonferenz, Bd. 5, Neuaufl. nach der 2. Aufl.). Bielefeld: UVW, Webler.
- Barnette, J. J. (2000). Effects of stem and likert response option reversals on survey internal consistency. If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, 60(3), 361–370.
- Bauer, W., Bleck-Neuhaus, J., Dombois, R. & Wehrtmann, I. (2013). *Forschungsprojekte entwickeln - von der Idee bis zur Publikation* (UTB, Bd. 4019, 1. Aufl.). Baden-Baden: Nomos.
- Birkelbach, K. (2005, September). *Über das Messen von Kompetenzen. Einige Überlegungen im Anschluss an ein BMBF-Projekt*. Vortrag auf der Herbsttagung der Sektion Berufs- und Wirtschaftspädagogik der DGFE 2005, Erfurt.

- Blömeke, S. & Zlatkin-Troitschanskaia, O. (2013). *Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor: Ziele, theoretischer Rahmen, Design und Herausforderungen des BMBF-Forschungsprogramms KoKoHs* (KoKoHs Working Papers, 1). Berlin & Mainz: Humboldt-Universität & Johannes Gutenberg-Universität.
- Bollen, K. A. & Davis, W. R. (2009). Two rules of identification for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 523–536.
- Böttcher, F. & Thiel, F. (2016, Juli). *Introduction of a new instrument to assess university students' research competences: R-Comp*. Vortrag auf der Higher Education Conference (HEC): The Scholarship of Learning, Teaching & Organizing, Amsterdam.
- Brew, A. (2006). *Research and teaching. Beyond the divide* (Universities into the 21st century). Houndmills, Großbritannien: Palgrave Macmillan.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research* (Methodology in the social sciences). New York, NY: Guilford Press.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). München: Pearson Studium.
- Butts, R. E. (1991). Methodology, metaphysics and the pragmatic unity of science. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 23–28). Berlin: W. de Gruyter.
- Chang, H.-P., Chen, C.-C., Guo, G.-J., Cheng, Y.-J., Lin, C.-Y. & Jen, T.-H. (2011). The development of a competence scale for learning science: Inquiry and communication. *International Journal of Science and Mathematics Education*, 9(5), 1213–1233.
- Dewey, J. (1938). *Unity of science as a social problem*. Chicago, IL: University of Chicago Press.
- Diekmann, A. (1995). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Reinbek bei Hamburg: Rowohlt Taschenbuchverlag
- Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (Lehrbuch, 2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ginns, P., Marsh, H. W., Behnia, M., Cheng, J. H. S. & Scalas, L. F. (2009). Using postgraduate students' evaluations of research experience to benchmark departments and faculties: issues and challenges. *British Journal of Educational Psychology*, 79(3), 577–598.

- Griffiths, R. (2004). Knowledge production and the research–teaching nexus: The case of the built environment disciplines. *Studies in Higher Education*, 29(6), 709–726.
- Groß Ophoff, J., Schladitz, S., Lohrmann, K. & Wirtz, M. (2014). Evidenzorientierung in bildungswissenschaftlichen Studiengängen. In K. Drossel, R. Strietholt & W. Bos (Hrsg.), *Empirische Bildungsforschung und evidenzbasierte Reformen im Bildungswesen* (S. 251–275). Münster: Waxmann.
- Hancock, G. R. & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck (Hrsg.), *Structural equation modeling. Present and future. A Festschrift in honor of Karl Jöreskog* (S. 195–216). Lincolnwood, IL: Scientific Software International.
- Healey, M. (2005). Linking research and teaching to benefit student learning. *Journal of Geography in Higher Education*, 29(2), 183–201.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Huber, L. (2009). Warum forschendes Lernen nötig und möglich ist. In L. Huber (Hrsg.), *Forschendes Lernen im Studium. Aktuelle Konzepte und Erfahrungen* (Motivierendes Lehren und Lernen in Hochschulen, Bd. 10, S.9–35). Bielefeld: UVW, Webler.
- Klieme, E., Hartig, J. & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 3–22). Göttingen: Hogrefe & Huber Publishers.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (Methodology in the social sciences, 2. Aufl.). New York, NY: Guilford Press.
- Kossek, B. (2009). *Survey: Die forschungsgelietete Lehre in der internationalen Diskussion*. Center for Teaching and Learning (CTL). Universität Wien. Zugriff am 11.07.2018. Verfügbar unter https://ctl-lectures-archiv.univie.ac.at/fileadmin/user_upload/elearning/Forschungsgelietete_Lehre_International_090414.pdf
- Langer, W. (2000). *Einführung in sozialwissenschaftliche Skalen-, Index- und Typenkonstruktion*. Zugriff am 11.07.2018. Verfügbar unter <https://langer.sozioologie.uni-halle.de/pdf/meth1/skalieren2.pdf>

- Marsh, H. W., Rowe, K. J. & Martin, A. (2002). PhD students' evaluations of research supervision: Issues, complexities, and challenges in a nationwide Australian experiment in benchmarking universities. *The Journal of Higher Education*, 73(3), 313–348.
- Mayer, R. E. (2003). What causes individual differences in cognitive performance? In R. J. Sternberg & E. Grigorenko (Hrsg.), *Perspectives on the psychology of abilities, competencies, and expertise* (S. 263–273). New York, NY: Cambridge University Press.
- Meijers, A. W. M., van Overveld, C. W. A. M. & Perrenet, J. C. (2005). *Criteria for academic bachelor's and master's curricula*. Delft: TU Delft.
- Mittelstraß, J. (1991). Einheit und Transdisziplinarität. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 12–22). Berlin: W. de Gruyter.
- Reise, S. P., Moore, T. M. & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92(6), 544–559.
- Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514.
- 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 Online*, 8(2), 23–74.
- Simonton, D. K. (2003). Expertise, competence, and creative ability. In R. J. Sternberg & E. Grigorenko (Hrsg.), *Perspectives on the psychology of abilities, competencies, and expertise* (S. 213–240). New York, NY: Cambridge University Press.
- Smyth, L., Davila, F., Sloan, T., Rykers, E., Backwell, S. & Jones, S. B. (2016). How science really works. The student experience of research-led education. *Higher Education*, 72(2), 191–207.
- Stein, J. A., Lee, J. W. & Jones, P. S. (2006). Assessing cross-cultural differences through use of multiple-group invariance analyses. *Journal of Personality Assessment*, 87(3), 249–258.
- Thiel, F. & Böttcher, F. (2014). Modellierung fächerübergreifender Forschungskompetenzen - Das RMKR-W-Modell als Grundlage der Planung und Evaluation von Formaten forschungsorientierter Lehre. In B. Berendt, A. Fleischmann, J. Wildt, N. Schaper & B. Szczyrba (Hrsg.), *Neues Handbuch Hochschullehre* (I 2.10, S. 109–124). Berlin: Raabe.

- Tremp, P. (2005). Verknüpfung von Lehre und Forschung: Eine universitäre Tradition als didaktische Herausforderung. *Beiträge zur Lehrerbildung*, 23(3), 339–348.
- Valter, K. & Akerlind, G. (2010). Introducing students to ways of thinking and acting like a researcher: A case study of research-led education in the sciences. *International Journal of Teaching and Learning in Higher Education*, 22(1), 89–97.
- Visser-Wijnveen, G. J., van der Rijst, Roeland M. & van Driel, J. H. (2015). A questionnaire to capture students' perceptions of research integration in their courses. *Higher Education*.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Seattle: Hogrefe & Huber.
- Weinert, F. E. (2002). Schulleistungen - Leistungen der Schule oder der Schüler? In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (Beltz Pädagogik, 2. unveränd. Aufl., S. 73–86). Weinheim: Beltz
- Willison, J. & O'Regan, K. (2007). Commonly known, commonly not known, totally unknown: a framework for students becoming researchers. *Higher Education Research & Development*, 26(4), 393–409.
- Yuan, K.-H. & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, 30(1), 165–200.



Kapitel 4

Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen - Ein Instrument zur Evaluation forschungsorientierter Lehr- Lernarrangements

This is a post-peer-review, pre-copyedit version of an article published in *Unterrichtswissenschaft*. The final authenticated version is available online at: <https://doi.org/10.1007/s42010-019-00053-8>

Böttcher-Oschmann, F., Groß-Ophoff, J. & Thiel, F. (2019). Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen - Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements. *Unterrichtswissenschaft*. Online veröffentlicht: 16.04.2019

Kapitel 5

How do we prepare teacher training students for Evidence-based Practice? Promoting students' research competences in Research-learning projects

Anmerkung: Dies ist die Version der Autorin von einer Arbeit, die im *International Journal of Educational Research* eingereicht wurde.

Böttcher-Oschmann, F., Groß-Ophoff, J. & Thiel, F. (*submitted for publication*). How do we prepare teacher training students for Evidence-based Practice? Promoting students' research competences in Research-learning projects.

5 How do we prepare teacher training students for Evidence-based Practice? Promoting students' research competences in Research-learning projects

Abstract

Research-learning projects (RLP) offer the possibility to prepare teacher training students for the requirements of Evidence-based Practice (EBP). Two models of student research competences serve as a basis for a newly developed RLP-format. In a longitudinal study, the acquisition of competences is evaluated with a self-assessment of competences and a competence test. Over- or underestimation of competences and response shift will be considered. A control group was included to increase internal validity of the study. The results indicate that students' preparation for EBP was successful. Further development opportunities for the RLP-format are discussed.

Keywords: Evidence-based Practice * Evidence-based decision-making * Research competences * Research-learning project * Longitudinal study

5.1 Introduction

At university, students are not only prepared for academic careers, but also for evidence-based professions (Brew, 2006) such as teaching. A central learning outcome for these professions is the acquisition of competences in dealing with evidence. In 2004, the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (German: Kultusministerkonferenz) decided that German students should be prepared for Evidence-based Practice (EBP) and therefore incorporated necessary competences in the Standards for Teacher Training in the Educational Sciences (Kultusministerkonferenz, 2004). Accordingly, teacher training students should learn a) about methods of self- and external evaluation to develop and control quality of teaching and school, and b) how to reflect and evaluate results of educational research. This should enable students in future practice c) to reflect on their own professional experiences and competences in order to draw conclusions from this reflection, and d) to use the results of educational research for their own professional practice.

Since EBP is becoming increasingly important in teacher training, university teaching formats need to be geared towards these requirements. Against this background, a Research-learning Project (RLP) was developed and evaluated for teacher training students based on existing competence models.

5.1.1 Evidence-based Practice and evidence-based decision making

Since the late 1990s, the concept of EBP has been widely discussed for the profession of teaching (e.g. Hammersley, 1997, 2007; Hargreaves, 1996, 1997). The concept of EBP has its roots in the medical field and describes a continuous, problem-based learning process in which every available evidence is used for diagnosis, prognosis and therapy. "The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research" (Sackett et al., 1996). In 1996, Hargreaves pointed out the parallels between professional requirements in the field of medicine and the teaching profession. He applied the concept of evidence-based medicine to the teaching profession, which can be labelled as *evidence-based teaching* or *research-based profession* (Hargreaves, 1996, 1997).

In both professions, medical and educational practitioners should be able to

- transform needed information into answerable questions,
- efficiently review the best current evidence to answer these questions,
- critically check the validity and usefulness of the evidence,
- apply the evidence to the problem in practice,
- evaluate the results of this application.

In accordance with Sackett and colleagues, Davies (1999) stated for the educational context that "evidence-based education means integrating individual teaching and learning expertise with the best available external evidence from systematic research" (p. 117). Moon (2000) urged to develop "a stronger knowledge and evidence base to teaching" (2000, p. 6). For this, the process of evidence-based decision-making in educational practice can be transferred from areas such as psychological diagnosis (Palisano, 2006; Schmidt-Atzert & Amelang, 2012). It consists of three phases: diagnosis, intervention and evaluation. The first phase, *diagnosis*, includes the current state of a person's or group's individual and contextual characteristics. The second phase, *intervention*, includes the actions to achieve the intended mechanisms, e.g. changes in cognition, motivation or in behaviour. The third phase, *evaluation*, includes the comparison of the initial state and the achieved effects.

Although the transfer of the EBP-concept from medicine to the teaching profession was viewed critically at the beginning (Hammersley, 1997), it is now widely considered in teacher training programs. However, there is a debate about the extent of EBP: According to Davies (1999) and Borg (2010), EBP can be divided into two levels. Davies (1999) distinguishes *using research* vs. *establishing research*. Borg (2010) describes these levels as *engagement with research* vs. *engagement in research*. The first level *using research* or *engagement with research* describes the reflection and use of evidence to solve problems in teaching practice. It is about *knowing* where and how to find systematic and comprehensible evidence, e.g. educational databases, and how such evidence can be used to answer questions. The second level *establishing research* or *engagement in research* describes the systematic investigation of a research question by *oneself* using research methods¹⁶. This can include, for example, the

¹⁶ In the following, the terms "using research" and "establishing research" by Davies (1999) will be used. However, they stand equally for the term "engagement with research" and "engagement in research" (Borg, 2010).

planning and preparation of a study, data collection and analysis. Davies (1999) is convinced that both levels – *using* as well as *establishing* research – are essential for EBP. Thus, it is necessary to impart both levels in university teaching formats for teacher training students: knowledge about current research and its use for practical decision making as well as basic research competences. Especially existing models of research competences can stimulate the design of teaching formats in higher education, which are introduced in the following.

5.1.2 Competence models as a basis for the development of a teaching format to promote EBP

A model that includes both *using* and *establishing research* is the RMRC-K-model (Böttcher & Thiel, 2018; Thiel & Böttcher, 2014) consisting of five dimensions: *Skills in Reviewing the State of Research*, *Methodological Skills*, *Skills in Reflecting on Research Findings*, *Communication Skills*, and *Content Knowledge*. It is a cross-disciplinary model of student research competences and is based on the pragmatic theory of science (Butts, 1991; Mittelstraß, 1991), which states that the research process shares common characteristics across various disciplines – even if research methodology is different (Böttcher & Thiel, 2018). The model is systemized and operationalized along the research process. Regardless of the discipline, in every research process a) current evidence has to be reviewed and evaluated in order to identify research needs (*Skills in Reviewing the State of Research*), b) research methods are systematically applied, which are typical for the discipline (*Methodological Skills*), c) results are interpreted and findings are reflected with regard to their practical and ethical implications (*Skills in Reflecting on Research Findings*), and d) findings are communicated with the *scientific community* via presentations or publications (*Communication Skills*). All this is based on d) *Content Knowledge* of central theoretical constructs, methods, and disciplinary standards for presenting research findings. The model's facets allow defining precise learning outcomes for teaching formats to promote competences for EBP and evaluate their achievement. Skills such as the systematic review of evidence (*Skills in Reviewing the State of Research*) as well as the reflection of evidence (*Skills in Reflecting on Research Findings*) correspond with the notion of *using research*. Since, however, generating evidence oneself by, for example, planning the research process, applying research methods (*Methodological Skills*) or reflecting on *own* collected evidence (*Skills in Reflecting on Research Findings*), *establishing research* is also an elementary component of the model. As the model is aimed for disciplines that prepare for an academic career in general, not for teaching practice, another model was

taken into account that is based along the research process, too, but differs as it focuses only on *using research*: Educational Research Literacy (ERL; Groß Ophoff, Schladitz, Lohrmann & Wirtz, 2014; Groß Ophoff, Wolf, Schladitz & Wirtz, 2017) was developed especially for the field of Educational Science and describes "the ability to purposefully access, comprehend, and reflect scientific information as well as apply the resulting conclusions to problems with respect to educational decisions" (Groß Ophoff et al., 2017, p.39). The corresponding conceptual model is based on four aspects of the research cycle: (1) *ask* appropriate questions, (2) *search* for relevant information, (3) *evaluate* and (4) *appraise* evidence. Thus, ERL is comprised of three competence facets situated between the single research phases comparable to the dimensions of the first model: *Information Literacy (IL)* – skills in formulating appropriate research questions as well as searching and evaluating needed information between (1) “ask” and (2) “search”, similar to *Skills in Reviewing the State of Research*; *Statistical Literacy (SL)* – skills in dealing with data and their form of representation between (2) “search” and (3) evaluate”, similar to *Methodological Skills*; *Evidence-based Reasoning (ER)* – skills in interpreting evidence to be applied in practice between (3) “evaluate” and (4) “appraise”, similar *Skills in Reflecting on Research Findings*. On the basis of these two competence models, the process of evidence-based decision-making can be further differentiated.

In order to make the *diagnosis* of the current state of individual and contextual characteristics, teachers need *Content Knowledge* of central theories about teaching and learning, current findings and central research methods in the area of the educational sciences. The systematic review of current evidence as well as the evaluation of relevant literature is a fundamental component of this phase (*Information Literacy* and *Skills in Reviewing the State of Research*). In order to assess this current state, teachers need to be able to select an appropriate survey design as well as to evaluate the corresponding data (*Statistical Literacy* and *Methodological Skills*). In the *intervention phase*, teachers make and implement decisions based on both new information acquired through diagnosis and existing professional knowledge (Shulman, 1986, 1987). In the *evaluation phase*, the achieved effects are finally assessed. For this purpose, it is necessary to analyse and interpret the results and to evaluate them regarding practical implications (*Statistical Literacy* and *Methodological Skills* as well as *Evidence-based Reasoning* and *Skills in Reflecting on Research Findings*). How a learning environment needs to be designed to promote the necessary competences is described below.

5.1.3 Research-learning projects to promote EBP

Research-learning projects (RLP) can offer university students the opportunity to acquire research competences. RLP generally are connected to existing research projects and usually last one to two semesters. As the introduction of a six-month practical semester in German teacher training offered a good opportunity to promote skills for EBP in teacher training, we adapted the RLP-format for teacher training students. The practical semester takes place in the third semester of the teacher training master's degree program. Most of the time, students are at school, do classroom observations, and improving their instructional skills – accompanied by their mentors. In addition, the students attend the RLP at university. The ungraded seminar is limited to 15 students, who work independently on a research question in educational science in groups of 2 to 4 persons and are supported by their lecturers. In the design of the RLP-format, there is a focus on the phases of *diagnosis* and *evaluation* for organisational reasons and to meet the challenge to reduce the complexity of the research process to one semester. Hence, the topic of the research questions is limited to the *diagnosis of non-cognitive characteristics* and the *evaluation of teaching quality*.

The RLP is divided into three phases: preparation for the research process, fieldwork, data presentation. During the semester, students are supported by their lecturers. The first phase *preparation for the research process* consists of seven sessions. Students acquire knowledge about the individual phases of the research process. Therefore, the central aspects of the research process are presented using concrete examples from educational research. Students learn about strategies for literature research, the structure of scientific articles, central educational research and evaluation methods, and how research results are presented on a poster. Thereby, students repeat and use their knowledge about current theories and findings on instructional quality and relevant non-cognitive characteristics. During this first phase, students define their research question (e.g. "How do girls and boys in my class differ in their perception of the motivational quality of biology classes?"; "How do my students and my mentor rate my classroom management compared to my own evaluation?") and develop an adequate research design. Therefore, students can choose from a range of pre-selected constructs and instruments that are suitable for diagnosing non-cognitive characteristics or evaluating quality of teaching. At the end of this phase, students prepare an exposé for submission and approval by their schools, for which they receive detailed feedback from their lecturer. The first phase of the RLP is therefore concerned with imparting *Content Knowledge*, practicing strategies for literature

research (*Skills in Reviewing the State of Research, Information Literacy*), selecting appropriate research methods, and developing their study design (*Methodological Skills*).

The second phase *fieldwork* lasts five weeks during which students are at school most of the time. They return to university solely for two 30- to 45-minute consultations with their lecturer. These consultations serve to support students in their research, data analysis (*Methodological Skills, Statistical Literacy*) and interpretation of evidence (*Skills in Reflecting on Research Findings, Evidence-based Reasoning*).

In the third and last phase *data presentation*, students prepare an academic poster about their research project. At the end of the RLP, posters are presented at a final event with conference character (*Communication Skills*).

The challenge in conducting the RLP is to reduce the complexity of the research process in such a way that it can be completed within one semester. To achieve this, students receive various support services, e.g. a manual for doing literature research or templates for the exposé, data analysis, and the poster. In addition, the learning environment was designed in such a way that students could go through the research process in the best possible way.

A study by Hollingsworth and Fassinger (2002) has shown that the lecturer's support during the research process is correlated with students' research productivity. That is why students in RLP are supported by lecturers in the sense of the *Cognitive Apprenticeship-approach* (Collins, Brown, & Newman, 1989). As the acquisition of competences in RLP is highly self-regulated, the lecturers should provide specific support. Therefore, the method of *modelling* is important in the first part of the RLP. By modelling, (meta-)cognitive processes can be made visible by the lecturer. In RLP, the phases of the research process are presented on the basis of concrete examples, e.g. introducing research strategies in databases, which are specific for the educational science, study designs and evaluation methods, which are especially suitable for the teaching context. The methods of *coaching* and *scaffolding* are important in the second and third phase of the RLP. Students are supported in their own research process by receiving concrete hints when they reach their limits, e.g. they receive advice on how to formulate and work on their research questions as well as hints if they do not know how to continue with the analysis of their collected data. Moreover, this RLP-format is in line with the theory of self-determined learning (Deci & Ryan, 2000). Students are provided with the possibility of a) experiencing autonomy by being able to freely choose a question from the

defined subject area, b) experiencing competence by going through the individual phases of the research process themselves and c) completing the RLP in a social community.

The extent to which the RLP-format to promote EBP is successful, can be evaluated based on the development in the aspired competences. For this purpose, different assessment procedures are available, some of which are outlined in the following section.

5.1.4 Challenges in measurement of competences

Competences can be assessed using so-called subjective and/or objective measures (Lucas & Baird, 2006). Objective measures such as competence tests allow to operationalize direct increases in competences (Lucas & Baird, 2006; Mabe & West, 1982). In subjective measures, participants rate the competences themselves, so indirect increases in competence can be measured. Both survey methods have strengths and weaknesses, which is why the combination of these measures is recommended (Lucas & Baird, 2006). For example, competence tests have a low test efficiency: a large pool of test items or several parallel tests are required to exclude recall bias in longitudinal surveys. Additionally, for reasons of time efficiency, only selected areas of knowledge can be tested (Cramer, 2010; Mertens & Gräsel, 2018). On the other hand, self-assessments of competences are often subject to different bias (e.g. social desirability), to misjudgement or to response shift (Kruger & Dunning, 1999; Lucas & Baird, 2006; Schwarz, 1999; Sprangers & Schwartz, 1999). On the one hand, the level of experience plays an important role in self-assessment: with increasing experience, people can rate their competences in a more differentiated way (Bach, 2013; Mertens & Gräsel, 2018). On the other hand, the level of competence affects self-assessment, too (Kruger & Dunning, 1999): Incompetent people *overestimate* their competences more than competent people, while competent people *underestimate* their abilities compared to their peers. Hence, in contrast to an objective competence test, the over- or underestimation of competences can lead to misjudgements (Chevalier, Gibbons, Thorpe, Snell, & Hoskins, 2009). Furthermore, in longitudinal studies, response shift may occur (Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999). It is defined as "a change in the meaning of one's self-evaluation of a target construct" (Sprangers & Schwartz, 1999, p. 1532) and is particularly well known in medical research but also in educational intervention research (Piwowar & Thiel, 2014; Meyer, Richter, & Raspe, 2013; Schwartz & Sprangers, 2010). Three types of response shift can be distinguished: a) *recalibration*, the change in a person's internal standards of measurement, b) *repriorisation*, the change in a person's values, c) *reconceptualisation*, a redefinition of the

target construct (Schwartz & Sprangers, 1999; Sprangers & Schwartz, 1999). There are design-based or statistical approaches to deal with response shift (Schwartz & Sprangers, 1999). The so-called "thentest" or "retrospective pretest" is a design-based approach to assess changes in internal standards (*recalibration*). During posttest, participants fill out the self-report questionnaire regarding how they evaluate themselves before an intervention. This allows to assess a retrospective estimation of one's own competence before a treatment. The difference in post values and values of retrospective pretest is interpreted as an increase in competences without the influence of the response shift (an "adjusted" or "quasi-indirect" measurement of change; Meyer et al., 2013; Schwartz & Sprangers, 2010). Hence, it can be assumed that both values share the same internal standards. The difference between the regular pretest and the retrospective pretest indicates, if a response shift exists and whether there is an initial over- or underestimation (Schwartz & Sprangers, 1999, 2010). In addition, the retrospective pretest is suitable for encouraging self-reflection on one's own increase in competence, which in turn can have a positive effect on self-efficacy beliefs (Hill & Betz, 2005). However, bias such as social desirability, recall bias and implicit theory of change can also influence differences in mean values when using a retrospective pretest (Hill & Betz, 2005; Schwartz & Sprangers, 2010). While the retrospective pretest cannot be used to identify repriorisation or reconceptualisation, they can be analysed with statistical approaches, e.g. testing measurement invariance (Piwowar & Thiel, 2014; Meredith, 1993). Thus, not only the occurrence of a response shift but also the *concrete* type of change can be analysed. Depending on the level of measurement invariance, recalibration, reprioritisation or reconceptualisation is indicated (Oort, 2005). For this statistical approach, however, there must be enough participants in a sample ($N \geq 100$) to reliably apply this method (Oort, 2005; Schwartz & Sprangers, 1999).

5.1.5 The current study

The aim of the presented study is to evaluate the above mentioned teaching format, which prepares teacher training students for the requirements of EBP. Therefore, a RLP-format based on the RMRC-K-model (Böttcher & Thiel, 2018; Thiel & Böttcher, 2014) and the ERL-model (Groß Ophoff et al., 2014, 2017) was developed. To evaluate the acquisition of competences, a combination of competence test and competence self-assessment was used, whereby misjudgements and response shift are considered.

5.2 Methodology and methods

5.2.1 Study design and sampling

In a longitudinal design, data were collected at the beginning and at the end of the winter semester 2016/17. Students participated in the RLP-format as described above. It should be noted that the sample of teacher training students in RLP underlies a certain pre-selection: students had the opportunity to register for RLP themselves. In case of overbooking, a maximum of 15 students were randomly distributed among the RLP on the basis preferences. Hence, it is a non-randomised sample.

In order to achieve a higher internal validity, a control group (CG) was included in the design. However, since all students of a cohort participated in the RLP, a group was sought that had characteristics as similar as possible compared to the group of students in the RLP. One lecturer offered three seminars with a maximum of each 30 students as a control group. All these seminars were based on the same seminar concept, dealing with teaching and learning. The focus of this concept was on the application of research-based knowledge for practice. The concept thus corresponds to the level of *using research* (Davies, 1999). Students in the CG were one academic year before students in the practical semester and completed the RLP one year later. Hence, they can be described as a waiting control group. Thus, the study design can be classified as quasi-experimental.

To evaluate an increase in competences, a combination of self-assessment of competences and competence testing was chosen. The self-assessment was carried out before ($_{self1}$) and after ($_{self2}$) the RLP in order to be able to discover an indirect increase in competence. At the end of the RLP, self-assessment was additionally used in a retrospective pretest ($_{self1}^{retro}$), whereby students assessed their competences *before* the seminar. This quasi-indirect measurement offers the possibility that students reflect their own increase in competence. Furthermore, the occurrence of a *response shift* can be considered in the analyses (Hill & Betz, 2005). The competence test was also used before ($_{test1}$) and after ($_{test2}$) the RLP in order to assess the direct increase in competence.

At the first assessment time point, 158 students participated in the survey in both seminar concepts ($n_{RLP} = 97$; $n_{CG} = 61$), 125 students at the second assessment time point ($n_{RLP} = 78$; $n_{CG} = 47$). In RLP, these students were spread over nine seminars with five lecturers, in CG over three seminars with one lecturer.

For longitudinal analyses, those students were selected who took part in the survey at both assessment time points. Hence, $n_{RLP} = 36$ students remained in the RLP-group, $n_{CG} = 23$ students in the CG-group (see Table 5.1). During the practical semester, RLP-students were predominantly in the third master semester ($M_{RLP} = 2.94$, $SD_{RLP} = 0.33$), CG-students in the first master semester ($M_{CG} = 1.09$, $SD_{CG} = 0.42$). 18 CG students (85.7%) attended a lecture on "Pedagogical Diagnostics" in the same semester. The two groups do not differ significantly regarding gender and second teaching subject, but regarding their first teaching subject (in RLP more languages as first teaching subject than in CG).

Table 5.1 Characteristics of participants in the longitudinal sample

	RLP	CG
	N = 59	
n	36	23
Gender		
Male	13	10
Female	22	13
Not reported	1	-
First / second teaching subject		
Languages	12 / 5	5 / 6
Social sciences	5 / 7	3 / 4
Natural sciences	14 / 14	11 / 10
Musical-creative	-	1
Primary education	-	1
Others	-	1
Not reported	5	1
Semester		
Mean / Standard deviation	2.94 0.33	1.09 0.42

5.2.2 Data collection and measures

In a paper pencil survey at the beginning and at the end of the semester, students completed a (1) self-assessment of competences followed by a (2) competence test. Prior to both surveys, students were informed about the study's aims, the voluntary nature of the survey, and their anonymity was guaranteed. At the end of each survey, students indicated some personal data (gender, first and second teaching subject¹⁷ and number of semesters). Students

¹⁷ During the survey, students were provided with the possibility to omit the indication of teaching subjects. Some students were worried that they could be identified by their rare combination of teaching subjects.

in CG were also asked to indicate whether they attended the lecture “Pedagogical Diagnostics” during the semester.

(1) *Self-assessment of competences.* The instrument to assess student research competences (*R-Comp*; Böttcher & Thiel, 2016; 2018) via self-ratings is based on the RMRC-K-model (Thiel & Böttcher, 2014). The *R-Comp* consists of 32 items (with a 5-point response scale ranging from “1 – strongly disagree” to “5 – strongly agree”) in five scales corresponding to the theoretical dimensions of the competence model: *Skills in Reviewing the State of Research* (four items; $\alpha = .87$; e.g. "I am able to systematically review the state of research regarding a specific topic."), *Methodological Skills* (eight items; $\alpha = .88$; e.g. "I am able to decide, which data/sources/materials I need to address my research question."), *Skills in Reflecting on Research Findings* (six items; $\alpha = .92$; e.g. "I am able to critically reflect on methodological limitations of my own research findings."), *Communication Skills* (five items; $\alpha = .89$; e.g. "I am able to write a publication in accordance with the standards of my discipline."), and *Content Knowledge* (nine items; $\alpha = .88$; e.g. "I am informed about the main (current) theories in my discipline."). The *R-Comp* mainly assesses *establishing research* (Davies, 1999) in all phases of the research process. In the *R-Comp*'s instructions, students were requested to answer questions with regard to the educational part of their studies. This should ensure that the self-assessment of research competences is carried out specifically for the domain of educational sciences. Two versions of the *R-Comp* were used: at the beginning and at the end of the semester in its original version; at the end of the semester as a retrospective pretest, to assess how students evaluate their skills and knowledge “before the RLP” (Schwartz & Sprangers, 1999, 2010).

(2) *Competence test.* In order to assess research competences on the basis of external criteria, the competence test corresponding to the ERL-model (Groß Ophoff et al., 2014, 2017) was used in pre- and posttest. This test assesses the level of *using research* (Groß Ophoff et al., 2017; Davies, 1999). Two test booklets were used consisting of 18 items at the pretest (*Information Literacy: 7 items; Statistical Literacy: 7 items; Evidence-based Reasoning: 4 items*) and 17 items at posttest (*Information Literacy: 7 items; Statistical Literacy: 5 items; Evidence-based Reasoning: 5 items*). These included items on the use of research strategies for *Information Literacy* and statistical/numerical tasks for *Statistical Literacy*, both mainly in multiple-choice formats. Items in the field of *Evidence-based Reasoning* included, for example,

two abstracts that had to be evaluated regarding the admissibility of several statements¹⁸. Test items were selected from a large item pool of 193 items that was standardized in a large-scale study ($N_i = 1360$, cf. Groß Ophoff et al., 2014). In item selection, care has been taken to ensure a broad spectrum of competences and sufficient discriminatory power of selected items ($M(r_{it}) = .31$).

5.2.3 Analyses

Statistical Characteristics

R-Comp: For each version of the R-Comp (pre - t_1 , post - t_2 , retrospective pretest - t_1^{retro}) five scale scores were calculated according to the five competence dimensions of the RMRC-K-Model. Thereby, the hierarchical structure of the empirical model (Böttcher & Thiel, 2018) was taken into account. Internal consistency for all R-Comp scales was evaluated inspecting Cronbach's α .

ERL-competence test: Scale scores computed on the three-dimensional ERL-model (*Information Literacy, Statistical Literacy, Evidence-based Reasoning*), which is viable for course evaluation (cf. Groß Ophoff et al., 2017). Person measures for each subdimension were determined on the basis of a dichotomous response format of the items ("1 = correctly solved" and "0 = not correctly solved") using the WLE estimator (Warm, 1989). Thereby, the person measures from manifest data are estimated using a maximum likelihood function (Hartig & Kühnbach, 2006; Strobl, 2010). Item difficulties were fixed (external anchor design; see Mittelhaeuser, Béguin, & Sijtsma, 2011; Wright & Douglas, 1996) in order to equate results of the current study to the standardisation study (Groß Ophoff et al., 2014).

Difference values and effect sizes

In order to compare the difference (Δ) between the assessment time points, the following values were calculated. For self-assessment: (a) $\Delta_{\text{self}t_2-t_1}$ to determine indirect differences in competences; (b) $\Delta_{\text{self}t_2-t_1^{\text{retro}}}$ to determine the quasi-indirect differences in competences via retrospective assessment without response shift; (c) $\Delta_{\text{self}t_1-t_1^{\text{retro}}}$ to identify response shift and misjudgements, where positive difference scores indicate over-estimation, negative difference scores under-estimation (Schwartz & Sprangers, 2010). For the competence test, (d) $\Delta_{\text{test}t_2-t_1}$ was calculated to determine the direct differences in competences.

¹⁸ A more detailed description of the items can be found in Groß Ophoff et al. (2017).

The calculation of effect sizes for the differences was carried out according to Lakens' recommendations inspecting d_{AV} (2013). The effects were interpreted in line with Cohen's benchmarks (1992): $\geq .20$ "small", $\geq .50$ "medium", and $\geq .80$ "large".

Longitudinal Analyses

Multivariate and multifactorial variance analyses (MANOVA) with repeated measurement were conducted in SPSS 25 to evaluate group specific effects. Because of the small sample size in both groups, the hierarchical structure of the data was not considered. Hence, the comparison is based on differences between the seminar *concepts*. MANOVA with repeated measurement was performed according to the General Linear Model (GLM) with the factors time and group. Variance homogeneity was given (Levene test). If the Mauchly test was not given, the Huynh field corrected degrees of freedom were used (see Table 5.3; Field, 2009). Afterwards, univariate analyses were performed to identify major effects. Additionally, individual comparisons between factor levels were determined considering the Bonferroni correction to account for the problem of multiple comparisons. The overall significance level was set at $p = .05$, while effect sizes η^2_p were interpreted according to Cohens (1988) benchmarks with $\geq .01$ "small", $\geq .06$ "medium", and $\geq .14$ "large".

For the analyses of self-assessment, a 2x2x3 MANOVA included the assessment time points t_1 , t_1^{retro} , t_2 for the five R-Comp scales in both groups. Even though t_1^{retro} is a theoretical assessment time point, it was included in MANOVA to investigate the occurrence of response shift. For the analyses of the competence test, a 2x2x2 MANOVA included test values for the three ERL scales at t_1 and t_2 in both groups.

Missing Values

None response occurs as participants either do not take part at one of the assessment time points or do not answer individual items (Sax, Gilmartin, & Bryant, 2003). Only 37% of the students, who took part in the first survey, remained in the sample at the second assessment time point. There was no systematic dropout according to students' personal data or competences at first assessment time point¹⁹. Unanswered items in self-assessment of competences did not occur more than two times per variable at any (theoretical) assessment time point and were not imputed or replaced. For the analyses of the competence test, omissions

¹⁹ However, students of our study were the target group of two surveys, which took place at the same time.

were not imputed or replaced but treated as missing values (Groß Ophoff et al., 2017). With regard to personal information, altogether, 9.7% missing values occurred, most of them for first (10.2%) and second (22.0%) teaching subject²⁰.

5.3 Results

MANOVA

For *self-assessed research competences*, significant and large multivariate effects were found for time and the interaction of time and group on the within level (Table 5.2). No significant, but a small effect was found for the group on the between level. Moreover, results of univariate analysis indicated that the effect of time was caused by all five skills and knowledge dimensions with medium to small effect sizes (Table 5.3). The interaction of time and group was caused by *Communication Skills*, and moreover tended to go back to *Skills in Reviewing the State of Research* as well as *Methodological Skills*, which were marginally significant. All dimension showed small effect sizes.

Table 5.2 Multivariate effects for self-assessed and tested research competences

	Pillai's Trace	<i>F</i>	<i>df</i>	Error <i>df</i>	<i>p</i>	η^2_p
Self-assessed Research Competences						
<i>Between</i>						
Intercept	0.97	343.02	5	53	0.00	0.97
Group	0.08	0.98	5	53	0.44	0.08
<i>Within</i>						
Time	0.67	9.81	10	48	0.00	0.67
Time x group	0.44	3.77	10	48	0.00	0.44
Tested Research Competences						
<i>Between</i>						
Intercept	0.37	10.89	3	55	0.00	0.37
Group	0.15	3.15	3	55	0.03	0.15
<i>Within</i>						
Time	0.20	4.55	3	55	0.01	1.99
Time x group	0.02	0.30	3	55	0.82	0.02

Notes: *F* = F-statistic; *df* = degrees of freedom; *p* = level of significance; η^2_p = effect size

For *tested research competences*, the results showed that they were significantly affected by group with a large effect on the between level, and time with a large effect on the

²⁰ Probably, these missing values occurred, since students were provided with the opportunity to omit the corresponding items (see above).

within level, not significantly but with a small effect affected by the interaction of time and group on the within level (Table 5.2). Moreover, results of univariate analysis indicated that the effect of time was caused by *Information Literacy* and *Statistical Literacy* with both medium effects, and *Statistical Literacy* had a small effect on the interaction of time and group (Table 5.3).

Table 5.3 Univariate effects for self-assessed and tested research competences

	<i>df</i>	SS	MS	<i>F</i>	<i>p</i>	η^2_p
Self-assessed Research Competences						
<i>Time</i>						
Skills in Reviewing the State of Research	1.75 ^a	5.44	3.11	11.97	0.00	0.17
Methodological Skills	2	3.58	1.79	11.97	0.00	0.17
Skills in Reflecting on Research Findings	1.74 ^a	5.20	2.99	13.41	0.00	0.19
Communication Skills	1.60 ^a	2.38	1.48	6.57	0.00	0.10
Content Knowledge	1.71 ^a	15.91	9.30	20.10	0.00	0.26
<i>Time x group</i>						
Skills in Reviewing the State of Research	1.75 ^a	1.18	0.68	2.60	0.09	0.04
Methodological Skills	2	0.82	0.41	2.75	0.07	0.05
Skills in Reflecting on Research Findings	1.74 ^a	0.70	0.40	1.81	0.17	0.03
Communication Skills	1.60 ^a	1.18	0.74	3.26	0.05	0.05
Content Knowledge	1.71 ^a	1.52	0.89	1.92	0.16	0.03
Tested Research Competences						
<i>Time</i>						
Information Literacy	1	5.77	5.77	6.57	0.01	0.10
Statistical Literacy	1	5.30	5.30	7.29	0.01	0.11
Evidence-based Reasoning	1	0.00	0.00	0.00	0.99	0.00
<i>Time x group</i>						
Information Literacy	1	0.01	0.01	0.01	0.93	0.00
Statistical Literacy	1	0.66	0.66	0.91	0.34	0.02
Evidence-based Reasoning	1	0.00	0.00	0.01	0.94	0.00

Notes. *df* = degrees of freedom; SS = Sum of Squares; MS = Mean Square; *F* = F-statistic; *p* = level of significance; η^2_p = effect size; ^a Huyhn-Feldt corrected *df*

Differences in mean values

All mean values, standard deviations, differences in mean values and effect sizes can be found in Table 5.4 for students in RLP, and in 5.5 for students in CG.

Indirect differences in competences ($\Delta_{\text{self}t_2-t_1}$): For students in RLP, significant increases in *Skills in Reviewing the State of Research*, *Methodological Skills*, and *Communication Skills* occurred. Thereby, *Skills in Reviewing the State of Research* showed a large effect, *Communication Skills* a medium effect, *Methodological Skills* and *Skills in Reflecting on Research Findings* a small, positive effect, and *Content Knowledge* no effect. For students in CG, no significant difference occurred. Whereby, *Skills in Reviewing the State of Research* and *Content Knowledge* showed a small positive effect, *Skills in Reflecting on Research Findings* a small negative effect, and *Methodological Skills* and *Communication Skills* no effect. These results are consistent with results of univariate analyses, which indicated that the group-specific effect was caused by *Skills in Reviewing the State of Research*, *Methodological Skills*, and *Communication Skills*.

Quasi-indirect differences in competences ($\Delta_{\text{self}t_2-t_1}^{\text{retro}}$): For students in RLP, significant increases occurred in all skills and knowledge scales. Thereby, *Skills in Reviewing the State of Research* showed a large effect, *Methodological Skills*, *Communication Skills*, and *Content Knowledge* a medium effect, *Skills in Reflecting on Research Findings* a small effect. For students in CG, significant increases occurred for *Skills in Reflecting on Research Findings* and *Content Knowledge*. Thereby, *Content Knowledge* showed a large effect, *Skills in Reviewing the State of Research*, *Skills in Reflecting on Research Findings*, and *Methodological Skills* a small effect, *Communication Skills* no effect.

Over- and underestimation of competences ($\Delta_{\text{self}t_1-t_1}^{\text{retro}}$): For students in RLP, positive but not significant differences occurred. Thereby, *Content Knowledge* showed a medium effect, all skills scales a small effect. For students in CG, significant increases occurred for *Methodological Skills*, *Skills in Reflecting on Research Findings*, and *Content Knowledge*. Thereby, *Content Knowledge* showed a large effect, *Skills in Reflecting on Research Findings* a medium effect, *Methodological Skills* a small effect, *Skills in Reviewing the State of Research* and *Communication Skills* no effects. In both groups, all identified effects indicate an initial over-estimation.

Direct differences in competences ($\Delta_{\text{test}t_2.t_1}$): For students in RLP, significant increases occurred for *Information Literacy* and *Statistical Literacy*, both with large, positive effects. Whereas, *Evidence-based Reasoning* showed no effect. For students in CG, no significant differences occurred. Nevertheless, for *Information Literacy* and *Statistical Literacy* showed large, positive effects. *Evidence-based Reasoning* showed no effect. The effect of the difference for *Information Literacy* tends to be higher in CG, whereas the effect for *Statistical Literacy* tends to be higher in the group of RLP. These results are consistent with results of univariate analyses, which indicated that a time effect was caused by *Information Literacy* and *Statistical Literacy* and that no group-specific effect occurred.

5.4 Discussion

The aim of this study was to develop, implement and evaluate a Research-learning project (RLP)-format to promote Evidence-based Practice (EBP). Teacher training students should acquire competences in diagnosis and evaluation for evidence-based decision-making. With the RMRC-K-model (Thiel & Böttcher, 2014) and the ERL-model (Groß Ophoff et al., 2014), two models of student research competences were chosen as foundation of this teaching format because both models provided differentiated competence dimensions as well as corresponding instruments to evaluate the acquisition of competences. Anyway, they differ in their focus on levels of teachers' Evidence-based Practice (Davies, 1999): The RMRC-K-model covers both *using* and *establishing research*, whereas the ERL model focuses primarily on *using research*. In a longitudinal study, the increase in competences was evaluated with a competence test (direct measurement) and a self-assessment of competences (indirect measurement). Furthermore, a retrospective pretest was added to consider over- or underestimation in self-assessments and response shift (quasi-indirect measurement).

The indirect measurement showed that students in RLP evaluated their skill dimensions on a higher level than at the beginning of the seminar, especially *Skills in Reviewing the State of Research*. The quasi-indirect measurement indicated that students perceived an increase in all skill dimensions as well as in *Content Knowledge*. In all five dimensions, students initially overestimated their competences before the RLP. The direct measurement showed that students improved in *Information Literacy* and *Statistical Literacy*. The results of the competence test correspond to the results of the self-assessment of competences. Students showed an increase in similar dimensions: *Information Literacy* and *Skills in Reviewing the State of Research* as well as *Statistical Literacy* and *Methodological Skills*. In general, it can be noted that students

KAPITEL 5 HOW DO WE PREPARE TEACHER TRAINING STUDENTS FOR EVIDENCE-BASED PRACTICE? PROMOTING STUDENTS' RESEARCH COMPETENCES IN RESEARCH-LEARNING PROJECTS

Table 5.4 Mean values, standard deviations, differences in mean values, effect sizes, and Cronbach's alpha for students in RLP

Research Learning Projects (RLP, $n = 36$)													
	self t_1		self t_1^{retro}		self t_2		self $t_1 - t_1^{\text{retro}}$		self $t_2 - t_1$		self $t_2 - t_1^{\text{retro}}$		$\alpha_{t_1/t_1^{\text{retro}}/t_2}$
	M	SD	M	SD	M	SD	Δ	d_{av}	Δ	d_{av}	Δ	d_{av}	
Skills in Reviewing the State of Research	3.02	0.62	2.76	0.89	3.40	0.65	0.26	0.34	0.38**	0.80	0.64***	0.83	.64 / .83 / .72
Methodological Skills	2.93	0.67	2.68	0.82	3.17	0.66	0.25	0.34	0.24*	0.48	0.49***	0.67	.87 / .92 / .88
Skills in Reflecting on Research Findings	3.16	0.75	2.89	0.98	3.27	0.91	0.27	0.31	0.11	0.26	0.38***	0.40	.85 / .93 / .91
Communication Skills	2.80	0.83	2.63	0.91	3.12	0.87	0.17	0.20	0.32**	0.67	0.49***	0.55	.80 / .87 / .83
Content Knowledge	2.61	0.73	2.20	0.83	2.68	0.83	0.41	0.52	0.07	0.17	0.48***	0.58	.82 / .85 / .91
	test t_1		test t_2		test $t_2 - t_1$								
	M	SD	M	SD	Δ	d_{av}							
Information Literacy	0.35	1.19	0.82	1.01	0.47*	0.79							
Statistical Literacy	0.02	1.12	0.61	1.11	0.59**	1.17							
Evidence-based Reasoning	0.40	0.82	0.39	0.77	-0.01	-0.02							

Notes: self t_1 = values of self-assessment at pretest; self t_1^{retro} = values of self-assessment at retrospective pretest; self t_2 = values of self-assessment at posttest; self $t_1 - t_1^{\text{retro}}$ = differences between self-assessed pretest and retrospective pretest; self $t_2 - t_1$ = differences between self-assessed posttest and pretest; self $t_2 - t_1^{\text{retro}}$ = differences between self-assessed posttest and retrospective pretest; α = Cronbach's alpha; M = means; SD = standard deviations; Δ = difference values; d_{AV} = effect size for differences by Lakens (2013); test t_1 = values of tested competences at pretest; test t_2 = values of tested competences at posttest; test $t_2 - t_1$ = differences between tested posttest and pretest; * $p < .05$, ** $p < .01$, *** $p < .001$

Table 5.5 Mean values, standard deviations, differences in mean values, effect sizes, and Cronbach's alpha for students in CG

Control Group (CG, $n = 23$)													
	self t_1		self t_1^{retro}		self t_2		self $t_1 - t_1^{\text{retro}}$		self $t_2 - t_1$		self $t_2 - t_1^{\text{retro}}$		$\alpha_{t_1/t_1^{\text{retro}}/t_2}$
	M	SD	M	SD	M	SD	Δ	d_{av}	Δ	d_{av}	Δ	d_{av}	
Skills in Reviewing the State of Research	2.84	0.81	2.79	0.77	3.02	0.84	0.04	0.05	0.18	0.38	0.23	0.28	.82 / .82 / .82
Methodological Skills	2.83	0.51	2.58	0.65	2.78	0.80	0.25*	0.43	-0.05	-0.17	0.20	0.27	.74 / .78 / .86
Skills in Reflecting on Research Findings	3.30	0.63	2.78	0.75	3.11	0.82	0.51**	0.75	-0.19	-0.49	0.33**	0.41	.82 / .90 / .91
Communication Skills	2.67	0.72	2.65	0.88	2.73	0.85	0.02	0.03	0.06	0.15	0.08	0.10	.66 / .81 / .79
Content Knowledge	2.82	0.83	2.09	0.88	3.02	0.94	0.73**	0.85	0.21	0.47	0.94***	1.03	.82 / .89 / .89
	test t_1		test t_2		test $t_2 - t_1$								
	M	SD	M	SD	Δ	d_{av}							
Information Literacy	0.26	0.91	0.69	1.17	0.44	1.13							
Statistical Literacy	0.67	0.75	0.95	1.13	0.28	0.84							
Evidence-based Reasoning	0.32	0.65	0.32	0.81	0.01	0.02							

Notes: self t_1 = values of self-assessment at pretest; self t_1^{retro} = values of self-assessment at retrospective pretest; self t_2 = values of self-assessment at posttest; self $t_1 - t_1^{\text{retro}}$ = differences between self-assessed pretest and retrospective pretest; self $t_2 - t_1$ = differences between self-assessed posttest and pretest; self $t_2 - t_1^{\text{retro}}$ = differences between self-assessed posttest and retrospective pretest; α = Cronbach's alpha; M = means; SD = standard deviations; Δ = difference values; d_{AV} = effect size for differences by Lakens (2013); test t_1 = values of tested competences at pretest; test t_2 = values of tested competences at posttest; test $t_2 - t_1$ = differences between tested posttest and pretest; * $p < .05$, ** $p < .01$, *** $p < .001$

seem to benefit from the RLP-format in almost every competence aspect. However, to what extent this is due to the RLP-format and not to general development processes, can only be inserted by inspecting results of the control group (CG) who visited a seminar based on research-based, practical knowledge about teaching and learning.

The results of MANOVA indicated a group-specific effect in the area of self-assessed research competences, but not in the area of tested competences. In self-assessment, students in RLP tended to report a higher increase in indirect and quasi-indirect measurement of competences than the students in CG. In direct measurement via competence testing, differences between groups seem less pronounced. Although *Information Literacy* and *Statistical Literacy* increased over time in both groups, this change cannot be explicitly attributed to a seminar concept. The increase in *Statistical Literacy* tended to be higher for students in RLP, whereas *Information Literacy* increased more for students in CG. There are no effects for *Evidence-based Reasoning* in either group. Two points are worth mentioning here:

The CG students started on a higher level of *Statistical Literacy* at the beginning of the semester and continued to improve over the time. However, high variances for *Statistical Literacy* as well as for *Information Literacy* at the end of the semester indicated that this improvement does not apply equally to all students in CG. Furthermore, in RLP, high variances occurred for *Information Literacy* and *Statistical Literacy* in RLP at both assessment time points. These results may be due to influences of confounding factors, such as teaching subjects of teacher training students: On the one hand, their teaching subjects could have an influence on the ability to solve test items. For example, students from mathematics and natural sciences, who often deal with numerical forms of representation, might be more able to solve statistical / numerical problems in the field of *Statistical Literacy*. This would be supported by the high proportion of students from mathematical and scientific teaching subjects in CG. In RLP, the proportion of students from mathematics, natural sciences, philology and social sciences was equally distributed, which could explain high variances for *Information Literacy* and *Statistical Literacy* at both assessment time points. In future, the teaching subjects' influence should be investigated further.

In both groups a response shift occurred, but to varying degrees. On the one hand, in indirect measurement some dimensions did not change over time in both groups, on the other hand, a decrease in *Skills in Reflecting on Research Findings* occurred (Kruger & Dunning,

1999; Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999). Therefore, it is useful to take the results of quasi-indirect measurement into consideration, which showed an increase in competence dimensions without response shift in line with expectations: for students in RLP an increase in all competence dimensions occurred. For students in CG, *Skills in Reflecting on Research Findings* and *Content Knowledge* increased, which are corresponding with the seminar's aims to acquire research-based knowledge for practice. Furthermore, results of quasi-indirect measurement indicate the extent to which the students misjudged themselves in the first assessment time point: In RLP, an overestimation occurred in all five dimensions, in CG only for *Methodological Skills*, *skills in reflecting in research findings*, and *Content Knowledge*. For *skills in reflecting in research findings*, this overestimation lead to the decrease in indirect measured competences over time. In both groups, the reason for the initial overestimation may be the students' low level of experience (Mertens & Gräsel, 2018) at the beginning of the semester. Students might have had a vague idea of *research competences* and therefore misjudged themselves, although some of them achieved good test scores. The fact that students in CG, in contrast to students in RLP, misjudged only three of five dimensions might be explained by the different seminar concepts. While all five dimensions were addressed in the RLP-format, the focus of the CG's seminar was mainly on *Content Knowledge* and *skills in reflecting in research findings*. On these dimensions, students in both groups might be more able to evaluate themselves in a differentiated way at the end of the semester. It seems that a change in person's internal standards of measurement had occurred (recalibration response shift) in these dimensions. However, a reconceptualisation response shift would also be conceivable, since both seminars aim exactly at the acquisition of knowledge in the aforementioned dimensions. Nevertheless, this cannot be answered based on available data. Therefore, measurement invariance testing as a statistical approach would be necessary (Piwoski & Thiel, 2014; Meredith, 1993). Furthermore, the question arises, why then, in CG, there is an effect for *Methodological Skills*, and why not for *Skills in Reviewing the State of Research* and *Communication Skills*. As the seminar of the CG did not focus on research methods, the effect for *Methodological Skills* might have occurred due to attending other courses during the semester (e.g. the lecture "Pedagogical Diagnostics", which promotes *Methodological Skills*). Further analyses are needed to be conducted here. The absence of response shift for *Skills in Reviewing the State of Research* and *Communication Skills* might have two reasons: Either students, at the beginning of the semester, were already able to

differentially evaluate these dimensions based on previous experience. Or students were still unable to make a differentiated assessment of themselves because literature research and the writing of academic publications and presentations did not have been explicitly promoted. Thus, the level of experience in these dimensions would still be low at the end of the semester. Then students would continue to overestimate themselves in posttest. Hence, this would result in no mean difference. In future studies, information about past research experiences could be integrated in a study.

Since group-specific differences between students in RLP and CG only occurred for self-assessed competences but not for tested competences, the question remains whether the RLP-format did not fulfil its purpose to prepare for EBP. We are not of this opinion. We rather interpret the results as an indicator for *using research* to be already an integral part of teacher training studies since the announcement by the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany in 2004. In both groups, there was an increase in *Information Literacy* and *Statistical Literacy*. The competence test primarily examines *using research*, which was a component of both seminar concepts: In RLP-format, *using research* is the main topic phase *preparation for the research process*. In seminars of the CG, *using research* was addressed by imparting knowledge about teaching and learning and by transferring it into practice.

The difference between the seminar concepts seems to be rather in *establishing research*, which was especially important in the RLP-format - besides the aspect of *using research*. In RLP, students were provided with the opportunity to undergo the research process. This is also reflected in results of self-assessed competences: these students had a higher increase compared to students of CG. Regardless of the debate about the ability to assess oneself in a differentiated way, the assessment of one's own competence has an impact on one's own actions (Bach, 2013). Furthermore, the quasi-indirect measurement of competence may have an impact on self-efficacy beliefs (Hill & Betz, 2005). As self-assessed research competences are highly correlated with self-efficacy beliefs (Böttcher-Oschmann, Groß Ophoff & Thiel, 2019; Mertens & Gräsel, 2018), it would be a desirable aim that the self-evaluation of competence as well as the belief to successfully undergo the research process have positive effects on future EBP. In future studies, a follow-up study could be included in the design to examine if the effects were transferred into practice.

Especially striking are the missing effects for *Skills in Reflecting on Research Findings* and *Evidence-based Reasoning*. Reasoning is a highly complex process and so are the corresponding test items highly challenging (Groß Ophoff et al., 2014, 2017). Furthermore, reasoning is not explicitly promoted in RLP by an exercise. The interpretation of results and the reflection on the implications were only discussed in consultations on the basis of students' concrete findings. In future, we will adapt the RLP-format in that direction. There will be specific exercises for reflection, argumentation, how to relate one's own findings back to the state of research, and how to draw conclusions and implications for one's own practice. The work on *Scientific Reasoning* by Fischer et al. (2014) provides important information in this respect.

One strength of this study is its design. The longitudinal approach while students were in the field indicated good external validity. The combination of competence self-assessments and a competence test offered the opportunity to create a comprehensive picture of the increase in competence. Both survey methods have certain advantages and disadvantages. Hence, they complement each other well (Lucas & Baird, 2006). In addition, the operationalisations of the R-comp and the ERL-test made it possible to assess both *using research* and *establishing research* (Davies, 1999). To develop a competence test to assess *establishing research* would certainly be a great challenge. Furthermore, the additional use of a retrospective pretest offered the possibility of inspecting over- and underestimation as well as response shift in the longitudinal assessment of competences. However, it should be noted that in retrospective measurement other bias may occur, e.g. social desirability, recall bias and implicit theory of change (Hill & Betz, 2005). It is therefore one of the study's strengths that a combination of these survey methods was used. In general, the combination of competence testing, self-assessment, and regarding a retrospective pretest provides comprehensive information. A combination of these survey methods should therefore be maintained in future studies. The addition of a control group was useful as an increase in competences due to general development processes or in the course of studies seems less probable. Hence, the study has a good internal validity.

Some limitations to our study need to be considered. On the one hand, the explanatory power is limited by the small sample size in both groups. This can be seen, for example, from the fact that sometimes large effects did not reach the significance level. Although a larger

sample was originally planned, due to the high dropout only a small number of students could be reached at both assessment time points. Students may have been saturated with other surveys that took place at university at the same time (Sax et al., 2003). Furthermore, invariance testing for response shift was not possible due to small sample size. This would be useful to compare results from our design-based approach to results from a statistical approach (Piwovar & Thiel, 2014; Schwartz & Sprangers, 1999).

Additionally, further analyses are necessary to empirically confirm the results found here. An implementation check should be carried out to take account of composition effects, influence of parallel courses attended, teaching subject combinations, school type or previous experience in the research process. Moreover, due to the small sample size, the hierarchical structure of the data has not been considered.

To secure the results, especially against influences of other factors, an experimental design would have to be carried out. However, this is difficult to implement in higher education. An optimisation of the quasi-experimental design could, however, include the choice of a different control group. Teacher training students who complete their practical semester abroad take part in the RLP at a different time, often during semester break. A quasi-experimental design with a real control group could therefore be carried out. The students in the RLP would then be the experimental group, the students in the semester break who do not attend any further courses (and start the RLP after semester break), as the control group. In this way, the influence of other factors on competence increase could be minimised.

5.5 Conclusion

Our study provides evidence that students can be prepared for EBP in RLP. The learning environment based on the ERL-model and the RMRC-K-model offered the opportunity to acquire the necessary competences during teacher training studies. Students as future teachers are prepared to apply methods of self- an external evaluation to develop and control quality of teaching and school and know how to reflect and evaluate results of educational research for their own practice. Hence, they are enabled to reflect on their own professional experiences and competences in order to draw conclusions from this reflection, and to use the results of educational research for their own professional practice.

5.6 References

- Bach, A. (2013). *Kompetenzentwicklung im Schulpraktikum. Ausmaß und zeitliche Stabilität von Lerneffekten hochschulischer Praxisphasen* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 87, 1. Aufl.). Münster: Waxmann.
- Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429.
- Böttcher, F. & Thiel, F. (2018). Evaluating research-oriented teaching: A new instrument to assess university students' research competences. *Higher Education*, 75(1), 91–110.
- Böttcher, F. & Thiel, F. (2016). Der Fragebogen zur Erfassung studentischer Forschungskompetenzen - Ein Instrument auf der Grundlage des RMRK-W-Modells zur Evaluation von Formaten forschungsorientierter Lehre. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre* (I 2.11, S. 57–74). Berlin: DUZ Medienhaus.
- Böttcher-Oschmann, F., Groß Ophoff, J. & Thiel, F. (2019). Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen – Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements. *Unterrichtswissenschaft*, 1–27. <https://doi.org/10.1007/s42010-019-00053-8>
- Brew, A. (2006). *Research and teaching. Beyond the divide* (Universities into the 21st century). Houndmills, Großbritannien: Palgrave Macmillan.
- Butts, R. E. (1991). Methodology, metaphysics and the pragmatic unity of science. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 23–28). Berlin: W. de Gruyter.
- Chevalier, A., Gibbons, S., Thorpe, A., Snell, M. & Hoskins, S. (2009). Students' academic self-perception. *Economics of Education Review*, 28(6), 716–727.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship. Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Hrsg.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (S. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Cramer, C. (2010). Kompetenzerwartungen Lehramtsstudierender. Grenzen und Perspektiven selbsteingeschätzter Kompetenzen in der Lehrerbildungsforschung. In A. Gehrmann, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle. Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 85–97). Bad Heilbrunn: Klinkhardt Verlag.
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108–121.
- Deci, E. L. & Ryan, R. M. (2000). The " what" and " why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
- Field, A. (2009). *Discovering statistics using SPSS. (and sex and drugs and rock 'n' roll)* (3. Aufl.). Los Angeles, CA: SAGE.
- Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R. et al. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, 2(3), 28–45.
- Groß Ophoff, J., Schladitz, S., Lohrmann, K. & Wirtz, M. (2014). Evidenzorientierung in bildungswissenschaftlichen Studiengängen. In K. Drossel, R. Strietholt & W. Bos (Hrsg.), *Empirische Bildungsforschung und evidenzbasierte Reformen im Bildungswesen* (S. 251–275). Münster: Waxmann.
- Groß Ophoff, J., Wolf, R., Schladitz, S. & Wirtz, M. (2017). Assessment of educational research literacy in higher education: Construct validation of the factorial structure of an assessment instrument comparing different treatments of omitted responses. *Journal for Educational Research Online/Journal für Bildungsforschung Online*, 9(2), 37–68.
- Hammersley, M. (1997). Educational research and teaching: a response to David Hargreaves' TTA lecture. *British Educational Research Journal*, 23(2), 141–161.
- Hammersley, M. (Ed.). (2007). *Educational research and evidence-based practice*. Los Angeles, CA: SAGE.
- Hargreaves, D. H. (1996). *Teaching as a research-based profession: Possibilities and prospects*. Teacher Training Agency Annual Lecture, London.
- Hargreaves, D. H. (1997). In defence of research for evidence-based teaching: a rejoinder to Martyn Hammersley. *British Educational Research Journal*, 23(4), 405–419.

- Hartig, J. & Kühnbach, O. (2006). Schätzung von Veränderung mit „plausible values“ in mehrdimensionalen Rasch-Modellen. In A. Ittel & H. Merkens (Hrsg.), *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* (1. Aufl., S. 27–44). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH Wiesbaden.
- Hill, L. G. & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation*, 26(4), 501–517.
- Hollingsworth, M. A. & Fassinger, R. E. (2002). The role of faculty mentors in the research training of counseling psychology doctoral students. *Journal of Counseling Psychology*, 49(3), 324–330.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.
- Kultusministerkonferenz (2004). *Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004.*
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 1–12.
- Lucas, R. E. & Baird, B. M. (2006). Global self-assessment. In M. Eid & E. Diener (Hrsg.), *Handbook of multimethod measurement in psychology* (S. 29–42). Washington, DC: American Psychological Association.
- Mabe, P. A. & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280–296.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543.
- Mertens, S. & Gräsel, C. (2018). Entwicklungsbereiche bildungswissenschaftlicher Kompetenzen von Lehramtsstudierenden im Praxissemester. *Zeitschrift für Erziehungswissenschaft*, 21(6), 1109–1133.
- Meyer, T., Richter, S. & Raspe, H. (2013). Agreement between pre-post measures of change and transition ratings as well as then-tests. *BMC Medical Research Methodology*, 13(52), 1–10.

- Mittelhaeuser, M.-A., Béguin, A. A. & Sijtsma, K. (2011). *Comparing the effectiveness of different linking designs: The internal anchor versus the external anchor and pre-test data*. Arnheim, Niederlande: Cito.
- Mittelstraß, J. (1991). Einheit und Transdisziplinarität. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 12–22). Berlin: W. de Gruyter.
- Moon, B. (2000). The changing agenda for professional development in education. In B. Moon, J. Butcher & E. Bird (Hrsg.), *Leading professional development in education* (S. 3–8). London: RoutledgeFalmer.
- Oort, F. J. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research*, 14(3), 587–598.
- Palisano, R. J. (2006). A collaborative model of service delivery for children with movement disorders: a framework for evidence-based decision making. *Physical Therapy*, 86(9), 1295–1305.
- Piwovar, V. & Thiel, F. (2014). Evaluating response shift in training evaluation: comparing the retrospective pretest with an adapted measurement invariance approach in a classroom management training program. *Evaluation Review*, 38(5), 420–448.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, 312(7023), 71–72.
- Sax, L. J., Gilmartin, S. K. & Bryant, A. N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education*, 44(4), 409–432.
- Schmidt-Atzert, L. & Amelang, M. (2012). Einleitung. In L. Schmidt-Atzert & M. Amelang (Hrsg.), *Psychologische Diagnostik* (5. Aufl., S. 1–33). Berlin: Springer.
- Schwartz, C. E. & Sprangers, M. A. G. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11), 1531–1548.
- Schwartz, C. E. & Sprangers, M. A. G. (2010). Guidelines for improving the stringency of response shift research using the thentest. *Quality of Life Research*, 19(4), 455–464.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93–105.

- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23.
- Sprangers, M. A. G. & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research. A theoretical model. *Social Science & Medicine*, 48(11), 1507–1515.
- Strobl, C. (2010). *Das Rasch-Modell. Eine verständliche Einführung für Studium und Praxis* (Sozialwissenschaftliche Forschungsmethoden, Bd. 2, 1. Aufl.). München: Hampp.
- Thiel, F. & Böttcher, F. (2014). Modellierung fächerübergreifender Forschungskompetenzen - Das RMKR-W-Modell als Grundlage der Planung und Evaluation von Formaten forschungsorientierter Lehre. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre* (I 2.10, S. 109–124). Berlin: Raabe.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Wright, B. D. & Douglas, G. A. (1996). Estimating Rasch (person, ability, theta) measures with known dichotomous item difficulties: Anchored Maximum Likelihood Estimation (AMLE). *Rasch Measurement Transactions*, 10(2), 499. Zugriff am 03.08.2018.
Verfügbar unter <https://www.rasch.org/rmt/rmt102t.htm>



Kapitel 6

Gesamtdiskussion

6 Gesamtdiskussion

Forschungsorientierte Lehre (FoL) stellt eine Möglichkeit zur Stärkung des Research-Teaching-Nexus (Brew, 2006; Griffiths, 2004; Healey, 2005; Healey & Jenkins, 2009) dar, indem Forschung in die universitäre Lehre integriert wird und studentische Forschungskompetenzen gefördert werden können. Studentische Forschungskompetenzen stellen als Spezialform des wissenschaftlichen Denkens und Argumentierens einen wichtigen generischen Outcome universitärer Lehre dar, und zwar sowohl für den wissenschaftlichen Nachwuchs als auch für evidenzbasierte ExpertInnen (vgl. Kapitel 2.1.2). Kompetenzmodelle eignen sich als Grundlage für eine universitätsweite Implementation von FoL-Programmen zur Förderung von Forschungskompetenzen, wie sie im Rahmen der Exzellenzinitiative vielfach entwickelt wurden (Deutsche Forschungsgemeinschaft & Wissenschaftsrat, 2015). Diese bieten einerseits die Möglichkeit, Kompetenzziele zu definieren und darauf abgestimmte Lehrpläne zu entwickeln (Erpenbeck & von Rosenstiel, 2007; Hetmanek et al., 2018; Klieme & Leutner, 2006). Andererseits können Kompetenzmodelle die Grundlage für die Entwicklung von Instrumenten zur Erfassung von Kompetenzen bilden (Erpenbeck & von Rosenstiel, 2007; Klieme & Hartig, 2008).

Die vorliegende Arbeit ist im Rahmen eines interdisziplinären FoL-Programms entstanden und hatte das Ziel, ein valides Instrument zur universitätsweiten Erfassung von Forschungskompetenzen bereitzustellen. Dieses Instrument sollte zur Evaluation der Wirksamkeit dieser forschungsorientierten Lehr-Lernformate eingesetzt werden. Die Grundlage dieses Evaluationsansatzes bildet das RMRK-W-Modell (Thiel & Böttcher, 2014) mit seinen Dimensionen *Recherche-*, *Methoden-*, *Reflexions-* und *Kommunikationskompetenzen* sowie *Fachliches Wissen*. Diese Dimensionen spiegeln ausgehend von der pragmatischen Wissenschaftstheorie (vgl. Kapitel 2.2.1, 3.1 und 4.1; Butts, 1991; Dewey, 1938; Mittelstraß, 1991) fach- und disziplinübergreifend gemeinsame Aspekte des Forschungsprozesses wider. In Anlehnung an die funktional-pragmatische Auffassung von Kompetenzen (vgl. Kapitel 2.1.1 und 3.1) werden Kompetenzen im RMRK-W-Modell als kognitive Dispositionen aufgefasst (Klieme, Hartig & Rauch, 2008), bestehend aus Wissen und Fertigkeiten (Nusche, 2008; Simonton, 2003) bzw. aus deklarativem und prozeduralem Wissen (R. Mayer, 2003). Ausgehend vom RMRK-W-Modell wurde der Fragebogen zur Erfassung studentischer Forschungskompetenzen (F-Komp) entwickelt, den es in dieser Arbeit zu validieren galt. Dieser Fragebogen umfasst 32 Items in fünf Skalen entsprechend des Kompetenzmodells und erfasst Forschungskompetenzen via Selbsteinschätzungen. Zwar

bringen Kompetenzselbsteinschätzungen einige Herausforderungen bei deren Messung mit sich (z.B. Verzerrungstendenzen, Fehleinschätzungen oder Response-Shift; Chevalier, Gibbons, Thorpe, Snell & Hoskins, 2009; Kruger & Dunning, 1999; Lucas & Baird, 2006; Mertens & Gräsel, 2018; Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999; vgl. Kapitel 4.1.3 und 5.1.4), allerdings eignen sie sich insbesondere für eine fach- und disziplinübergreifende Erfassung von Forschungskompetenzen (vgl. Kapitel 2.2.2). Im Vorfeld dieser Arbeit wurde bereits eine Inhaltsvalidierung vorgenommen (Thiel & Böttcher, 2014), die in Studie 1 im Rahmen der Darstellung des Konstruktions- und Überarbeitungsprozesses des F-Komp vorgestellt wurde (vgl. Kapitel 3.1.2). ExpertInnen aus unterschiedlichen Disziplinen beurteilten die Items des F-Komp hinsichtlich ihrer Passung für das eigene Fach. Damit konnte bestätigt werden, dass aus Sicht der ExpertInnen einerseits die Dimensionen des RMRK-W-Modells und damit auch die theoretisch hergeleiteten, fach- bzw. disziplinübergreifenden Phasen des Forschungsprozesses für verschiedene Disziplinen zutreffen. Andererseits wurde bestätigt, dass die Items des F-Komp in den Disziplinen der ExpertInnen zur Erfassung von Forschungskompetenzen geeignet sind.

Ein weiteres Ziel dieser Arbeit war es, den F-Komp einzusetzen, um eine neu konzipierte forschungsorientierte Lehrveranstaltung zu evaluieren. Diese Veranstaltung sollte so gestaltet sein, dass der Kompetenzerwerb begünstigt wird (vgl. Kapitel 2.3.1). Als eine besondere Zielgruppe wurden Lehramtsstudierende identifiziert, die bereits im Laufe ihres Studiums Forschungskompetenzen für eine evidenzbasierte Praxis (vgl. Kapitel 2.3.2) erwerben sollen.

Im Rahmen dieser Arbeit wurden die beiden zuvor dargestellten Desiderata anhand der beiden Fragestellungen aus Kapitel 2.4 bearbeitet:

- *Forschungsfrage 1:* Können Forschungskompetenzen mit dem F-Komp valide erfasst werden (Studie 1 und 2)?
- *Forschungsfrage 2:* Lassen sich Kompetenzen von Lehramtsstudierenden für eine evidenzbasierte Praxis in einer kompetenzmodellbasierten, forschungsorientierten Lernumgebung fördern (Studie 3)?

6.1 Zusammenfassung und Diskussion der Ergebnisse

Die Ergebnisse der drei durchgeführten empirischen Studien wurden bereits im Einzelnen diskutiert. Daher erfolgt nun eine Zusammenfassung der Ergebnisse entlang der in Kapitel 2.4 dargestellten Forschungsfragen. Im Anschluss daran werden die Befunde integrierend diskutiert.

Forschungsfrage 1: Können Forschungskompetenzen mit dem F-Komp valide erfasst werden?

Im Rahmen dieser Arbeit lag der Fokus auf der Konstruktvalidierung des F-Komp. Die Ergebnisse der Studie 1 (Kapitel 3) haben gezeigt, dass der F-Komp *faktoriell valide* ist. Es zeigte sich, dass sich die fünf auf Grundlage der pragmatischen Wissenschaftstheorie (Butts, 1991; Dewey, 1938; Mittelstraß, 1991) identifizierten Dimensionen des RMRK-W-Modells (Thiel & Böttcher, 2014) *Recherche-, Methoden-, Reflexions- und Kommunikationskompetenzen* sowie *Fachliches Wissen* auch empirisch mit dem F-Komp abbilden lassen. Darüber hinaus wurden zwölf Subdimensionen identifiziert, welche die Facetten des RMRK-W-Modells widerspiegeln und zum Teil kombinieren²¹. Damit werden die deklarativen und prozeduralen Wissensbereiche (R. Mayer, 2003) bzw. das Wissen und die Fertigkeiten (Nusche, 2008; Simonton, 2003) zur Absolvierung der Anforderungen des Forschungsprozesses empirisch abgebildet. Es standen jedoch noch zwei Alternativmodelle zur Auswahl, die prinzipiell ebenfalls annehmbare Fit-Indizes aufwiesen: ein Modell mit einem Faktor höherer Ordnung und ein Bi-Faktor-Modell. Aufgrund der teils knappen Entscheidung gegen diese möglichen Alternativmodelle wurde vermutet, dass die Faktorenstruktur in Abhängigkeit der Stufen des Studiums (Bachelor, Master und Promotion) und damit potenziell auch in Abhängigkeit der Kompetenzniveaus der Studierenden stehen könnte. In der Stichprobe befanden sich Studierende auf drei unterschiedlichen Stufen des Studiums und damit potenziell auch auf drei unterschiedlichen Erfahrungsstufen: Bachelorstudierende, die als NovizInnen bezeichnet werden können (Simonton, 2003) und bislang wahrscheinlich wenig Forschungserfahrungen sammeln konnten, Masterstudierende, welche wahrscheinlich erste Forschungserfahrungen sammeln konnten sowie Promotionsstudierende, die wahrscheinlich die meisten Forschungserfahrungen, möglicherweise bereits auch in eigenen Forschungsprojekten, sammeln konnten.

²¹ Eine detaillierte Darstellung dieser Sub-Dimensionen ist in Kapitel 3.4 und 4.1.1 sowie den Abbildungen 3.2 bis 3.4 zu finden.

In Studie 2 (Kapitel 4) wurden zur Überprüfung der *konvergenten* und *diskriminanten Validität* zwei Forschungsfragen bearbeitet. Zum einen wurde überprüft, dass mit dem F-Komp keine allgemeinen Lern- und Studienkompetenzen erfasst werden. Zum anderen wurde die Validität im Hinblick auf die gewählte Methodik der Selbsteinschätzung überprüft. Die Ergebnisse in Bezug auf die erste Forschungsfrage haben gezeigt, dass mit dem F-Komp *Forschungskompetenzen* und keine allgemeinen Lern- und Studienkompetenzen gemessen werden. Dabei zeigte sich, dass mit dem F-Komp keine zeitstabilen, kontext- und domänenunabhängigen kognitiven Fähigkeiten (Klieme & Hartig, 2008; Weinert, 2001), wie visuelle Wahrnehmung und logisches Schlussfolgern, erfasst wurden. Die in Abbildung 2.2 dargestellte Systematik von *domänenübergreifenden* vs. *domänenunabhängigen* Kompetenzen (in Anlehnung an Hetmanek et al., 2018) scheint in diesem Fall zuzutreffen. Weiterhin zeigte sich, dass prozessbezogene kognitive Aspekte selbstgesteuerten Lernens, wie Elaboration und kritisches Prüfen, relevant sind, um den Forschungsprozess zu durchlaufen, sich jedoch auch von Forschungskompetenzen abgrenzen lassen. Dieses Ergebnis lässt sich ebenfalls gut in diese Systematik integrieren, da sowohl selbstreguliertes Lernen als auch Forschungskompetenzen als *domänenübergreifende* Kompetenzen eingeordnet werden können, allerdings auf unterschiedlichen Ebenen der Allgemeingültigkeit (Abbildung 2.2). Weiterhin deutete sich an, dass Überzeugungen hinsichtlich der Unsicherheit des wissenschaftlichen Wissens (Hofer & Pintrich, 1997; Schommer, 1998; Trautwein & Lüdtke, 2007) tendenziell wichtig sind für das Durchlaufen des Forschungsprozesses, da im Rahmen dieses dynamischen Problemlöseprozesses bestehendes Wissen durch neues Wissen abgelöst werden kann (Trautwein & Lüdtke, 2007).

Darüber hinaus wurde mit der zweiten Forschungsfrage der Studie 2 die Validität der gewählten Methodik der Selbsteinschätzung von Kompetenzen überprüft. Dazu wurden in einem ersten Teil dieses Validierungsaspektes Zusammenhänge zu Forschungskompetenzen untersucht, die mit einer anderen Methode erfasst wurden, nämlich einem Kompetenztest (bildungswissenschaftliche Forschungskompetenz; Groß Ophoff, Schladitz, Lohrmann & Wirtz, 2014; Groß Ophoff, Wolf, Schladitz & Wirtz, 2017). In einem zweiten Teil dieses Validierungsaspektes wurden andererseits Zusammenhänge zu einem verwandten Konstrukt untersucht, welches ebenfalls über die Methode der Selbsteinschätzung erfasst wird (forschungsbezogene Selbstwirksamkeitserwartungen). Im Falle des Kompetenztests wurden kleine bis mittlere Zusammenhänge zu den disziplinspezifisch erfassten bildungswissenschaftlichen Forschungskompetenzen gefunden. Dies sind zwar höhere

Zusammenhänge als in vergleichbaren Studien (für Forschungskompetenzen: Schladitz, Groß Ophoff & Wirtz, 2015; für Selbsteinschätzung vs. Kompetenztest: Kunter & Klusmann, 2010; König & Tachtsoglou, 2012; für Selbsteinschätzung und externe Kriterien: Braun, 2008; Cassidy & Eachus, 2000), jedoch stellen sie keinen Nachweis der Übereinstimmung beider Konstrukte dar. Es wurde vermutet, dass die Fehleinschätzung von Kompetenzen diesen Ergebnissen zugrunde liegen könnte (Kapitel 4.4). Diese Fehleinschätzung wiederum ist nach Ansicht von Mertens und Gräsel (2018) abhängig von der bisherigen Erfahrung bzw. nach Ansicht von Kruger und Dunning (1999) abhängig von der eigenen Kompetenz. Mit zunehmender Erfahrung können Kompetenzselbsteinschätzungen differenzierter vorgenommen werden (Bach, 2013; Mertens & Gräsel, 2018). Darüber hinaus *überschätzen* inkompetente Personen ihre Kompetenzen stärker als kompetente Personen, kompetente Personen hingegen *unterschätzen* ihre Kompetenzen im Verhältnis zu ihren Peers (Kruger & Dunning, 1999).

Im zweiten Teil des Validierungsaspektes hinsichtlich der gewählten Methodik, im Falle des verwandten Konstruktes, zeigten Forschungskompetenzen große Zusammenhänge zu forschungsbezogenen Selbstwirksamkeitserwartungen. Damit ist die Abgrenzung zu einem Konstrukt, welches ebenfalls über die Methode der Selbsteinschätzung erfasst wird, nicht gelungen. Es stellte sich die Frage, was bei der Selbsteinschätzung von Kompetenzen tatsächlich gemessen wird. Cramer (2010) spricht davon, dass dabei weniger die eigene Kompetenz als vielmehr *Kompetenzerwartungen* gemessen werden. Dies würde bedeuten, dass mit dem F-Komp weniger *generische Fertigkeiten* (und fachliches Wissen zum Forschungsprozess) als vielmehr nicht-kognitive Outcomes gemessen werden, die unter der Kategorie *motivationale, volitionale und/oder soziale Outcomes* zusammengefasst wurden (vgl. Kapitel 2.1.1; Nusche, 2008; Weinert, 2001; Zlatkin-Troitschanskaia, Pant, Kuhn, Toepper & Lautenbach, 2016). Sowohl Kunter und Klusmann (2010) als auch Frey (2006) fassen jedoch Selbsteinschätzungen als subjektive Methode der Kompetenzmessung auf (vgl. Bach, 2013), sodass mit diesen Selbsteinschätzungen *generische Fertigkeiten* (und im Fall des F-Komp auch fachliches Wissen zum Forschungsprozess) erfasst werden sollten. Daher bleibt die Frage bestehen, wie die hohen Zusammenhänge der beiden Konstrukte zustande gekommen sein könnten. Letztlich wurde auch für dieses Ergebnis die Vermutung geäußert (vgl. Kapitel 4.4), dass das Kompetenzniveau einen Einfluss auf die Fähigkeit zur Unterscheidung der *aktuellen* Kompetenz (Kompetenzselbsteinschätzung) und der Überzeugung über die *künftige* Kompetenz (Selbstwirksamkeitserwartungen) haben könnte.

Zur Beantwortung der ersten Forschungsfrage wurde in den ersten beiden Studien dieser Arbeit die Konstruktvalidität überprüft. Wird die Inhaltsvalidierung, welche im Rahmen des Überarbeitungsprozesses durchgeführt wurde (Thiel & Böttcher, 2014), einbezogen, hat der F-Komp zwei der in Kapitel 2.2.2 dargestellten Schritte der Validitätsüberprüfung durchlaufen. Damit ist sowohl eine theoretisch-argumentative Beurteilung des Instrumentes erfolgt als auch in ersten Schritten eine empirisch-statistische Überprüfung, dass mit dem F-Komp auf das interessierende, theoretische Konstrukt Forschungskompetenzen geschlossen werden kann (Döring & Bortz, 2016).

Forschungsfrage 2: Lassen sich Kompetenzen von Lehramtsstudierenden für eine evidenzbasierte Praxis in einer kompetenzmodellbasierten, forschungsorientierten Lernumgebung fördern?

In Studie 3 (Kapitel 5) wurde eine forschungsorientierte Lernumgebung in Form eines Lernforschungsprojektes entwickelt. Angehende Lehrkräfte sollten auf ihre Rolle als evidenzbasierte ExpertInnen (Kultusministerkonferenz, 2004) vorbereitet werden, für die sie sowohl bereits bestehende Evidenz *nutzen* sollen (*using research*; Davies, 1999) als auch selbst Evidenz *generieren* sollen (*establishing research*; Davies, 1999). Die forschungsorientierte Lernumgebung wurde auf Basis von zwei Kompetenzmodellen entwickelt: dem generischen RMRK-W-Modell (Thiel & Böttcher, 2014), welches sowohl die Nutzung als auch das Generieren von Evidenz beinhaltet, sowie dem disziplinspezifischen Modell zur bildungswissenschaftlichen Forschungskompetenz (Groß Ophoff et al., 2014, 2017), welches insbesondere die Nutzung von Evidenz durch Lehrkräfte beinhaltet. Die Lernumgebung wurde auf kognitiver, motivationaler und sozialer Ebene so gestaltet, dass der Erwerb von Kompetenzen zur Nutzung und zum Generieren von Evidenz gefördert werden sollte: Als kognitive Komponenten wurden insbesondere das *Modeling*, *Coaching* und *Scaffolding* aus dem Cognitive-Apprenticeship-Ansatz (Collins, Brown & Newman, 1989) eingesetzt; als motivationale Komponenten wurden Kompetenzerleben, Autonomieerfahrungen und soziale Eingebundenheit ermöglicht (Deci & Ryan, 2000); als soziale Komponente wurde die Gruppenarbeit mit Unterstützung durch ExpertInnen gewählt. Anhand konkreter Aufgaben haben die Studierenden den gesamten Forschungsprozess durchlaufen: So sollten sie beispielsweise Daten erheben und analysieren sowie ein Poster zur Präsentation ihrer Ergebnisse erstellen. Der Kompetenzerwerb wurde erfasst durch die Kombination von Kompetenzselbsteinschätzungen und einem Kompetenztest, was in der Literatur angeraten wird, um mögliche Schwächen der Erhebungsmethoden

auszugleichen (Lucas & Baird, 2006). Dabei fokussierte der Kompetenztest vor allem die Nutzung von Evidenz (Groß Ophoff et al., 2014, 2017), während die Kompetenzselbsteinschätzung mit dem F-Komp sowohl die Nutzung als auch das Generieren von Evidenz umfasst. Zudem wurde das Response-Shift-Phänomen (Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999) berücksichtigt, welches bei Selbsteinschätzungen in längsschnittlichen Studien auftreten kann und eine Veränderung der Kriterien für die Selbsteinschätzung eines internalen Konstruktes beschreibt.

Die Ergebnisse dieser Studie deuteten darauf hin, dass es keine Unterschiede zwischen den beiden untersuchten Gruppen im Kompetenztest gab, gleichwohl aber Unterschiede in der Selbsteinschätzung von Forschungskompetenzen mit dem F-Komp: Die Studierenden der Lernforschungsprojekte zeigten insbesondere in denjenigen Kompetenzbereichen einen Zuwachs, die auf das Generieren von Evidenz abzielen. Die Studierenden in der Kontrollgruppe zeigten hingegen einen Zuwachs in denjenigen Kompetenzbereichen, welche die Nutzung von Evidenz repräsentieren. Der Unterschied zwischen den Gruppen könnte daher auf die Ermöglichung von *Forschungserfahrung* zurückzuführen sein. Die vorliegenden Ergebnisse geben somit einen ersten Hinweis darauf, dass sich Kompetenzen für eine evidenzbasierte Praxis gezielt in einer forschungsorientierten Lernumgebung fördern lassen.

Die Ergebnisse der drei im Rahmen dieser Arbeit durchgeführten Studien verweisen darauf, dass das Kompetenzniveau der Studierenden bedingt durch deren bisherige Forschungserfahrung einen Einfluss auf die Selbsteinschätzung von Kompetenzen haben könnte: In Studie 1 deuteten die alternativen Faktorenstrukturen darauf hin, in Studie 2 die hohen Zusammenhänge zu forschungsbezogenen Selbstwirksamkeitserwartungen bzw. die mittleren Zusammenhänge zum Kompetenztest, in Studie 3 der Unterschied zwischen dem Kompetenzzuwachs der Studierenden der Lernforschungsprojekte im Gegensatz zu dem der Studierenden der Kontrollgruppe. Darüber hinaus sind allerdings auch motivationale Einflussfaktoren denkbar. Mögliche Einflussfaktoren werden im folgenden Abschnitt integrierend diskutiert.

Einflussfaktoren auf Selbsteinschätzungen von Kompetenzen

Die Selbsteinschätzung von Kompetenzen kann durch verschiedene Faktoren beeinflusst werden, beispielsweise durch Erinnerungseffekte (*Recall Bias*), durch implizite Theorien über Veränderungen oder durch soziale Erwünschtheit (Hill & Betz, 2005; Lucas & Baird, 2006; Schwartz & Sprangers, 2010). Darüber hinaus wäre es denkbar, dass

Wertkomponenten wie Interesse, Nützlichkeit, Wichtigkeit, Freude und Kosten/Aufwand (Thoren et al., in Druck) einen Einfluss auf die Ergebnisse der Studien haben könnten. Dies könnte insbesondere für diejenigen Studien zutreffen, in denen Lehramtsstudierende die Stichproben gebildet haben (insbesondere in Studie 2 und 3). Gerade (angehende) Lehrkräfte werden immer wieder dahingehend befragt, welchen Wert Forschungskompetenzen oder eine evidenzbasierte Praxis für sie hat (z.B. für Lehramtsstudierende: Thoren et al., in Druck) und wie sie Evidenz für ihre Praxis wirklich nutzen (z.B. für Lehrkräfte in der Praxis: Hetmanek et al., 2015). Diese Wertkomponenten stammen aus dem Erwartungs-Wert-Modell (Wigfield & Eccles, 2000): Neben der Einschätzung, wie groß die Wahrscheinlichkeit ist, eine Handlung erfolgreich durchführen zu können (Erwartungen) entscheidet auch die Einschätzung über die subjektive Bedeutung, die einer Handlung und ihren Folgen zugemessen wird (Wert), über die Bildung von Motivation. Daher könnten sich Wertkomponenten auf die Selbsteinschätzung von Kompetenzen auswirken.

Darüber hinaus ist ein konsistentes Ergebnis der drei Studien, dass das Kompetenzniveau (in Anlehnung an Kruger & Dunning, 1999) und die bisherige Erfahrung mit Forschung (in Anlehnung an Mertens und Gräsel, 2018) einen Einfluss auf die Kompetenzselbsteinschätzung haben könnten. Möglicherweise ist dies auf die Stufen des Kompetenzerwerbs zurückzuführen. Hinweise zu diesen Stufen bzw. zum „Erwerb komplexer kognitiver Fertigkeiten“ (Nückles & Wittwer, 2014, S. 241) lassen sich in verschiedenen Forschungsbereichen finden: beispielsweise bei Schraw (2006) im Kontext der Expertiseforschung, der jedoch die Stufen des Kompetenzerwerbs auf den Kontext des Verfassens wissenschaftlicher Publikationen überträgt, oder bei Anderson (1982), dessen Theorie auf Basis von „Wenn-Dann-Produktionsregeln“ (Renkl, 2015) zur Entwicklung von computerbasierten Lernprogrammen genutzt wird (Nückles & Wittwer, 2014). Grundlage beider Ansätze ist die Unterscheidung von deklarativem und prozeduralem Wissen, also dem Wissen über Fakten und Konzepte („Wissen, dass“) vs. dem Wissen über Strategien oder Algorithmen („Wissen, wie“; R. Mayer, 2003; Renkl, 2015).

Beide Autoren sprechen von drei Phasen oder Stufen, also unterschiedlichen Kompetenzniveaus: Auf dem ersten Niveau ist zunächst deklaratives Wissen vorhanden. Dies kann Wissen zu einzelnen Bereichen des Forschungsprozesses sein, wie beispielsweise Fakten und Konzepte zu Forschungsfragen, Forschungsmethoden, Datenanalyse, Interpretation von Daten, das Ziehen von Schlussfolgerungen (vgl. Schraw, 2006). Auf diesem Niveau werden zur Lösung von Aufgaben zu diesen Bereichen (z.B. zur Analyse von Daten) lediglich

allgemeine Problemlösemethoden angewandt (Anderson, 1982; Nückles & Wittwer, 2014). Eine dieser Problemlösemethoden kann das analoge Schließen sein (Nückles & Wittwer, 2014), bei dem allerdings viele kognitive Kapazitäten verbraucht werden. Auf diesem Niveau werden beispielsweise Korrelationsanalysen deswegen durchgeführt, weil sich diese in vergangenen, ähnlichen Situationen zur Analyse des Zusammenhangs zwischen Variablen bewährt haben.

Auf dem zweiten Niveau werden einzelne deklarative Wissensbestandteile mit entsprechenden prozeduralen Wissensbereichen zusammengeführt in Form von Schemata oder Skripten (Anderson, 1982; Nückles & Wittwer, 2014; Schraw, 2006). Bewähren sich beispielsweise Korrelationsanalysen als Analysemethode für Zusammenhänge zwischen Variablen, so werden entsprechende deklarative und prozedurale Wissensbestandteile zur Korrelationsanalyse gemeinsam in einem Schema abgespeichert. Dieses Schema kann dann wiederum mit anderen Schemata, z.B. zu Mittelwertanalysen, in ein Skript zur Analyse deskriptiver Statistiken integriert werden. Verschiedene Skripte lassen sich wiederum zu größeren konzeptuellen Modellen zusammenfügen (Schraw, 2006), z.B. zu einem konzeptuellen Modell zur Datenanalyse. Zur Lösung von Aufgaben werden auf dieser Stufe keine allgemeinen Problemlösemethoden mehr angewendet. Es wird stattdessen auf das entsprechende deklarative Wissen zurückgegriffen, was eine schnellere Bearbeitung von Aufgaben ermöglicht (Anderson, 1982; Nückles & Wittwer, 2014).

Auf dem dritten Niveau werden diese Schemata, Skripte und konzeptuellen Modelle weiter eingeübt. Es fällt einerseits leichter, in einer neuen Situation zu entscheiden, auf welche Skripte zurückgegriffen wird, welche Skripte sich bewährt haben und welche davon wiedereingesetzt werden (Anderson, 1982; Nückles & Wittwer, 2014). Andererseits werden durch das Einüben nur noch geringe kognitive Kapazitäten zur Durchführung einer Aufgabe benötigt, sodass diese Kapazitäten für andere Prozesse verwendet werden können (Schraw, 2006). So würde im Fall der Datenanalyse das Skript zur Analyse deskriptiver Statistiken effizienter durchgeführt werden können. Darüber hinaus könnten einerseits noch kognitive Kapazitäten für komplexere Analysen verbleiben. Andererseits ist auch denkbar, dass noch kognitive Kapazitäten für Prozesse in anderen konzeptuellen Modellen, wie z.B. der Interpretation von Daten, genutzt werden können.

In Studie 1 könnten sich die Kompetenzniveaus in den unterschiedlichen Alternativmodellen widerspiegeln und damit auch abbilden, inwiefern Studierende vom

Forschungsprozess als einem zyklischen, zusammenhängenden Prozess ausgehen: Ein Modell mit fünf Dimensionen und ausdifferenzierten Subdimensionen könnte für Bachelorstudierende zutreffen, die sich auf einem niedrigeren Kompetenzniveau befinden. Diese ausdifferenzierten Subdimensionen könnten ein Anzeichen dafür sein, dass die Studierenden die einzelnen Phasen des Forschungsprozesses als separate Schritte empfinden, und zwar aufgrund fehlender Verknüpfungen von deklarativen mit prozeduralen Wissensbeständen in Schemata, Skripten und konzeptuellen Modellen. So wäre es für Studierende auf diesem Niveau wahrscheinlich schwierig, zwei komplexe Prozesse, wie einerseits das Recherchieren von Literatur und andererseits das Erstellen eines Forschungsdesigns, gleichzeitig durchzuführen. Im Gegensatz dazu könnte, wie in der Diskussion der ersten Studie angedeutet (vgl. Kapitel 3.4), ein Modell mit einem Faktor höherer Ordnung ohne Subdimensionen bzw. ein Bi-Faktormodell für erfahrenere Studierende auf einem höheren Kompetenzniveau zutreffen. Diese Verbindung der einzelnen Kompetenzaspekte könnte darauf hindeuten, dass Studierende über bereits angelegte Schemata und Skripte und damit über die kognitiven Kapazitäten verfügen, den Forschungsprozess in einem größeren konzeptuellen Modell zu sehen. So ließen sich beispielsweise bereits während der Literaturrecherche Hinweise auf die Gestaltung des eigenen Forschungsdesigns finden. Sollte das Kompetenzniveau einen Einfluss auf die Faktorenstruktur haben, ließen sich möglicherweise diese Niveaus auch in Kompetenzstufen modellieren.

In Studie 2 scheinen sich die erzielten Ergebnisse an die Annahmen von Mertens und Gräsel (2018) bzw. Kruger und Dunning (1999) anzuschließen, dass das Erfahrungs- bzw. Kompetenzniveau die Fehleinschätzung von Kompetenzen beeinflusst. Eine *Unterschätzung* von Kompetenzen, die nach Kruger und Dunning (1999) auf einem hohen Kompetenzniveau auftritt, wird den Ergebnissen wahrscheinlich nicht zugrunde liegen, da in den Stichproben dieser Studie vorrangig Masterstudierende vertreten waren. Ein hohes Kompetenzniveau würde wohl eher auf einem Promotionsniveau oder gar darüber, ab einem Post-Doc-Niveau, vorkommen. Es wird sich wahrscheinlich vielmehr um die *Überschätzung* von Kompetenzen handeln aufgrund eines niedrigen Kompetenzniveaus (Kruger & Dunning, 1999) bzw. um die undifferenzierte Einschätzung aufgrund fehlender Erfahrung (Bach, 2013; Mertens & Gräsel, 2018) mit Forschung. Auf diesem Niveau sind sicherlich schon deklarative Wissensbestände zu den einzelnen Phasen des Forschungsprozesses vorhanden. Womöglich gibt es erste Verknüpfungen mit den prozeduralen Wissensbeständen in Schemata oder Skripten (Schraw, 2006). Nach Kruger und Dunning (1999) wird die Überschätzung der eigenen Kompetenz erst

mit steigendem Kompetenzniveau rückblickend erkannt. Übertragen auf ein Beispiel des Forschungsprozesses bedeutet die Überschätzung von Kompetenzen Folgendes: Auf einem geringen Kompetenzniveau kann möglicherweise noch keine differenzierte Einschätzung zu einer Frage wie „Ich bin in der Lage, den Forschungsstand zu einem bestimmten Thema zu sichten und systematisch aufzubereiten“ abgegeben werden. Dazu würde deklaratives *und* prozedurales Wissen, integriert in Schemata und Skripten gehören, und zwar zu Bereichen wie beispielsweise Recherchestrategien, Kenntnis über wichtige Datenbanken, Wissen und Fertigkeiten zur Einschätzung der Plausibilität und Qualität der recherchierten Literatur, etc. (vgl. Kapitel 2.2.1; Trempler et al., 2015). Dass alle diese Wissensbestände zur Sichtung und Aufbereitung des Forschungsstandes gehören, wird vermutlich mit zunehmender Erfahrung erkannt, sodass dann erst eine differenzierte Selbsteinschätzung möglich sein sollte.

Wie könnte sich diese Überschätzung nun auf die Ergebnisse der Zusammenhänge zum Kompetenztest und zu den forschungsbezogenen Selbstwirksamkeitserwartungen auswirken? Für die Bearbeitung des eingesetzten Kompetenztests (Groß Ophoff et al., 2014, 2017) sind neben deklarativen auch prozedurale Wissensbestandteile notwendig. So werden beispielsweise Merkmale von verschiedenen Messinstrumenten wie Fragebögen, Interviews und Beobachtungsverfahren vorgegeben (deklaratives Wissen). Die Testpersonen müssen dann entscheiden, welches Verfahren für bestimmte Kontexte der empirischen Sozialforschung geeignet ist (prozedurales Wissen). Möglicherweise führt die Annahme, eine solche Aufgabe richtig gelöst zu haben, dazu, dass die eigene Kompetenz in diesem Bereich, z.B. der Auswahl von Forschungsmethoden, überschätzt wird.

Die in Studie 2 gefundenen kleinen bis mittleren Zusammenhänge des Kompetenztests und der Kompetenzselbsteinschätzung sind tendenziell höher als diejenigen Ergebnisse, die in bisherigen Studien gefunden wurden (keine Zusammenhänge bis hin zu kleinen Zusammenhängen; Cassidy & Eachus, 2000; König & Tachtsoglou, 2012; Kunter & Klusmann, 2010; Schladitz et al., 2015). In der Diskussion der zweiten Studie wurde bereits die Vermutung geäußert, dass höhere Zusammenhänge dann vorkommen können, wenn die Studierenden bereits eine forschungsorientierte Lehrveranstaltung besucht haben (vgl. Kapitel 4.4). Die Ermöglichung eigener Forschungserfahrungen scheint eine differenziertere Selbsteinschätzung von Forschungskompetenzen zu ermöglichen (vgl. Bach, 2013; Mertens & Gräsel, 2018) und somit möglicherweise auch höhere Zusammenhängen zur objektiven Erfassung von Forschungskompetenzen.

Die hohen Zusammenhänge zu forschungsbezogenen Selbstwirksamkeitserwartungen könnten ebenfalls auf die Überschätzung von Kompetenzen aufgrund eines geringen Kompetenzniveaus zurückzuführen sein. Bei der Selbsteinschätzung von Forschungskompetenzen wird die *aktuelle* Kompetenz erfasst, bei der Erfassung von Selbsteinschätzungen die Überzeugung, eine *künftige* Handlung durchzuführen (vgl. Kapitel 4.1.3; Wigfield & Eccles, 2000). Da beide Konstrukte über Selbsteinschätzungen erfasst werden, kann die Überschätzung hier gleichermaßen zugrunde liegen. Mit steigender Kompetenz könnte es möglicherweise besser gelingen, die beiden Konstrukte zu unterscheiden.

In Studie 3 könnten die Unterschiede zwischen den Studierenden der Lernforschungsprojekte und denen der Kontrollgruppe auf die *Forschungserfahrung* zurückzuführen sein. Das Durchlaufen des Forschungsprozesses, und damit das Sammeln von *Erfahrungen*, war in den Lernforschungsprojekten ein explizites Seminarziel. Während der Phase *Vorbereitung auf den Forschungsprozess* haben die Studierenden wahrscheinlich deklaratives Wissen zu den einzelnen Phasen des Forschungsprozesses erworben. Während der praktischen Anteile dieser Phase (z.B. dem Formulieren von Forschungsfragen), in der Phase des *Feldzugangs* sowie der *Datenpräsentation* hatten die Studierenden die Möglichkeit, erste Anteile dieses deklarativen Wissens mit dem prozeduralen Wissen in Schemata oder Skripten miteinander in Verbindung zu bringen. Im Gegensatz dazu stand in den Seminaren der Kontrollgruppe die Anwendung forschungsbasierten Wissens für die Lehr-Lern-Praxis und damit die Nutzung von Evidenz im Vordergrund, was zu neuen Verknüpfungen von deklarativem und prozeduralem Wissen in diesen Bereichen geführt haben könnte. Der Einfluss dieser Forschungserfahrung wird vor allem durch die Ergebnisse verdeutlicht, die auf das Auftreten des Response-Shift-Phänomens (Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999) hinweisen. Dieses wiederum lässt sich gut mit der Annahme von Kruger und Dunning (1999) verbinden, dass erst retrospektiv durch das Erlangen von Kompetenz die eigene Überschätzung von Kompetenzen deutlich wird und Kompetenzen genauer eingeschätzt werden können. Denn Response-Shift gibt einen Hinweis darauf, dass sich durch Erfahrungen (a) interne Standards der Messung (Rekalibrierung), (b) die Werte einer Person (Repriorisierung) oder (c) die Bedeutung des Konstruktes verändern können (Rekonzeptualisierung; Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999). Welche Art von Response-Shift aufgetreten ist, kann aufgrund der vorhandenen Datenbasis nicht genau ermittelt werden. Durch die intensive Beschäftigung der Studierenden mit den

Phasen des Forschungsprozesses könnte sich jedoch durchaus die Bedeutung des Konstruktes *Forschungskompetenzen* verändert haben (Rekonzeptualisierung; vgl. Kapitel 5.4). Die Studierenden hätten am Anfang des Semesters, bedingt durch ihre fehlende Erfahrung mit dem Forschungsprozess, Schwierigkeiten, sich selbst differenziert einzuschätzen (Bach, 2013; Mertens & Gräsel, 2018), sodass sie ihre Kompetenzen überschätzen (Kruger & Dunning, 1999). Im Laufe des Semesters hätten sie dann durch die Beschäftigung mit dem Forschungsprozess die Möglichkeit gehabt, Schemata und Skripte zu den Phasen des Forschungsprozesses zu etablieren. Sie hätten damit ein genaueres mentales Bild dieses Prozesses und seiner Anforderungen erlangen können und damit auch die Fähigkeit, ihre eigenen Kompetenzen für diesen Prozess besser einschätzen zu können. Diese Vermutung wird durch die Ergebnisse der Kontrollgruppe bekräftigt, bei deren Studierenden ein Response-Shift nur in denjenigen Bereichen aufgetreten ist, die auf die Nutzung von Evidenz abzielten²².

Nachdem die Ergebnisse der drei Studien dieser Arbeit zusammengefasst und integrierend diskutiert wurden, bleiben noch einige Fragen offen. Diese sollen im nachfolgenden Kapitel dargestellt werden.

6.2 Grenzen und Stärken der Arbeit

In den einzelnen Studien wurde bereits auf einige Einschränkungen der Ergebnisse verwiesen. Nachfolgend soll auf zentrale, studienübergreifende Einschränkungen inhaltlicher und methodischer Art eingegangen werden, um anschließend die Stärken der Arbeit darzustellen.

Grenzen

Forschungskompetenzen als generische Kompetenzen. Dem RMRK-W-Modell (Thiel & Böttcher, 2014), welches die Grundlage für die Entwicklung des F-Komp, aber auch für die Gestaltung der forschungsorientierten Lernumgebung war, liegt die Annahme der pragmatischen Wissenschaftstheorie (Butts, 1991; Dewey, 1938; Mittelstraß, 1991) zugrunde, dass sich fach- bzw. disziplinübergreifend gemeinsame Aspekte des Forschungsprozesses identifizieren lassen sollten. Im Sinne der Systematisierung von Hetmanek und Kolleginnen (2018) sollten Forschungskompetenzen ein *domänenübergreifendes* Konstrukt darstellen (vgl.

²² Zwar ist auch im Bereich der Methodenkompetenzen ein Response-Shift aufgetreten. Allerdings wurde hierfür in Kapitel 5.4 eine mögliche Alternativerklärung identifiziert: der Einfluss des Wissenserwerbs in einer Parallelveranstaltung.

Abbildung 2.2). Diese Annahme konnte in dieser Arbeit nicht abschließend beantwortet werden. Die Überprüfung der fach- bzw. disziplinübergreifenden Annahme steht aufgrund zu geringer Stichprobengrößen noch aus. Darüber hinaus sind Studierende aus geisteswissenschaftlichen Studiengängen bislang noch unterrepräsentiert.

Einflussfaktoren auf Kompetenzselbsteinschätzungen. Im vorherigen Kapitel wurde ausgeführt, inwiefern verschiedene Faktoren einen Einfluss auf die Selbsteinschätzung von Kompetenzen haben könnten. Diese Einflüsse wurden bislang jedoch noch nicht ausreichend empirisch abgesichert. So wurden die genannten Verzerrungstendenzen oder Wertkomponenten bislang noch nicht über entsprechende Skalen erfasst oder mit geeigneten Analysen überprüft. Darüber hinaus konnten die Annahmen über die Kompetenzniveaus der Studierenden bislang noch nicht untersucht werden, da neben der nötigen Stichprobengröße beispielsweise Angaben zu bisherigen Forschungserfahrungen fehlten. Zwar konnte die in Studie 1 identifizierte Faktorenstruktur in Studie 2 für Masterstudierende repliziert werden, jedoch konnten bislang noch keine Analysen für Bachelor- und Promotionsstudierende durchgeführt werden.

Forschungskompetenzen als kognitive Outcomes. Laut der Definition von Kompetenzen, die dieser Arbeit zugrunde liegt (vgl. Kapitel 2.1.1 und Kapitel 3.1), sollte es sich bei Forschungskompetenzen um *kognitive Leistungsdispositionen* (Klieme et al., 2008) handeln. Allerdings deuten die hohen Zusammenhänge mit forschungsbezogenen Selbstwirksamkeitserwartungen an, dass es sich bei Forschungskompetenzen, die mit dem F-Komp via Selbsteinschätzungen erfasst werden, weniger um kognitive als vielmehr um *motivationale und/oder volitionale Outcomes* (vgl. Kapitel 2.1.1), und zwar *Kompetenzerwartungen* handeln könnte (Cramer, 2010; Mertens & Gräsel, 2018). Dazu sind weitere Analysen notwendig.

Schritte der Validitätsüberprüfung. Bislang kann nur eine vorläufige Aussage über die Validität des F-Komp getroffen werden. Einerseits steht die Überprüfung der Kriteriumsvalidität durch die Hinzunahme von externen Kriterien bislang noch aus. Andererseits konnte die Validität der Methodik der Selbsteinschätzung noch nicht final beantwortet werden. Darüber hinaus wurden im Rahmen der Konstruktvalidierung bislang vor allem Konstrukte verwendet, die den gesamten Forschungsprozess thematisieren oder allgemeine Lern- und Studienkompetenzen darstellen. Bisher nicht berücksichtigt wurde die Hinzunahme von Teilkompetenzen des Forschungsprozesses (Kapitel 2.2.1).

Stichproben. Im Hinblick auf die Stichproben, die den drei Studien zugrunde liegen, sind folgende Einschränkungen zu beachten: Es handelt sich überwiegend um Gelegenheitsstichproben, da Lehrende akquiriert wurden, die freiwillig Erhebungszeit in ihren Lehrveranstaltungen zur Verfügung gestellt haben. Es könnte ein *participation bias* (Döring & Bortz, 2016) vorliegen. Darüber hinaus mussten einige Erhebungen deutlich verkürzt vorgenommen werden, um möglichst wenig Zeit einer Lehrveranstaltung in Anspruch zu nehmen. So war insbesondere in Studie 2 die Bearbeitung der Forschungsfragen nur in Teilstudien möglich. Weiterhin muss beachtet werden, dass die Ergebnisse in Studie 3, obwohl diese eine gezielt ausgewählte Stichprobe beinhaltetete, ebenfalls einigen Einschränkungen unterliegen: Die Stichprobe der Studie 3 unterlag einem hohen Dropout, welcher die Aussagekraft der Ergebnisse durch die geringe Stichprobengröße eingeschränkt haben könnte (Döring & Bortz, 2016). Darüber hinaus sind die Ergebnisse nicht auf andere Universitäten übertragbar, da diese Universitäten zum Teil andere Schwerpunkte für das Praxissemester im Lehramtsmaster legen (z.B. König & Rothland, 2018; Mertens & Gräsel, 2018). Daher muss festgehalten werden, dass für diese Arbeit mit ihren drei Studien aufgrund der Stichprobenproblematik eine eingeschränkte externe Validität vorliegt und die Ergebnisse damit nur bedingt generalisierbar sind. Darüber hinaus war es aufgrund der teils zu kleinen Stichproben (z.B. hinsichtlich der Stufen des Studiums, einzelner Fächer oder Seminargruppen) nicht möglich, komplexere Analysen vorzunehmen.

Stärken

Die Dissertation erweitert mit ihren drei Studien den Forschungsstand in folgenden Punkten: Zum einen wird ein Verfahren zur Erfassung von Forschungskompetenzen vorgestellt, welches theoretisch fundiert ist und sich empirisch entsprechend des RMRK-W-Modells (Thiel & Böttcher, 2014) abbilden lässt: der Fragebogen zur Erfassung studentischer Forschungskompetenzen, welcher eine fach- und disziplinübergreifende Erfassung ermöglicht. Dieses Verfahren basiert auf der pragmatischen Wissenschaftstheorie (Butts, 1991; Dewey, 1938; Mittelstraß, 1991) und damit auf einer gut begründeten Theorie über gemeinsame, fach- und disziplinübergreifende Phasen des Forschungsprozesses. Zudem lassen sich bereits einige Hinweise auf die fach- bzw. disziplinübergreifende Modellierung von Forschungskompetenzen finden durch (1) den Nachweis der Inhaltsvalidität (vgl. Thiel & Böttcher, 2014 und Kapitel 3.1); (2) die faktorielle Validierung in Studie 1, in welcher die Faktorenstruktur des F-Komp überprüft wurde, und zwar anhand einer Stichprobe von

Studierenden aus unterschiedlichen Disziplinen; (3) die Replikation dieses Modells in Studie 2 an einer Stichprobe von Masterstudierenden aus unterschiedlichen Disziplinen.

Darüber hinaus wurde auf der Grundlage des RMRK-W-Modells (Thiel & Böttcher, 2014) eine forschungsorientierte Lernumgebung für angehende Lehrkräfte konzipiert. Durch diese Modellbasierung ließen sich einerseits konkrete Kompetenzziele definieren, welche auf die Standards für die Lehrerbildung in den Bildungswissenschaften (Kultusministerkonferenz, 2004) abgestimmt sind, und andererseits entsprechende Lehrpläne gestalten. Das Erreichen der intendierten Ziele wurde in dieser Studie umfassend evaluiert, und zwar mit einer Kombination aus einem Kompetenztest und Kompetenzselbsteinschätzungen unter Berücksichtigung des Response-Shift-Phänomens (Schwartz & Sprangers, 1999, 2010; Sprangers & Schwartz, 1999).

6.3 Ausblick

Zum Abschluss dieser Arbeit sollen mögliche Implikationen für Forschung und Praxis abgeleitet werden.

Implikationen für künftige Forschung

Die vorliegende Arbeit hat gezeigt, dass der F-Komp ein großes Potenzial für eine valide Erfassung studentischer Forschungskompetenzen mit sich bringt. Daher wären weitere Untersuchungen gewinnbringend, die sich aus den zuvor dargestellten Grenzen dieser Arbeit ableiten lassen sowie weiterführende Fragen adressieren.

Größere, systematisch zusammengesetzte Stichproben. Einige der zuvor dargestellten Grenzen dieser Arbeit lassen sich durch eine systematische Stichprobenziehung und größere, repräsentative (Sub-)Stichproben überwinden. So kann der Nachweis von *Forschungskompetenzen als generische Kompetenzen* erbracht werden, indem fach- bzw. disziplinspezifische Substichproben mit ausreichend Studierenden aus verschiedenen Fächern und Disziplinen akquiriert werden, um latente Analysen, wie die Überprüfung der Faktorenstruktur und Invarianzmessungen, vornehmen zu können. Sodann wäre auch die Mehrebenenstruktur der Daten besser zu berücksichtigen. Darüber hinaus kann der *Einfluss des Kompetenzniveaus auf Kompetenzselbsteinschätzungen* anhand von größeren Stichproben mit Studierenden aus unterschiedlichen Stufen des Studiums überprüft werden. Möglicherweise lassen sich sogar Kompetenzstufen entsprechend der Kompetenzniveaus modellieren.

Weiterhin könnten *Forschungskompetenzen als kognitive Outcomes* ausführlicher analysiert werden, indem sowohl Forschungskompetenzen über Selbsteinschätzungen und Kompetenztests erfasst werden als auch Selbstwirksamkeitserwartungen mitberücksichtigt werden. Diese könnten in einem längsschnittlichen, (quasi-)experimentellen Design erhoben werden. Sofern die Stichprobe ausreichend groß wäre, könnten latente Veränderungsmessungen vorgenommen werden, um die Genauigkeit der Kompetenzselbsteinschätzung, die Messinvarianz über die Zeit zur Identifikation von Response-Shift sowie Aussagen über die genaue Beziehung der Konstrukte untereinander zu überprüfen.

Darüber hinaus könnten weitere Skalen und Analysemethoden eingesetzt werden, um den *Einfluss* der vormals genannten *Faktoren auf Kompetenzselbsteinschätzungen* zu untersuchen. So könnten beispielsweise die Skalen von Thoren und KollegInnen (in Druck) zur Erfassung der Wertkomponenten in künftige Untersuchungen mit aufgenommen werden, oder die Vorerfahrung mit Forschung erfasst werden, um die Analyse von Kompetenzniveaus zu unterstützen. Faktorenanalysen könnten genutzt werden, um Einflüsse wie soziale Erwünschtheit zu untersuchen (Lucas & Baird, 2006).

In künftigen Studien sollten weitere *Schritte der Validitätsüberprüfung* des F-Komp vorgenommen werden. So könnten für eine weitere Konstruktvalidierung Zusammenhänge zu weiteren Teilkompetenzen des Forschungsprozesses (vgl. Kapitel 2.2.1) erfasst werden. Zu diesen Teilkompetenzen stehen zum Teil bereits validierte Instrumente zur Verfügung, z.B. ein Kompetenztest zur Erfassung von *Information Literacy* (Trempler et al., 2015) oder (offene und geschlossene) Testaufgaben zur Erfassung des wissenschaftlichen Schreibens (Winter-Hözl, Wäschle, Wittwer, Watermann & Nückles, 2015). Zudem könnte die Kriteriumsvalidität überprüft werden, indem (a) bisherige Forschungserfahrungen oder Noten von Bachelorarbeiten zur Überprüfung der retrospektiven Validität, (b) die Analyse eines Forschungsberichts, welcher zur forschungsorientierten Lehrveranstaltung verfasst werden muss, zur Überprüfung der konkurrenten Validität, oder (c) die Note des Masterabschlusses zur Überprüfung der prognostischen Validität in die oben skizzierten Forschungsvorhaben hinzugenommen wird.

Der auf dem RMRK-W-Modell (Thiel & Böttcher, 2014) basierende Ansatz umfasst bisher lediglich die Selbsteinschätzung von Kompetenzen. Dieser Ansatz könnte weiter ausgebaut werden: Einerseits könnte, wie in Studie 2 vorgeschlagen, ein entsprechendes

Instrument zur Erfassung von Selbstwirksamkeitserwartungen entwickelt werden. Andererseits könnte ein Kompetenztest entwickelt werden, der sich an Aufgaben aus bisherigen fach- bzw. disziplinspezifischen Tests orientiert, z.B. zu Information Literacy (Trempler et al., 2015), bildungswissenschaftliche (Groß Ophoff et al., 2014, 2017) oder sozialwissenschaftliche Forschungskompetenz (Gess, Wessels & Blömeke, 2017; Gess, Geiger & Ziegler, 2018). Eine große Herausforderung bestünde sicherlich in der Problematik der *domänenübergreifenden* Konstruktion von Aufgaben. Diese Aufgaben müssten so *domänenübergreifend*, gleichzeitig aber auch so *domänenspezifisch* wie möglich gestaltet werden, um eine differenzierte Kompetenzeinschätzung vornehmen zu können. Möglicherweise ließen sich verschiedene Testversionen entwickeln, welche dieselben Aufgabentypen und *-schwierigkeiten* zu bestimmten Kompetenzbereichen umfassen, inhaltlich allerdings fachspezifisch ausgestaltet werden. Darüber hinaus ließen sich noch Versionen des F-Komp zur Fremdbeurteilung entwickeln – entweder für externe Beobachter oder für Lehrende, die den Kompetenzerwerb von Studierenden einschätzen.

Im Bereich der Kompetenzen für eine evidenzbasierte Praxis könnte der Impact von forschungsorientierten Lernformanten auf SchülerInnen-Outcomes überprüft werden. Es könnte längsschnittlich überprüft werden, inwiefern von denjenigen Lehramtsstudierenden, die eine entsprechende Intervention besucht haben, künftig eine evidenzbasierte Praxis umgesetzt wird und inwiefern SchülerInnen davon profitieren. Dazu könnten die Lehramtsstudierenden im Referendariat erneut in einer Follow-Up-Erhebung befragt werden und die Erfassung der SchülerInnen-Leistungen in ein entsprechendes Studiendesign mit aufgenommen werden.

Implikationen für die Praxis

Die Förderung von Kompetenzen für eine evidenzbasierte Praxis wird für (angehende) Lehrkräfte immer relevanter (J. Bauer & Prenzel, 2012; J. Bauer, Prenzel & Renkl, 2015; Borg, 2010; Davies, 1999; Hargreaves, 1996, 1997). Diese Kompetenzen der Nutzung und des Generierens von Evidenz können angehenden Lehrkräften dabei helfen, in ihrer späteren beruflichen Praxis evidenzbasierte Entscheidungen zu treffen und die Qualität ihrer Lehre zu verbessern (vgl. Kippers et al., 2018), um SchülerInnen bestmöglich zu fördern. Die Bereitschaft für diese evidenzbasierte Praxis könnte gestärkt werden, indem es bereits während der Lehramtsausbildung gelingt, dass Lehramtsstudierende den persönlichen Mehrwert dieser Praxis für die Qualität des eigenen Unterrichts erkennen und verinnerlichen.

Mit dem Fragebogen zur Erfassung studentischer Forschungskompetenzen steht einerseits ein Instrument zur Evaluation forschungsorientierter Lernformate zur Verfügung. Andererseits können aber auch universitätsweite FoL-Programme, wie sie beispielsweise im Rahmen der Exzellenzinitiative entstanden sind, mit dem Fragebogen evaluiert werden. Durch die fach- bzw. disziplinübergreifende Erfassung sind universitätsweite Handlungsempfehlungen ableitbar (Böttcher & Thiel, 2017) oder Vergleiche zwischen Universitäten möglich. Der Fragebogen stößt daher bereits auf große Nachfrage. So existieren mittlerweile Übersetzungen oder Adaptionen des Instrumentes (z.B. eine türkische Version von Alkan, 2018; eine Version für technisch orientierte Fachhochschulen von Hauser, Reuter, Gruber & Mottok, 2018) oder befinden sich in Planung (z.B. Eberhard Karls Universität Tübingen; Georg-August-Universität Göttingen; University of Bialystok; Universität Bremen; Universitätsklinikum Heidelberg; Universität von Valencia). Darüber hinaus zeigt die insgesamt große Bandbreite an bisher existierenden Ansätzen zur Modellierung (Kapitel 2.2.1) sowie Beispiele zur Förderung von Forschungskompetenzen (Kapitel 2.3.1), dass ein großes Interesse daran besteht, die Verknüpfung von Forschung und Lehre zur Stärkung des Research-Teaching-Nexus (Brew, 2006; Griffiths, 2004; Healey, 2005; Healey & Jenkins, 2009) durch Programme wie Forschungsorientierte Lehre weiter voranzutreiben.

Literatur II

- Alkan, M. F. (2018). R-COMP Araştırma Yeterlikleri Ölçeğinin Türk Dili ve Kültürüne Uyarlanması [Anpassung der Forschungsskala an die türkische Sprache und Kultur]. *International Journal of Turkish Education Sciences*, 6(11), 15–28.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369–406.
- Bach, A. (2013). *Kompetenzentwicklung im Schulpraktikum. Ausmaß und zeitliche Stabilität von Lerneffekten hochschulischer Praxisphasen* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 87, 1. Aufl.). Münster: Waxmann.
- Bauer, J. & Prenzel, M. (2012). Science education. European teacher training reforms. *Science*, 336(6089), 1642–1643.
- Bauer, J., Prenzel, M. & Renkl, A. (2015). Evidenzbasierte Praxis - im Lehrerberuf?! Einführung in den Thementeil. *Unterrichtswissenschaft*, 43(3), 188–192.
- Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429.

- Böttcher, F. & Thiel, F. (2017). *Ergebnisse der Evaluation der Forschungsorientierten Lehre (FoL) an der Freien Universität Berlin*. Berlin: Freie Universität Berlin.
- Braun, E. (2008). *Das Berliner Evaluationsinstrument für selbsteingeschätzte studentische Kompetenzen (BEvaKomp)*. Göttingen: V & R Unipress.
- Brew, A. (2006). *Research and teaching. Beyond the divide (Universities into the 21st century)*. Houndmills, Großbritannien: Palgrave Macmillan.
- Butts, R. E. (1991). Methodology, metaphysics and the pragmatic unity of science. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 23–28). Berlin: W. de Gruyter.
- Cassidy, S. & Eachus, P. (2000). Learning style, academic belief systems, self-report student proficiency and academic achievement in higher education. *Educational Psychology*, 20(3), 307–322.
- Chevalier, A., Gibbons, S., Thorpe, A., Snell, M. & Hoskins, S. (2009). Students' academic self-perception. *Economics of Education Review*, 28(6), 716–727.
- Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship. Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Hrsg.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (S. 453–494). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cramer, C. (2010). Kompetenzerwartungen Lehramtsstudierender. Grenzen und Perspektiven selbsteingeschätzter Kompetenzen in der Lehrerbildungsforschung. In A. Gehrmann, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle. Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 85–97). Bad Heilbrunn: Klinkhardt Verlag.
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108–121.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238.
- Deutsche Forschungsgemeinschaft & Wissenschaftsrat (2015). *Bericht der Gemeinsamen Kommission zur Exzellenzinitiative an die Gemeinsame Wissenschaftskonferenz*. Bonn. Zugriff am 03.06.2019. Verfügbar unter <http://www.gwk-bonn.de/fileadmin/Papers/DFG-WR-Bericht-Juni2015.pdf>

- Dewey, J. (1938). *Unity of science as a social problem*. Chicago, IL: University of Chicago Press.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (Springer-Lehrbuch, 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Berlin: Springer.
- Erpenbeck, J. & von Rosenstiel, L. (2007). Einführung. In J. Erpenbeck & L. von Rosenstiel (Hrsg.), *Handbuch Kompetenzmessung. Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (2., überarbeitete und erweiterte Auflage, S. XVII–XLV). Stuttgart: Schäffer-Poeschel.
- Frey, A. (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften – eine erste Standortbestimmung zu bereits publizierten Instrumenten. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (Zeitschrift für Pädagogik Beiheft, Bd. 51, S. 30–46). Weinheim: Beltz.
- Gess, C., Geiger, C. & Ziegler, M. (2018). Social-scientific research competency. Validation of Test Score Interpretations for Evaluative Purposes in Higher Education. *European Journal of Psychological Assessment*, 1–14.
- Gess, C., Wessels, I. & Blömeke, S. (2017). Domain-specificity of research competencies in the social sciences. Evidence from differential item functioning. *Journal for Educational Research Online*, 9(2), 11–36.
- Griffiths, R. (2004). Knowledge production and the research–teaching nexus: The case of the built environment disciplines. *Studies in Higher Education*, 29(6), 709–726.
- Groß Ophoff, J., Schladitz, S., Lohrmann, K. & Wirtz, M. (2014). Evidenzorientierung in bildungswissenschaftlichen Studiengängen. In K. Drossel, R. Strietholt & W. Bos (Hrsg.), *Empirische Bildungsforschung und evidenzbasierte Reformen im Bildungswesen* (S. 251–275). Münster: Waxmann.
- Groß Ophoff, J., Wolf, R., Schladitz, S. & Wirtz, M. (2017). Assessment of educational research literacy in higher education: Construct validation of the factorial structure of an assessment instrument comparing different treatments of omitted responses. *Journal for Educational Research Online/Journal für Bildungsforschung Online*, 9(2), 37–68.
- Hargreaves, D. H. (1996). *Teaching as a research-based profession: Possibilities and prospects*. Teacher Training Agency Annual Lecture, London.

- Hargreaves, D. H. (1997). In defence of research for evidence-based teaching: a rejoinder to Martyn Hammersley. *British Educational Research Journal*, 23(4), 405–419.
- Hauser, F., Reuter, R., Gruber, H. & Mottok, J. (2018). Research competence: Modification of a questionnaire to measure research competence at universities of applied sciences. In *Proceedings of 2018 IEEE Global Engineering Education Conference (EDUCON)*. Date and venue: 17-20 April, 2018, Santa Cruz de Tenerife, Canary Islands, Spain (S. 109–117). Piscataway, NJ: IEEE.
- Healey, M. (2005). Linking research and teaching to benefit student learning. *Journal of Geography in Higher Education*, 29(2), 183–201.
- Healey, M. & Jenkins, A. (2009). *Developing undergraduate research and inquiry*. Heslington, Großbritannien: Higher Education Academy.
- Hetmanek, A., Engelmann, K. F., Opitz, A. & Fischer, F. (2018). Beyond intelligence and domain knowledge. Scientific reasoning and argumentation as a set of cross-domain skills. In F. Fischer, C. A. Chinn, K. F. Engelmann & J. Osborne (Hrsg.), *Scientific reasoning and argumentation. The roles of domain-specific and domain-general knowledge* (S. 203-226). New York, NY: Routledge.
- Hetmanek, A., Wecker, C., Kiesewetter, J., Trempler, K., Fischer, M. R., Gräsel, C. et al. (2015). Wozu nutzen Lehrkräfte welche Ressourcen? Eine Interviewstudie zur Schnittstelle zwischen bildungswissenschaftlicher Forschung und professionellem Handeln im Bildungsbereich. *Unterrichtswissenschaft*, 43(3), 193–208.
- Hill, L. G. & Betz, D. L. (2005). Revisiting the retrospective pretest. *American Journal of Evaluation*, 26(4), 501–517.
- Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140.
- Kippers, W. B., Poortman, C. L., Schildkamp, K. & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention? *Studies in Educational Evaluation*, 56, 21–31.
- Klieme, E. & Hartig, J. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (Zeitschrift für Erziehungswissenschaft Sonderheft 8 | 2007, S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Klieme, E., Hartig, J. & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 3–22). Göttingen: Hogrefe & Huber Publishers.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876–903.
- König, J. & Rothland, M. (2018). Das Praxissemester in der Lehrerbildung: Stand der Forschung und zentrale Ergebnisse des Projekts Learning to Practice. In J. König, M. Rothland & N. Schaper (Hrsg.), *Learning to Practice, Learning to Reflect?* (S. 1–62). Wiesbaden: Springer Fachmedien Wiesbaden.
- König, J. & Tachtsoglou, S. (2012). Pädagogisches Professionswissen und selbsteingeschätzte Kompetenz. In J. König & A. Seifert (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung* (S. 284–297). Münster: Waxmann.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.
- Kultusministerkonferenz (2004). *Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004.*
- Kunter, M. & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften–Methodische Herausforderungen. *Unterrichtswissenschaft*, 38(1), 68–86.
- Lucas, R. E. & Baird, B. M. (2006). Global self-assessment. In M. Eid & E. Diener (Hrsg.), *Handbook of multimethod measurement in psychology* (S. 29–42). Washington, DC: American Psychological Association.
- Mayer, R. E. (2003). What causes individual differences in cognitive performance? In R. J. Sternberg & E. Grigorenko (Hrsg.), *Perspectives on the psychology of abilities, competencies, and expertise* (S. 263–273). New York, NY: Cambridge University Press.
- Mertens, S. & Gräsel, C. (2018). Entwicklungsbereiche bildungswissenschaftlicher Kompetenzen von Lehramtsstudierenden im Praxissemester. *Zeitschrift für Erziehungswissenschaft*, 21(6), 1109–1133.

- Mittelstraß, J. (1991). Einheit und Transdisziplinarität. In Akademie der Wissenschaften zu Berlin (Hrsg.), *Einheit der Wissenschaften. Internationales Kolloquium der Akademie der Wissenschaften zu Berlin (Bonn, 25.-27. Juni 1990)* (Forschungsbericht, Bd. 4, S. 12–22). Berlin: W. de Gruyter.
- Nückles, M. & Wittwer, J. (2014). Lernen und Wissenserwerb. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie. Mit Online-Materialien* (Psychologie 2014, 6., vollständig bearb. Aufl., S. 225–252). Weinheim: Julius Beltz.
- Nusche, D. (2008). *Assessment of learning outcomes in higher education. A comparative review of selected practices* (OECD Education Working Papers, Bd. 15). Frankreich: OECD publishing.
- Renkl, A. (2015). Wissenserwerb. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (Springer-Lehrbuch, S. 3–24). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schladitz, S., Groß Ophoff, J. & Wirtz, M. (2015). Konstruktvalidierung eines Tests zur Messung bildungswissenschaftlicher Forschungskompetenz. *Zeitschrift für Pädagogik*, 61, 167–184.
- Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology*, 68(4), 551–562.
- Schraw, G. (2006). Knowledge: Structures and processes. In P. A. Alexander & P. H. Winne (Hrsg.), *Handbook of Educational Psychology* (2. Aufl., S. 245–260). Mahwah, N.J.: Erlbaum.
- Schwartz, C. E. & Sprangers, M. A. G. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11), 1531–1548.
- Schwartz, C. E. & Sprangers, M. A. G. (2010). Guidelines for improving the stringency of response shift research using the thentest. *Quality of Life Research*, 19(4), 455–464.
- Simonton, D. K. (2003). Expertise, competence, and creative ability. In R. J. Sternberg & E. Grigorenko (Hrsg.), *Perspectives on the psychology of abilities, competencies, and expertise* (S. 213–240). New York, NY: Cambridge University Press.
- Sprangers, M. A. G. & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research. A theoretical model. *Social Science & Medicine*, 48(11), 1507–1515.

- Thiel, F. & Böttcher, F. (2014). Modellierung fächerübergreifender Forschungskompetenzen - Das RMKR-W-Modell als Grundlage der Planung und Evaluation von Formaten forschungsorientierter Lehre. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre* (I 2.10, S. 109–124). Berlin: Raabe.
- Thoren, K., Wißmann, J., Harks, M., Wenger, M., Kinder, A. & Hannover, B. (in Druck). Förderung von Datennutzungskompetenzen in der Lehramtsausbildung. Konzeption und Evaluation dreier Seminare. In I. Gogolin, B. Hannover & A. Scheunpflug (Hrsg.), *Evidenzbasierung in der Lehramtsausbildung* (Edition der Zeitschrift für Erziehungswissenschaft).
- Trautwein, U. & Lüdtke, O. (2007). Epistemological beliefs, school achievement, and college major: A large-scale longitudinal study on the impact of certainty beliefs. *Contemporary Educational Psychology*, 32(3), 348–366.
- Trempler, K., Hetmanek, A., Wecker, C., Kiesewetter, J., Wermelt, M., Fischer, F. et al. (2015). Nutzung von Evidenz im Bildungsbereich. Validierung eines Instruments zur Erfassung von Kompetenzen der Informationsauswahl und Bewertung von Studien. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik, Beiheft, Bd. 61, S. 144–166). Weinheim: Beltz Juventa.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Seattle, WA: Hogrefe & Huber.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.
- Winter-Hölzl, A., Wäschle, K., Wittwer, J., Watermann, R. & Nückles, M. (2015). Entwicklung und Validierung eines Tests zur Erfassung des Genrewissens Studierender und Promovierender der Bildungswissenschaften. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik, Beiheft, Bd. 61, S. 185–202). Weinheim: Beltz Juventa.
- Zlatkin-Troitschanskaia, O., Pant, H. A., Kuhn, C., Toepper, M. & Lautenbach, C. (2016). *Messung akademisch vermittelter Kompetenzen von Studierenden und Hochschulabsolventen. Ein Überblick zum nationalen und internationalen Forschungsstand* (Edition ZfE, Band 1). Wiesbaden: Springer VS.

Abbildungs- und Tabellenverzeichnis

Abbildung 2.1 Übersetzung aus dem Englischen und Adaption der Grafik von Hetmanek et al. (2018, S. 210) zur Hierarchie von kognitiven Fertigkeiten	22
Abbildung 2.2 Domänenübergreifende Forschungskompetenzen als Spezialfall des wissenschaftlichen Denkens und Argumentierens (Adaption der übersetzten Abbildung von Hetmanek et al., 2018, S. 210; vgl. Abbildung 2.1)	25
Abbildung 2.3 Das RMRK-W-Modell	29
Figure 3.1 The RMRC-K-model	63
Table 3.1 Overview of sub-samples	65
Figure 3.2 Schematic illustration of the five-factor model and the five-factor model with sub-dimensions	68
Table 3.2 Model fit for the tested models	68
Table 3.3 Latent inter-factor correlations	69
Figure 3.3 Schematic illustration of the higher-order five-factor model with sub-dimensions	69
Figure 3.4 Schematic illustration of the bi-factor model with sub-dimensions and the bi-factor-model	71
Table 3.4 Factor loadings, means, standard deviations, values for construct reliability, and values for Cronbach's α of the dimensions and sub-dimensions	72
Abbildung 4.1 Schematische Darstellung des empirischen Faktorenmodells. Die Buchstabenbezeichnungen der Subfaktoren entspricht der Darstellung im Text; Ki = Item zu Kompetenzbereichen; Wi = Item zu Wissensbereichen	92
Tabelle 4.1 Übersicht zur Passung von Fragestellungen und den drei durchgeführten Studien	99
Tabelle 4.2 Stichprobenübersicht der Studien 1-3	99
Tabelle 4.3 Übersicht zu genutzten Skalen mit Antwortformat, Cronbachs α und Beispielimens	102
Tabelle 4.4 Latente und manifeste Korrelationen zwischen selbsteingeschätzten Forschungskompetenzen und den Validierungskonstrukten	107
Tabelle 4.5 Mittelwerte (Standardabweichungen) / Cronbachs α in den Studien 1-3	109
Table 5.1 Characteristics of participants in the longitudinal sample	136
Table 5.2 Multivariate effects for self-assessed and tested research competences	140
Table 5.3 Univariate effects for self-assessed and tested research competences	141
Table 5.4 Mean values, standard deviations, differences in mean values, effect sizes, and Cronbach's alpha for students in RLP	144
Table 5.5 Mean values, standard deviations, differences in mean values, effect sizes, and Cronbach's alpha for students in CG	145



Danksagung

Die vorliegende Arbeit entstand im Rahmen des Projektes „Forschungsorientierte Lehre an der Freien Universität Berlin“, welches durch das Präsidium der Freien Universität Berlin initiiert und gefördert wurde. Ein besonderer Dank gilt Felicitas Thiel, unter deren Leitung ich dieses Projekt durchgeführt habe und die mich während meiner Promotion betreut hat. Felicitas Thiel hat mich in all der Zeit enorm unterstützt: durch das Vertrauen, welches sie mir entgegengebracht hat, durch ihr wertvolles Feedback, durch ihre wissenschaftliche, aber auch persönliche Unterstützung, aber vor allem, weil sie mir ein Arbeiten ermöglicht hat, dass meiner persönlichen Situation entsprochen hat. Das alles weiß ich sehr zu schätzen, danke!

Darüber hinaus gilt mein Dank Rainer Watermann, der sich sowohl die Zeit für das Zweitgutachten dieser Arbeit nimmt als auch in Gesprächen immer wieder hilfreiche Hinweise für mich hatte, die in diese Arbeit eingeflossen sind.

Ein ausdrücklicher Dank gilt meinen (ehemaligen) KollegInnen am Arbeitsbereich Schulpädagogik/Schulentwicklungsforschung, die mich während meiner Promotionszeit begleitet haben, entweder in gemeinsamen Gesprächen, durch aufbauende Worte oder durch euer wertvolles Feedback zu meinen Beiträgen im Kolloquium. Ich nenne euch einfach in alphabetischer Reihenfolge: Anabel Bach, Victoria Barth, Anja Böhnke, Susanne Czaja, Sebastian Haase, Christopher Kellermann, Katharina Krysmanski, Anne-Katrin Lehmann, Juliane Müller, Max Nachbauer, Franziska Pfitzner-Eden, Jasmin Tarkian, Danguì Yue. Weiterhin möchte ich an dieser Stelle Jana Groß Ophoff danken, die mich in vielen Telefonaten und einigen persönlichen Treffen beim Erstellen zweier Manuskripte unterstützt hat.

An dieser Stelle möchte ich nicht vergessen, mich bei denjenigen Personen zu bedanken, die mich im Rahmen des FoL-Projekt als studentische MitarbeiterInnen, PraktikantInnen oder WerkvertragsnehmerInnen unterstützt haben bzw. im Rahmen ihrer Tätigkeit am Arbeitsbereich Schulpädagogik/Schulentwicklungsforschung: Chelsea Löffler, Jessica Meier, Paulina Lanäus, Victoria Kruzewitz, Clara Limberg und Moritz Nottrott sowie Marius Eckert, Elizaveta Firsova, Anne Ntinda, Julia Bulla, Patrick Sydow, Berkan Sarigül, Antonia Jonas. Danke!

Ich möchte mich sehr bei meiner Familie und meinen Freunden bedanken, die mich auf diesem Weg begleitet haben. Dazu gehören insbesondere meine Eltern, die mich

stets in allem voller Liebe unterstützt haben und die Grundlage für meinen Weg gelegt haben; mein Bruder Stefan und seine Frau Katrin, die mir stets eine Inspiration waren; meine Freundinnen Anna und Dani, die mich seit dem Abitur jede Woche aufs Neue mit ihrer Freundschaft bereichern; meine Schwägerin Wiebke und mein Schwager Sören, mit denen mein Mann und ich ein vierblättriges Kleeblatt bilden; Martin Latsch (†) und André Nowakowski, die mich insbesondere in einer frühen Phasen meiner Promotion unterstützt haben.

Mein letzter Dank und dennoch der größte geht an meinen Ehemann Fabian. Du hast mich vom ersten Moment dieser Promotion an bedingungslos unterstützt, hast mich aufgefangen und mich aufgebaut. Du teilst einfach alle Momente mit mir. *Für immer und dich*²³.

²³ Quellenangabe des Liedzitats: „Für immer und dich“, written by Rio Reiser, 1986.

Selbstständigkeitserklärung

Hiermit versichere ich, die Dissertationsarbeit „Erfassung und Förderung studentischer Forschungskompetenzen - Validierung eines Instrumentes zur Erfassung studentischer Forschungskompetenzen sowie Gestaltung und Evaluation einer forschungsorientierten Lernumgebung“ selbstständig angefertigt zu haben. Sämtliche Hilfsmittel, die ich verwendet habe, sind angegeben. Sie ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Franziska Böttcher-Oschmann | Berlin, im Juni 2019



Anteile an Studien mit Co-Autorenschaft

Die Studien, die für diese Dissertationsarbeit durchgeführt wurden, sind in Co-Autorenschaft entstanden. Nachfolgend soll ein Überblick über die Fremdanteile an den Artikeln gegeben werden.

Studie 1: Evaluating research-oriented teaching: a new instrument to assess university students' research competences

Die Studie ist in Co-Autorenschaft mit Felicitas Thiel entstanden. Die Konzeption, Durchführung und Berichtsabfassung lag bei Franziska Böttcher-Oschmann. In den Überarbeitungsprozess des Artikels sind, neben Anmerkungen von anonymen Gutachterinnen und Gutachtern, Hinweise von Felicitas Thiel eingeflossen.

Studie 2: Validierung eines Fragebogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen - Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements & Studie 3: How do we prepare teacher training students for Evidence-based Practice? Promoting students' research competences in Research-learning projects

Die zweite und dritte Studie sind in Co-Autorenschaft mit Jana Groß Ophoff und Felicitas Thiel entstanden. Für diese beiden Studien wurden von Jana Groß Ophoff einerseits der verwendete Kompetenztest, andererseits statistische Kennwerte wie WLE-Schätzer und EAP/PV-Reliabilität zur Verfügung gestellt. Die in den Studien dargestellten Analysen mit diesen Kennwerten erfolgten durch Franziska Böttcher-Oschmann. Die Konzeption, Durchführung und Berichtsabfassung lag bei Franziska Böttcher-Oschmann. In den Überarbeitungsprozess des Artikels sind, neben Anmerkungen von anonymen Gutachterinnen und Gutachtern für die zweite Studie, Hinweise von Jana Groß Ophoff und Felicitas Thiel eingeflossen.