

Immanuel Ulrich & Carmen Heckmann

## **Taxonomien hochschuldidaktischer Designs und Methoden aus pädagogisch-psychologischer Sicht samt Musterbeispielen aus der aktuellen Forschung**

### **Zusammenfassung**

Durch den Qualitätspakt Lehre ist das Interesse an hochschuldidaktischer Forschung in Deutschland seit 2011 stark gestiegen. Aufgrund der Interdisziplinarität der Hochschuldidaktik nutzen empirische Studien unterschiedliche Forschungsansätze. Im deutschsprachigen Raum fehlen den hochschuldidaktisch Forschenden aber oftmals anregende Vorarbeiten an den jeweiligen Hochschulen vor Ort. Dieser Beitrag möchte dem entgegenwirken und breite Möglichkeiten hochschuldidaktischer Forschung aufzeigen. Zur Klassifikation hochschuldidaktischer Studien werden verschiedene Taxonomien samt Beispielen vorgestellt: Die Makro-, Meso- und Mikro-Ebene, Donabedians (1966) Klassifikation in Prozesse, Strukturen und Ergebnisse sowie die Taxonomie in Reaktion, Lernen, Verhalten und Ergebnisse von Kirkpatrick und Kirkpatrick (2006). Forschungsdesigns wie experimentelle Feldstudien werden vorgestellt und bewertet. Abschließend werden klassische Erhebungsmethoden (Beobachtung, Dokumentenanalyse, Fragebogen, Interview und Testverfahren) behandelt und je Methode eine hochschuldidaktische Studie exemplarisch vorgestellt. Nach Möglichkeit wurden von der Deutschen Gesellschaft für Hochschuldidaktik prämierte Arbeiten ausgewählt.

### **Schlüsselwörter**

Designs; Methoden; Taxonomien; Hochschuldidaktik; Forschungsbeispiele

Immanuel Ulrich & Carmen Heckmann

## **Taxonomies of Higher Education Designs and Methods in Pedagogy and Psychology including Prime Examples of recent Research**

### **Abstract**

Due to the „Qualitätspakt Lehre“ (Teaching quality pact), interest in higher educational research in Germany has increased since 2011. The multidisciplinary of higher education studies results in multiple research approaches. German researchers in higher education often lack inspiring preliminary work at their universities. This article aims at counteracting this circumstance by pointing out broad possibilities of higher educational research. To classify higher educational studies, different taxonomies will be introduced, and examples alongside: Macro-, meso- and micro-level, Donabedians (1966) classification of processes, structures and results, as well as Kirkpatrick's (cf. Kirkpatrick & Kirkpatrick, 2006) taxonomy of reaction, learning, behavior and results. Research designs like experimental field studies are introduced and assessed. Finally, classic survey methods (observations, document analyses, questionnaires, interviews and test procedures) are examined and an exemplary study is presented for each method. Research studies which won the research award of the German Association of Higher Education Didactics were selected, where possible.

### **Keywords**

Designs; Methods; Taxonomies; Higher Education; Research Examples

## 1 Ausgangslage

Während die Hochschuldidaktik in Deutschland trotz langer Tradition über Jahrzehnte nur an wenigen Hochschulen (inkl. Universitäten) etabliert und institutionalisiert war (vgl. Berendt, 2005; Dany, 2007), hat sich seit 2011 durch das *Bund-Länder-Programm für bessere Studienbedingungen und mehr Qualität in der Lehre* („Qualitätspakt Lehre“) die Situation diametral gewandelt. Dank der rund 2 Milliarden Euro starken Mittel des Qualitätspaktes Lehre wurden in der ersten Förderphase 2011-2016 insgesamt 186 Hochschulen, in der zweiten Förderphase von 2016-2020 insgesamt 156 Hochschulen Drittmittel für die Lehre gewährt (Bundesministerium für Bildung und Forschung, 2016).

Die einzelnen Programme haben dabei Fördervolumina zwischen unter 500.000 und über 20 Millionen Euro pro Förderphase und beinhalten dabei verschiedene Maßnahmen (vgl. Schmidt, Heinzemann, Schulze, Walkemeyer & Faaß, 2015):

- Hochschuldidaktische Weiterbildung für v.a. (Nachwuchs-)Lehrende, aber auch Professor(inn)en und studentische Tutor(inn)en
- Hochschuldidaktische Beratung für v.a. (Nachwuchs-)Lehrende, aber auch Professor(inn)en und studentische Tutor(inn)en
- Spezifische Angebote für Professor(inn)en (z.B. Coaching)
- Konzeption innovativer Studienmodelle und neuer Lehr-/Lernkonzepte (kompetenz-, problem- und/oder praxisorientiert)
- Zusätzliche innovative (Präsenz-)Lehrangebote für Studierende, inkl. studentischer Tutorien und Vorkursen (Hauptfokus Studieneingangsphase)
- Betreuung und Beratung von Studienbewerber(inne)n und Studierenden
- Mentoringprogramme für Studierende
- Entwicklung und Angebot neuer E-Learning-Angebote für Studierende
- Flexible Personalmittel/Fonds/Wettbewerbliche Vergabe für gute Lehre
- Einführung und Optimierung des Qualitätsmanagements in der Lehre
- Neueinstellung zusätzlichen Personals (Lehrende, zentrale Einrichtungen wie Prüfungsämter, Medientechnik etc.)
- Projektkoordination der Programme
- Hochschuldidaktische (Begleit-)Forschung, teils samt (inter-)nationaler Kooperation (nur bei einigen Maßnahmen)

Mit dieser Förderung gehen umfangreiche hochschuldidaktische Wirksamkeitsnachweise im Rahmen des Qualitätspaktes Lehre sowie gesonderte Begleitforschungen einher (Bundesministerium für Bildung und Forschung, 2015, 2016).

Folgend sollen der aktuelle Forschungsstand, mögliche Messebenen und konkrete Beispiele aktueller hochschuldidaktischer Forschungsarbeiten thematisiert werden.

## 2 Stand der (Hochschuldidaktik-)Forschung

Hochschuldidaktische Forschung analysiert das Lehren und Lernen an der Hochschule (Battaglia, 2010, S. 29). Als interdisziplinäres Feld umfasst es im weiteren Sinne alle Wissenschaftsdisziplinen, welche sich mit dem individuellen Lehren und Lernen und/oder dem System Hochschule und dessen Auswirkung auf Lehren und Lernen beschäftigen. Die sozialwissenschaftlichen Disziplinen, die sich insbesondere mit der Thematik beschäftigen, sind

- Erziehungswissenschaften
- Psychologie
- Politologie
- Soziologie
- Wirtschaftswissenschaften

Hochschuldidaktikforschende haben somit vielfältige Möglichkeiten, welche Forschungsfragen sie auf welchen Ebenen adressieren wollen (Übersichten vgl. Braun, Weiß & Seidel, 2014 & s.u.; Schaeper & Weiß, 2016; Schaper, 2014b; Wildt & Jahnke, 2010). Aufgrund unseres pädagogisch-psychologischen Forschungshintergrundes fokussieren wir uns auf das individuelle Lehren und Lernen, welches qualitative, quantitative und mixed-method Forschung umfasst.

Aus quantitativen Meta-Analysen zum Lehren und Lernen wissen wir, dass rund 80% der Varianz des studentischen Lernerfolgs durch individuelle Faktoren der Studierenden und Lehrenden bedingt werden, während Faktoren des Systems Hochschule nur 20% ausmachen (Hattie, 2015). Hochschuldidaktische Forschung, welche sich auf studentischen Lernerfolg fokussiert, wird entsprechend dieser Befunde meist in Teildisziplinen durchgeführt, welche Individualfaktoren des Lehrens und Lernens untersuchen, d.h. in erster Linie in den Erziehungswissenschaften und der Psychologie.

Während die hochschuldidaktische Forschung in Deutschland erst seit 2011 durch den Qualitätspakt Lehre eine breite Förderung erhalten hat, sind auf internationaler Ebene bereits in den Jahren zuvor umfangreiche hochschuldidaktische Forschungsstudien vorhanden und publiziert worden (Übersicht vgl. Ulrich, 2016, Kap. 3). Zwar dominieren Schulstudien die pädagogisch-psychologische Forschung (Hattie, 2009, 2015), die Menge der hochschulbezogenen pädagogisch-psychologischen Forschung ist aber so umfangreich, dass valide Schlussfolgerungen zu Individualfaktoren des Lehrens und Lernens gezogen werden können.

Des Weiteren sind die im Qualitätspakt Lehre angedachten Maßnahmen (vgl. Kap. 1) im internationalen Kontext bereits erfolgreich umgesetzt worden, um studentisch Lernerfolge zu steigern. Dies umfasst die Weiterbildung und Beratung der Lehrenden (Hattie, 2015), innovative Lehrkonzepte wie z.B. problembasiertes Lernen (Gijbels, Dochy, Van den Bossche & Segers, 2005), Beratung für Studierende (Reese, Prout, Zirkelback & Anderson, 2010), Mentoring für Studierende (Eby, Allen, Evans, Ng & DuBois, 2008), E-Learning-Angebote für Studierende (Tamim, Bernard, Borokhovski, Abrami & Schmid, 2011) sowie bessere finanzielle Ausstattung der Lehre und mehr Lehrpersonal (Hattie, 2015).

Neben dem generell positiven Einfluss solcher Maßnahmen kommt es substantiell auf deren effektive Ausgestaltung an, wozu aber nicht immer Daten für den Hochschulkontext vorliegen. Gehen wir exemplarisch von einem hochschuldidaktischen Weiterbildungsprogramm aus, welches Nachwuchslehrende durchlaufen, und welches einer Begleitforschung unterliegt. In diesem Fall ermöglichen uns die bisher publizierten englisch- und deutschsprachigen Metaanalysen, die für den studentischen Lernerfolg bedeutsamsten Faktoren guter Lehre abzuleiten. Hierzu sind aber teilweise nur Daten aus Schulkontexten gegeben. Je nach Messung der guten Lehre erhalten wir unterschiedliche Faktoren guter Lehre: Man könnte hier Fragebögen, z.B. die studentische Lehrevaluation heranziehen. Auch wären die Leistungstests der Studierenden, z.B. die Klausurnoten, eine mögliche Messgröße. Tabelle 1 zeigt die – nach aktuellem Forschungsstand der Metaanalysen – fünf wirksamsten Faktoren guter Hochschullehre auf, differenziert nach Auswirkung auf Lehrevaluationen und Leistungsdaten der Studierenden. Die Metaanalysen umfassen Variablen der Lehrenden und Studierenden sowie Rahmenbedingungen der Lehre und deren Wirkung auf die Lernergebnisse der Studierenden. Die Lernergebnisse werden dabei in der Regel in Form von Noten und teilweise durch standardisierte Leistungstests erhoben. Kursiv dargestellt sind die Aspekte, welche nur Daten aus dem Schulkontext aufweisen. Diese sind mit Bedacht zu interpretieren und in künftiger hochschuldidaktischer Forschung zu überprüfen.

Tab. 1: TOP 5 der wirksamsten Aspekte guter Lehre bzgl. studentischen Lernerfolgs, aus Ulrich (2016, S. 34)

Platz	Lehrevaluation	Noten ( <i>kursiv = Befunde aus der Schule</i> )
1	Art, Qualität und Häufigkeit des Feedbacks des Lehrenden an die Studierenden ( $d = 3.53$ )	Anpassung der Lehrmethode je Fähigkeit des Studierende ( $d = 1.68-1.34$ , gute vs. schwache Studierende)
2	Hilfsbereitschaft und Erreichbarkeit des Lehrenden ( $d = 2.20$ )	Gute Vorbereitung (Lehrende) und Struktur Lehrveranstaltung ( $d = 1.39$ )
3	(a) Faire und objektive Notenvergabe ( $d = 2.08$ ) (b) Art und Nutzen der Lehrveranstaltungsmaterialien ( $d = 2.08$ )	<i>Lehrender präsentiert den Studierenden zu Beginn der Stunde seine spezifischen Lernziele für die Lehrveranstaltung (<math>d = 1.37</math> für Kompetenzen)</i>
4	Relevanz und Nutzen der Lehrveranstaltungsinhalte ( $d = 1.96$ )	Klarheit und Verständlichkeit des Lehrenden ( $d = 1.35$ )
5	Freundlichkeit des Lehrenden sowie Interesse an und Respekt gegenüber Studierenden ( $d = 1.71$ )	<i>Klare Betonung von Gemeinsamkeiten und Unterschieden (versch. Theorien) (<math>d = 1.32</math>)</i>

Anm.:  $d$  = Effektstärke Cohens  $d$  (Cohen, 1988)

Es zeigt sich, dass die in Lehrevaluationen am relevantesten eingeschätzten Wirkfaktoren von Lehre das Feedback des Lehrenden an die Studierenden und dabei neben der Art des Feedbacks insbesondere die Qualität, aber auch die Häufigkeit als wichtige Einflussgröße gelten können. Zusätzlich wird eine faire und objektive Benotung der studentischen Leistung als wichtiger Einflussfaktor genannt. Darüber hinaus wird die Lehre nach ihrer Nütz-

lichkeit im Hinblick auf Lehrveranstaltungsmaterialien und -inhalte beurteilt und Beziehungsfaktoren wie die Hilfsbereitschaft, Erreichbarkeit, Freundlichkeit, Interesse und Respekt der Lehrperson als besonders wichtig eingeschätzt.

Für die Leistungsmessung der Studierenden werden meist Noten herangezogen. Für die Erreichung guter Lernleistung konnten in den Metaanalysen insbesondere Hinweise auf die Bedeutung von adaptiv genutzte Lehrmethoden, einer guten Lehrveranstaltungsstrukturierung durch die Lehrperson, die Transparenz von Lernzielen zu Beginn des Lehrveranstaltungstermins sowie Klarheit und Verständlichkeit des Lehrenden in Bezug auf eine klare Einordnung der Inhalte in Theoriebezüge festgestellt werden.

Da die genannten Punkte aber teils großen Interpretationsspielraum lassen (z.B. „Art des Feedbacks“), ist es wichtig, diese in qualitativen Studien (zu Feedback z.B. Johnson, 2000) zu konkretisieren. Bezogen auf die Begleitforschung zu dem o.g. hochschuldidaktischen Weiterbildungsprogramm sollten sich die Inhalte der Weiterbildungen wie auch die Messinstrumente auf die relevanten Faktoren in Tabelle 1 konzentrieren, welche durch qualitative Studien handlungsnah für Lehrende spezifiziert wurden.

Da hochschuldidaktische Forschung aber weitaus mehr Möglichkeiten bietet als „nur“ die Wirksamkeitsüberprüfung der Programme, werden folgend mögliche Messebenen, Forschungsdesigns und Erhebungsmethoden behandelt.

### 3 Forschungsmethoden: Taxonomien

Am Beginn jeder Forschungsarbeit steht zunächst eine Forschungs idee, welche dann im Rahmen von Theorierecherche zu einer spezifischen Fragestellung konkretisiert wird. Diese ist zunächst noch unabhängig davon, ob für die hochschuldidaktische Forschungsarbeit qualitative (Przyborski & Wohrab-Sahr, 2014), quantitative (Stein, 2014) oder mixed-method Studien (Kelle, 2014) konzipiert werden. Da Hochschuldidaktikforschung (meist) in komplexen, sich wechselseitig beeinflussenden Settings stattfindet, sollen zunächst mögliche Messebenen präsentiert und als Anregung erläutert werden.

#### 3.1 Messebenen

Hochschuldidaktische Forschung kann auf verschiedenen Ebenen erfolgen. Zunächst kann man zwischen einer Makro-, Meso- und Mikroebene differenzieren. Diese Ebenen umfassen in hochschuldidaktischen Forschungsdesigns v.a. folgende Bereiche (vgl. Bundesministerium für Bildung und Forschung, 2013; Merkt, 2014):

1. Makro-Ebene: Diese Ebene umfasst die gesamte Hochschule und dort v.a. das strategische Hochschulmanagement (Qualitätsmanagement der Lehre, Governancestrukturen, Professionalisierungsprozesse und Qualifizierungsansätze des (Lehr-)Personals)
2. Meso-Ebene: Diese Ebene beinhaltet Teilbereiche der Hochschule, z.B. Fachbereiche. Hier werden Studiengänge inhaltlich (Curriculumsgestaltung, Entwicklung innovativer Studienmodelle), organisatorisch (Studienorganisation, Studiengangmanagement) und auch infrastrukturell (z.B. Medienausstattung) in den Blick genommen.

3. Mikro-Ebene: Diese Ebene fokussiert sich auf das individuelle Lehren, Lernen und Prüfen der Lehrenden und Studierenden. Allgemeine Lehr- und Lernprozesse, fächerspezifische und fachübergreifende Maßnahmen, spezifische Phasen im Studienverlauf (z.B. Studieneingangsphase) sind von Interesse.

In der Hochschuldidaktik wird meistens die Mikro-Ebene in den Blick genommen. Auf dieser Ebene überwiegen psychologische und erziehungswissenschaftliche Forschungsstudien. In der Meso-Ebene und Makro-Ebene kommen neben psychologischen und erziehungswissenschaftlichen stärker soziologische, politologische und wirtschaftswissenschaftliche Forschungsstudien vor. Insgesamt ist diese fachliche Differenzierung bei hochschuldidaktischer Forschung aufgrund thematischer Überschneidungen nicht immer sinnvoll.

Eine weitere Taxonomie differenziert zwischen Prozessen, Strukturen und Ergebnissen (Donabedian, 1966) und adressiert sowohl die Makro- als auch die Meso- und Mikroebene eines Systems:

- Prozesse umfassen die Maßnahmen und Aktivitäten der Lehrenden, Studierenden und weiteren Mitarbeitenden der Hochschule. Dies sind z.B. die Implementation einer Lehrevaluation (Makro-Ebene), die Akkreditierung eines Studiengangs (Meso-Ebene) sowie die Lehraktivitäten des Lehrenden und Lernaktivitäten der Studierenden (Mikro-Ebene). Prozesse sind durch die jeweiligen Initiatorinnen und Initiatoren unmittelbar steuerbar.
- Strukturen beinhalten die Rahmenbedingungen und die Qualifikation der Personen. Dies sind z.B. ein Neubau des Hochschulcampus (Makro-Ebene), die Ausstattung der Labore (Meso-Ebene) sowie die hochschuldidaktische Qualifikation der Lehrenden und das Vorwissen der Studierenden zu Lehrveranstaltungsbeginn (Mikro-Ebene). Die Strukturen beeinflussen die Performanz der Prozesse und somit die Ergebnisse.
- Ergebnisse stellen die Wirkungen der Prozesse dar, wobei die Prozesse durch die Strukturen positiv wie negativ mitbeeinflusst werden können. Dies sind z.B. eingeworbene Drittmittel für die Lehre oder die Gesamtzahl der Absolvent(innen) der Hochschule (Makro-Ebene) bzw. je Studiengang (Meso-Ebene) sowie der Lernerfolg und Kompetenzgewinn der Studierenden (Mikro-Ebene).

Braun, Weiß & Seidel (2014) kombinieren die Makro-, Meso- und Mikroebene mit Donabedians (1966) Taxonomie und erweitern dies zu einem Angebots-Nutzungs-Modell zur Erklärung von Lehr-Lern-Prozessen an der Hochschule. Da diese v.a. auf der Mikro- bzw. Individualebene stattfinden, ist diese besonders ausdifferenziert:

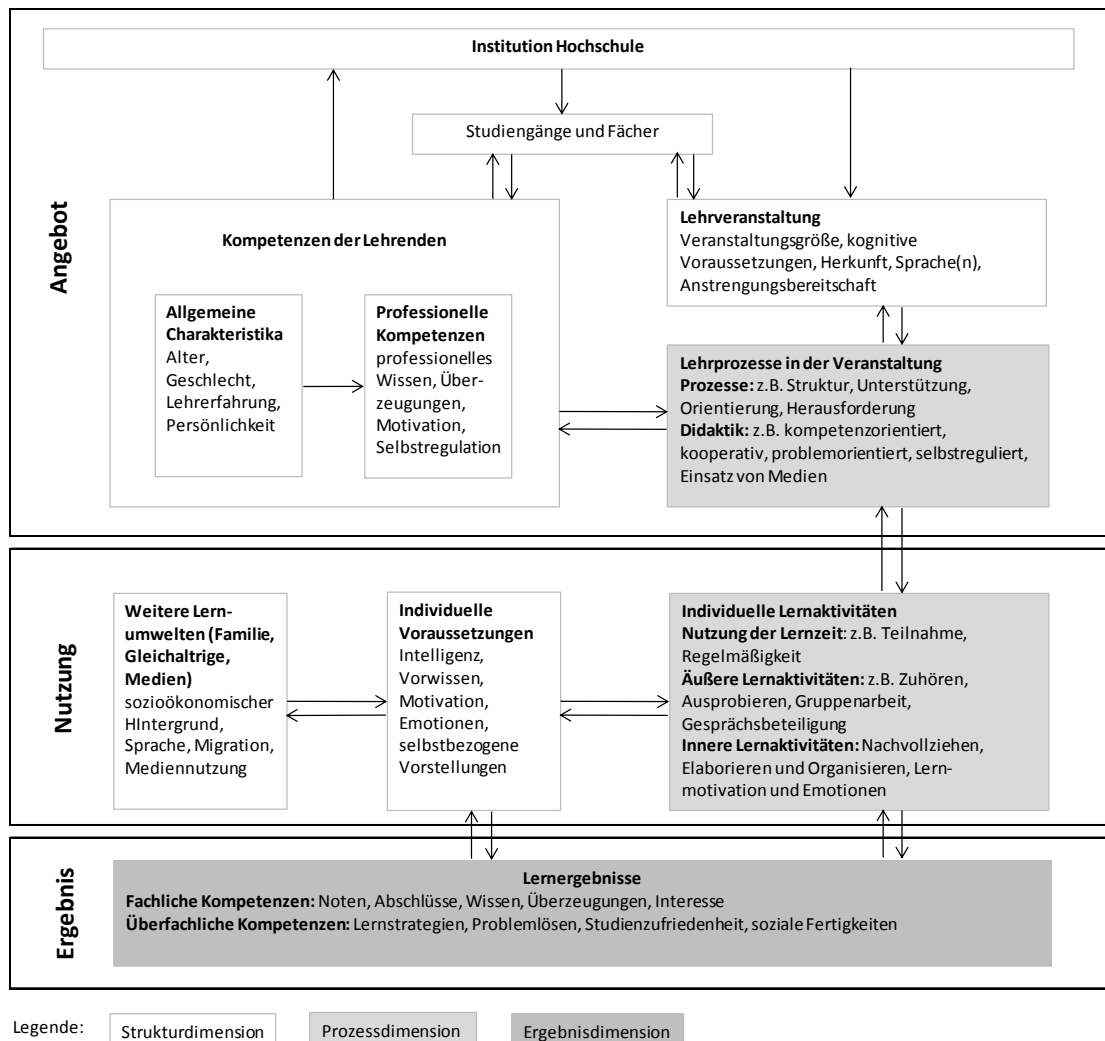


Abb. 1: Angebots-Nutzungs-Modell zur Erklärung von Lehr-Lern-Prozessen an der Hochschule (Braun, Weiß & Seidel, 2014, S. 438)

Eine andere Taxonomie von Kirkpatrick und Kirkpatrick (2006), welche Ähnlichkeiten zu der von Donabedian (1966) hat, differenziert zwischen Reaktion, Lernen, Verhalten und Ergebnissen. Ihre Taxonomie stammt aus dem Weiterbildungsbereich, ist dort bis heute Standard (vgl. Kauffeld, 2016; Schaper, 2014a) und tendenziell nur auf der Mikro-Ebene zu verorten. Bezogen auf die Hochschuldidaktik eignet sich diese Taxonomie v.a. für die hochschuldidaktische Weiterbildung und dazugehörige Begleitforschung zu den Veränderungen, welche die Weiterbildung hervorgerufen hat:

1. **Reaktion:** Diese Ebene umfasst die Akzeptanz und Zufriedenheit mit der Maßnahme (z.B. eine hochschuldidaktische Weiterbildung) und wird meist unmittelbar danach erhoben.
2. **Lernen:** Diese Ebene fokussiert sich auf den Lernfortschritt bzw. Kompetenzgewinn der Teilnehmenden der Maßnahme.
3. **Verhalten:** Diese Ebene umfasst das (auf Basis des Kompetenzgewinns neu) gezeigte Verhalten der Teilnehmenden der Maßnahme nach Abschluss der Maßnahme. Idealerweise ist das Verhalten nun besser als vorher.



4. Ergebnisse: Diese Ebene konzentriert sich auf die Ergebnisse, welche durch das geänderte Verhalten der Teilnehmenden entstanden sind.

Bezogen auf Donabedians Taxonomie wären Kirkpatrick's Ebenen 1 und 3 eher Donabedians „Prozessen“ und Kirkpatrick's Ebenen 2 und 4 eher Donabedians „Ergebnissen“ zuzuordnen. Kirkpatrick's Ebenen sind zeitversetzt, d.h. die Ebenen 1 und 2 sind unmittelbar nach der Maßnahme, die Ebenen 3 und 4 davon nur zeitlich versetzt zu erheben.

Kirkpatrick setzt mit seinen Ebenen eine Hierarchie fest – die untere Ebene ist die Voraussetzung für die höhere – was sich gut dazu eignet, Weiterbildungsprogramme und deren Wirksamkeit zu erheben, die Empirie zur dieser hierarchischen Festlegung ist jedoch schwach (Bates, 2004). Bezogen auf typische Zielgruppen wie Studierende, Lehrende und deren Kollegen sind in Kirkpatrick's Taxonomie folgende Untersuchungsgegenstände denkbar (vgl. Tabelle 2):

Tab. 2: Mögliche Untersuchungsgegenstände je Ebene nach Kirkpatrick (Kirkpatrick & Kirkpatrick, 2006), aus Ulrich & Heckmann (2013, S. 11)

Ebene	Zielgruppen		
	Lehrende	Studierende	Kollegen
<b>Reaktion</b>	Zufriedenheit mit der hochschuldidaktischen Qualifikation	Zufriedenheit mit der Lehre des Dozierenden	Arbeitszufriedenheit durch weniger notwendige Lehrunterstützung des lehrenden Kollegen
<b>Lernen</b>	Hochschuldidaktische Kompetenz	Lernerfolg in der Veranstaltung	Didaktische Anregungen durch Lehrhospitation beim „frisch“ qualifizierten Lehrenden
<b>Verhalten</b>	Lehre	Motivation bei der Klausurvorbereitung	Gründung eines abteilungsinternen Lehrqualitätszirkels
<b>Ergebnisse</b>	Lehrevaluations-ergebnisse, Lehrpreis	Klausurergebnisse	Lehrpreis für die gesamte Arbeitseinheit

### 3.2 Designs

Hinsichtlich des Forschungsdesigns bestehen unterschiedliche Möglichkeiten. Bei allen Designs finden mindestens eine Intervention (z.B. eine hochschuldidaktische Weiterbildung) und eine Messung (z.B. Befragung der Teilnehmenden der hochschuldidaktischen Weiterbildung) statt. Unsere Designs (vgl. Döring & Bortz, 2016, Kap. 7) können wir differenzieren in

- Querschnitts- vs. Längsschnittuntersuchung
- Labor- vs. Feldstudie
- Experimentelle, quasi-experimentelle und nicht-experimentelle Studien

Eine Querschnittsuntersuchung erfasst die untersuchten Subjekte nur einmal, während eine Längsschnittstudie diese mindestens zweimal, meist vor und nach einer Interventi-

on, erfasst. Aus wissenschaftlicher Sicht sind Längsschnittstudien gegenüber Querschnittstudien zu bevorzugen, da erstere die Abbildung eines Verlaufs erlauben (Ausnahme: stabile Merkmale, was bei Individuen selten ist). Nicht-experimentelle Quer- und Längsschnittstudien werden Ex-post-facto-Designs genannt, Beispiele für Nicht-experimentelle Längsschnittstudien sind Paneldesigns (gleiche Befragte je Messung), Trenddesigns (unterschiedliche Befragte je Messung) und Kohortendesigns (wie Panel, aber zusätzlich Differenzierung der Befragten in Kohorten).

Laborstudien finden unter möglichst vielen kontrollierten Bedingungen statt, um die Wirkung der Intervention möglichst störungsfrei zu erfassen. Jedoch ist die Generalisierbarkeit ‚auf die Realität‘ eingeschränkt. Feldstudien finden unter möglichst realen Bedingungen statt, um die Wirkung der Intervention ‚in der Realität‘ zu erheben. Jedoch ist die störungsfreie Messung der Intervention eingeschränkt. Anwendungsforschung bevorzugt die Feldforschung. Über Kontrollgruppen und Randomisierung kann die störungsfreie Messung der Intervention in der Feldforschung sichergestellt werden.

Experimentelle und quasi-experimentelle Studien umfassen Kontrollgruppen, nicht-experimentelle Studien verzichten auf diese. Um Aussagen über die Wirkung einer Intervention treffen zu können, ist es hilfreich, eine Vergleichsgruppe mit zu erheben, um andere Effekte neben der Intervention ausschließen zu können (z.B. Reifungseffekte). Hierbei können als Vergleichsgruppe idealerweise Kontrollgruppen gewählt werden, die keinerlei Intervention erhalten. Falls das Unterlassen einer Intervention ethisch nicht vertretbar wäre, kann man auf weitere Interventionsgruppen ausweichen, welche eine andere Intervention erhalten. Teils kann man dies umgehen, indem man mit Wartekontrollgruppen arbeitet – diese erhalten eine Intervention, aber erst nach einem gewissen Zeitraum, und können so zuvor als klassische Kontrollgruppe fungieren.

Ein Problem bei Kontrollgruppen besteht in den Selbstselektionsprozessen der Studienteilnehmer(innen): ob ein Lehrender an einer hochschuldidaktischen Weiterbildung (hier Interventionsgruppe) teilnimmt oder nicht (hier Kontrollgruppe), entscheidet der Lehrende meist selbst. So nehmen an hochschuldidaktischen Weiterbildungen eher Lehrende teil, denen Lehre wichtig ist. Dies erschwert den Vergleich zur Kontrollgruppe. Hier ist es wichtig, potenzielle Störfaktoren in beiden Gruppen mit zu kontrollieren (z.B. wäre hochschuldidaktisches Vorwissen ein „Störfaktor“, wenn man den hochschuldidaktischen Wissenserwerb durch einen hochschuldidaktischen Workshop erfassen möchte). In diesem Fall spricht man von einem quasi-experimentellen Design: eine Kontrollgruppe ist gegeben, aber die Zuordnung zur Kontroll- oder Interventionsgruppe erfolgt unkontrolliert.

Um die *störungsfreie* Messung der Intervention zu gewährleisten, erfolgt die Zuordnung der Teilnehmenden zu der Interventions- oder Kontrollgruppe idealerweise zufallsbasiert, d.h. randomisiert. In diesem Fall ist ein sog. experimentelles Design gegeben. Dieses Design ist wissenschaftlich zu bevorzugen, da es sichere Kausalschlüsse zur Intervention ermöglicht, aber schwierig umsetzbar.

In der hochschuldidaktischen Forschung bieten sich daher (meist) Studiendesigns an, welche als Längsschnitt, im Feld und experimentell stattfinden. Da diese aber auch sehr aufwändig und zugleich oftmals externe Restriktionen gegeben sind, wird dieser Standard nur teilweise erreicht.

Vor der Auswahl der Erhebungsmethode erfolgt zunächst noch die Operationalisierung und die Festlegung der Stichprobenziehung (vgl. Döring & Bortz, 2016, Kap. 8 & 9).

### 3.3 Erhebungsmethoden

Als qualitative und quantitative Erhebungsmethoden in den Sozialwissenschaften differenzieren Döring und Bortz (2016, Kap. 10) Beobachtung, Dokumentenanalyse, Fragebogen, Interview, physiologische Messung und Testverfahren. Meist ist ein Einsatz mehrerer Methoden günstig, um die Objektivität der Erhebung zu erhöhen, wenngleich die Datenerhebung durch den Einsatz mehrerer Methoden in der Durchführung aufwändiger ist. Da physiologische Messungen im Rahmen der Hochschuldidaktik sowohl unüblich als auch auf dem aktuellen Stand der Technik in Teilen unmöglich sind, verzichten wir folgend auf deren Ausarbeitung. Tabelle 3 gibt eine Übersicht:

Tab. 3: Mögliche Erhebungsmethoden nach Döring & Bortz (2016) je hochschuldidaktischer Zielgruppe, adaptiert aus Ulrich & Heckmann (2013, S. 9)

Ebene	Zielgruppen		
	Lehrende	Studierende	Kollegen
<b>Beobachtung</b>	Videoaufzeichnung der Lehre	Studentische/r Tutor/in im Kurs	Hospitierende/r Kollegin/Kollege im Kurs
<b>Dokumentenanalyse</b>	Vergleich Kursstruktur vor und nach der Weiterbildung (identischer Kurs alle zwei Semester)	Einschätzung der Qualität der Kurshausaufgaben vor und nach der Weiterbildung (identischer Kurs alle zwei Semester)	Einschätzung der Qualität der Kurshausaufgaben vor und nach der Weiterbildung (identischer Kurs alle zwei Semester)
<b>Fragebogen</b>	Selbsteinschätzung der eigenen Kompetenz vor und nach der Weiterbildung	Lehrevaluation (Fremdeinschätzung Dozierender und/oder Selbsteinschätzung Lernerfolg)	Anonymes, gemittelt Feedback aller Kollegen
<b>Interview</b>	Zufriedenheit der Weiterbildungsteilnehmer(innen)	Einzelinterview besonders engagierter vs. demotivierter Studierender bzgl. der Ursachen dessen	Einschätzung der didaktischen Stärken und Schwächen des Lehrenden
<b>Testverfahren</b>	Hochschuldidaktischer Wissenstest	Vergleich Klausurergebnisse vor und nach der Weiterbildung (identischer Kurs alle zwei Semester)	Stresstest zum Semesterbeginn (Lehrbelastung durch Einarbeitung neuer Kollegen)

Alle dargestellten Erhebungsmethoden verfügen über Stärken und Schwächen. Digital gestützte Beobachtungen ermöglichen genaue Analysen des Lehrprozesses im Feld, jedoch stellt der Beobachtende eine Störquelle dar, die zu potenziell sozial erwünschten Verhaltensweisen (vgl. Diekmann, 2009, S. 447 ff.) auf Seiten der beobachteten Lehrenden und Studierenden führt (was durch eine Befragung nach der Beobachtung kontrolliert werden kann).

Dokumentenanalysen ermöglichen genaueste Analysen schriftlichen Materials (sofern ausreichend vorhanden). Bei Unklarheiten bzw. mangelnder Kenntnis des Sachgegenstandes ist für den Untersuchenden aber die Möglichkeit, bei den Autor(inn)en der Dokumente nachzufragen, selten gegeben.

Fragebögen ermöglichen anonyme Auskünfte der Befragten in großer Zahl, v.a. bei vorgegebenen Antwortkategorien (z.B. „Lehren Sie gerne?“, Antwortskala von 1 „sehr gerne“ bis 5 „sehr ungerne“). Jedoch können durch die vorgegebenen Antwortkategorien Informationen im Erhebungsprozess entfallen (z.B. „Ich lehre gerne in Seminaren, aber ungerne in Vorlesungen.“). An deutschen Hochschulen wird die Qualität der Lehre aufgrund gesetzlicher Vorgaben per Lehrevaluationsfragebogen an die Studierenden erhoben. Auch wenn die Lehrevaluation ihre Grenzen hat (vgl. Wolbring, 2013; 2014), können diese ohnehin vorliegenden Daten – unter Berücksichtigung datenschutzrechtlicher und forschungsethischer Vorgaben – potenziell in hochschuldidaktische Studien integriert werden (z.B. als reflexives Feedbacktool).

Interviews bieten im Gegensatz zum Fragebogen interaktive Möglichkeiten, detailliert Personen zu befragen und bei unzureichenden Antworten gezielt nachzufragen. Jedoch besteht, ähnlich wie bei Beobachtungen, das Problem der sozialen Erwünschtheit (z.B. „Lehren Sie gerne?“). Hier könnte man z.B. zunächst per Fragebogen Beliebtheits- und Unbeliebtheitsfaktoren von Lehre erfragen und im Interview diese genauer eruieren (z.B. „Warum glauben Sie, schreiben einige Lehrende ‘Lehre nütze ihnen nichts?’“).

Testverfahren sind, sofern sie valide konzipiert sind, sehr hilfreich zu Erhebung von Fachkompetenzen. Jedoch ist ihre Akzeptanz gering, wenn sie nicht von übergeordneten Institutionen vorgegeben wurden. Die deutschen Hochschulen erfassen per Testverfahren (Klausur) die Fachkompetenzen ihrer Studierenden. Diese ohnehin vorliegenden Daten können – unter Berücksichtigung datenschutzrechtlicher und forschungsethischer Vorgaben – potenziell in hochschuldidaktische Studien integriert werden.

Nach der Erhebung werden die Daten aufbereitet, analysiert und präsentiert (vgl. Döring & Bortz, 2016, Kap. 11-13; Field, 2013; Hussy, Schreier & Echterhoff, 2013).

## 4 Konkrete hochschuldidaktische Forschungsbeispiele

In den Unterkapiteln dieses Abschnittes sollen exemplarisch hochschuldidaktische Forschungsarbeiten dargestellt werden, welche sich (v.a.) einer der o.g. Erhebungsmethoden bedienen. Soweit möglich, wurden Arbeiten ausgewählt, die den Johannes-Wildt-Nachwuchspreis für hochschuldidaktische Forschung der Deutschen Gesellschaft für Hochschuldidaktik (dghd) erhalten haben (Deutsche Gesellschaft für Hochschuldidaktik, 2016): Die prämierten Arbeiten sind alle in den letzten vier Jahren als Qualifikationsarbeiten eingereicht und gut benotet worden. Zugleich wurden sie in Konkurrenz zu anderen

hochschuldidaktischen Arbeiten von einer fachlich diversen Preisjury ausgewählt. Dies betrifft die Unterkapitel 4.2 Dokumentenanalyse (Centeno García, 2016), 4.3 Fragebogen (Ulrich, 2013) und 4.4 Interview (Trautwein, 2014).

Die vorgestellten Studien sind oftmals multimethodal (z.B. Interview und Fragebogen), haben aber einen Hauptfokus auf die Erhebungsmethode, unter der sie hier eingeordnet wurden. Der Vollständigkeit halber sei erwähnt, dass auch bei Ulrich (2013), Trautwein (2014) sowie Deslauriers, Schelew und Wieman (2011) Beobachtungen und bei Trautwein (2014) eine Dokumentenanalyse von Lehrportfolios durchgeführt wurden. Des Weiteren wurden auch bei Heckmann (eingereicht) Fragebögen und Interviews sowie bei Deslauriers, Schelew und Wieman (2011) Fragebögen eingesetzt.

#### 4.1 Beobachtung

Um die Weiterentwicklung von Lehrveranstaltungsgeschehen zu fördern und Kompetenzen von Lehrenden möglichst objektiv zu erfassen, werden in der schulischen Unterrichtsforschung seit vielen Jahren (vgl. Helmke, 2003, Kunter, 2005, Seidel & Prenzel, 2007) sowie seit einigen Jahren auch in der Hochschuldidaktik erfolgreich Videoanalysen als Forschungsmethode eingesetzt (z.B. Johannes & Seidel, 2010; Seidel & Hoppert, 2011; Wibbecke, 2015). Lehrkompetenzen lassen sich auf diese Weise auf Videos in Form real gezeigten Verhaltens abbilden. Die Möglichkeit der mehrfachen Betrachtung einer Situation, ggf. auch aus mehreren Perspektiven, ermöglicht die vielfältige Analyse der Lehrkompetenzen einer Lehrperson. Diese Form der Auswertung komplexer Prozesse, kann zu einer höheren prognostischen und inhaltlichen Validität der getroffenen Aussagen über das Lehrverhalten gegenüber Selbstbeobachtungsverfahren beitragen (Oser, Curcio & Düggeli, 2007). Allerdings müssen für videoanalytische Forschungsarbeiten vergleichsweise hohe personelle und finanzielle Ressourcen eingeplant werden.

Hierbei ist zu entscheiden, ob ein hoch oder ein niedrig inferentes Beobachtungsinstrument genutzt wird. Letzteres erhebt konkrete Ereignisse, Organisationsformen etc. der Lehrveranstaltungsgestaltung anhand von festgelegten Indikatoren, also direkt beobachtbaren Lehr- oder Lernaktivitäten innerhalb festgelegter Zeitabschnitte (Hugener, Pauli & Reusser, 2006). Anhand der qualitativen Inhaltsanalyse nach Mayring (2002) wird das Kategoriensystem entworfen und bzgl. Beobachterübereinstimmung geprüft. Hoch inferente Beobachtungssysteme hingegen berücksichtigen auch Aspekte des Lehrverhaltens, die über das konkret beobachtbare Verhalten hinaus auf die Verhaltenstendenzen der Lehrperson schließen lassen. Für diese Art der Unterrichtsbeurteilung müssen komplexe Merkmale berücksichtigt und interpretative Prozesse in den Beurteilungsprozess einbezogen werden (Hugener et al., 2006).

Ein niedrig inferentes Beobachtungssystem kam beispielsweise in einer Studie von Johannes und Seidel (2010) zum Einsatz, die ein einjähriges Trainingsprogramm für Lehranfänger(innen) entwickelten, in dem diese über zwei Semester beim Lehreinstieg begleitet wurden. Neben der Teilnahme an einem Workshopangebot zur Erweiterung ihres didaktischen Methodenrepertoires, wurde eine Videoaufzeichnung einer Lehrveranstaltungseinheit der teilnehmenden Lehrenden erstellt, zu der sie ein einstündiges individuelles Feedback erhielten (Johannes & Seidel, 2012). Zusätzlich wurden ausgewählte Videosequenzen für eine Gruppenreflexion in das Training integriert und die Selbstein-

schätzung der Lehrendenfokussierung der Lehrenden erhoben. Bei den Feedbackelementen dieser Studie kam auch das von Seidel und Hoppert (2011) entwickelte Beobachtungsinstrument zum Einsatz. Auf diese Weise wurden die Videoanalysen gezielt zur Schulung der Hochschullehrkräfte eingesetzt. Gleichzeitig wurde das Instrument genutzt, um mögliche Unterschiede zwischen selbsteingeschätzter und real beobachteter Lehrenden- bzw. Studierendenfokussierung zu untersuchen. Es zeigte sich, dass durch videobasierte Trainings die professionelle (Selbst-)Wahrnehmung von Lehrkräften geschult werden kann. Diese Verknüpfung videobasierter, hochschuldidaktischer Forschung und der (Weiter-) Qualifizierung von Hochschullehrenden führte demnach zu interessanten Erkenntnissen.

An der Goethe-Universität Frankfurt wurde anhand theoretischer Überlegungen von Heckmann (eingereicht) ein hoch inferentes Ratingverfahren entwickelt und in einer mixed-method-Studie eingesetzt, um anhand von Videoanalysen der Frage nachzugehen, inwiefern sich professionelles Verhalten von Lehrkräften an der Hochschule im Verlauf einer Weiterbildungsmaßnahme über die Zeit entwickelt. Die Forschungsmethodik der direkten Beobachtung mittels Videoanalysen wurde hier zur Betrachtung und Analyse von Lehrkompetenzen eingesetzt. Das hoch inferente Rating dient der Einschätzung der Lehrqualität anhand von festgelegten Kriterien. Diese interpretative Beurteilungsform erfordert ein relativ hohes Maß an Expertise im Beobachtungsfeld, um das komplexe (Lehrveranstaltungs-)Geschehen adäquat beurteilen zu können. Nach der theoriegeleiteten Konzeption des Ratings (u. a. nach Bos & Tarnai, 1999; Johannes, Fendler, Hoppert & Seidel, 2011) wurden insgesamt acht Bewertungsdimensionen (z.B. Lernzielorientierung und Aktivierung von Studierenden) festgelegt und jeder Bewertungsdimension mehrere Verhaltensindikatoren zugeordnet. Diese wurden anschließend zur Analyse von 36 videografierten Seminarveranstaltungs-sitzungen von insgesamt 12 Lehrpersonen eingesetzt, die am Weiterbildungsprogramm zum Zertifikat Hochschullehre des Interdisziplinären Kollegs Hochschuldidaktik der Goethe-Universität Frankfurt teilnahmen.

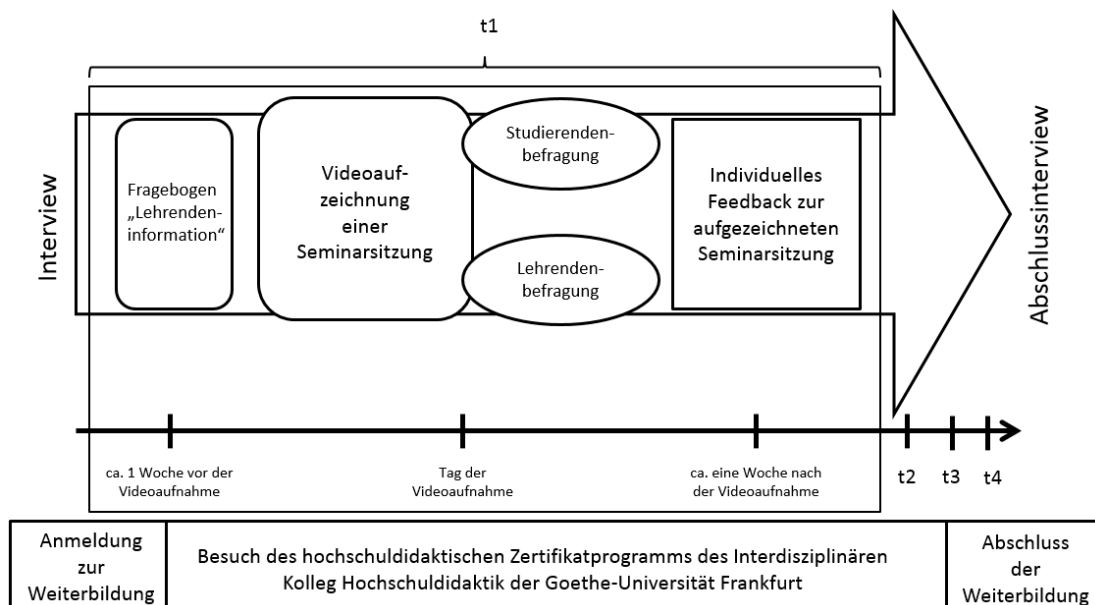


Abb. 2: Design der Studie von Heckmann (eingereicht).

Die Lehrpersonen wurden während ihrer Weiterbildungsteilnahme einmal pro Semester videografiert. Die Überprüfung der Interraterreliabilität des Instrumentes ergab bei der statistischen Prüfung eine zufriedenstellende bis exzellente Messgenauigkeit ( $p \geq .6$  bis  $.95$ ). Zur Analyse und Unterstützung der Kompetenzentwicklung von Lehrenden ist aufgrund der Verlaufsdaten davon auszugehen, dass individuelle Faktoren der Lehrperson, die in dieser Studie durch Interviews erfasst wurden, berücksichtigt werden sollten, um den besten Lernzuwachs zu erzielen und Einflussfaktoren auf den Lernzuwachs bestmöglich interpretieren zu können. Die Teilnehmer(innen) von hochschuldidaktischen Weiterbildungsveranstaltungen stellen u.a. hinsichtlich ihrer Lehrerfahrung, erhaltener Unterstützung im Fachbereich beim Lehreinstieg, aber auch im Hinblick auf die Flexibilität der Lehrgestaltung durch Rahmenvorgaben, eine sehr heterogene Gruppe dar. Dies führt dazu, dass mit reinen Beobachtungsdaten zwar Aussagen darüber getroffen werden können, ob eine Verhaltensänderung bei der einzelnen Lehrperson stattgefunden hat. Welche Einflussfaktoren jedoch dazu führten, dass eine Änderung abbildbar ist oder nicht, kann durch diese Daten nicht abschließend abgebildet werden. Für die hochschuldidaktische Beratung, beispielsweise durch ein (Video)Feedback nach Hospitationen, kann ein hoch inferentes Beobachtungssystem einen hilfreichen Beitrag leisten, der die beratene Lehrperson dabei unterstützen kann, ihr Potential bestmöglich auszuschöpfen (Heckmann, eingereicht).

Durch die Forschungsmethode Beobachtung (z.B. Heckmann, eingereicht; Johannes & Seidel, 2010) können demnach vielfältige Aspekte des Lehrgeschehens durch externe Beobachter(innen) ausgewertet werden: Zum einen im Rahmen hochschuldidaktischer Weiterbildungs- und Beratungsangebote für Reflexions- und Feedbackzwecke, zum anderen zur Prüfung hochschuldidaktischer Weiterbildungen, ob auf die theoretische und durch Übungen angeleitete Vermittlung bestimmter Aspekte guter Lehre eine tatsächliche Umsetzung im Lehrhandeln erfolgt. Eingeschränkt werden diese Beobachtung durch das Problem sozialer Erwünschtheit, da sich Lehrende in beobachteten Lehrveranstaltungen tendenziell stärker anstrengen. Dies kann in Teilen durch häufigere Beobachtung und/oder Einschätzungen der Studierenden (z.B. „Verhielt sich der/die Lehrende heute wie üblich?“) kompensiert werden.

#### **4.2 Dokumentenanalyse: Modulhandbücher**

Für die Datenerhebungsmethode der Dokumentenanalyse werden schriftliche Quellen verwendet, um Informationen zu analysieren. Hierzu können unterschiedlichste Formen von schriftlichen Dokumenten verwendet werden, wie beispielsweise Broschüren, Ratgeber, Zeitungsartikel oder ähnliches. Anhand der Dissertationsschrift von Centeno García (2016) soll das Vorgehen der Dokumentenanalyse als Datenerhebungsmethode näher erläutert werden.

Um herauszuarbeiten, welche Voraussetzungen für eine angemessene didaktisch-methodische Gestaltung der Textarbeit (diese umfasst Kriterien für Textauswahl und Textvernetzung) in der geisteswissenschaftlichen Lehre erforderlich sind, führte die Autorin einen iterativen Prozess von drei Phasen durch: Konturieren, theoretisches Sampling und Konzeptualisieren. Dabei wurden die Praxis- und die Theorieebene fortwährend aufeinander bezogen, um die gesichteten Dokumente der Praxisliteratur in der Theorie

derart zu verorten, dass durch eine Konturierung Schlüsselkategorien und Untersuchungsdimensionen aus der Literatur abgeleitet werden können.

Ausgewertet wurden verschiedene praxisbezogene Texte inkl. Modulhandbüchern, die sich mit dem Thema Textarbeit in der Hochschullehre für geisteswissenschaftliche Fächer beschäftigen. Textarbeit kann einerseits über die Auseinandersetzung mit einem Einzeltext erfolgen, und andererseits anhand übergeordneter Ziele Tätigkeits- und Kommunikationsprofile erarbeiten, für welche die Einhaltung der beschriebenen Prozessfolgen einzuhalten und angemessene Textsorten und Dokumente erfasst werden müssen. Von der Autorin gefundene Schlüsselkategorien und Untersuchungsdimensionen wurden im Prozess auf die Theorie bezogen und erneut hinterfragt, bis als Ergebnis des Prozesses die Konzeptualisierung zentraler Kategorien feststand. Das Forschungskonzept sieht jedoch die gefundenen Kategorien nicht als abgeschlossenes Ergebnis, sondern vielmehr als Möglichkeit der Anwendung eines systemisch-dynamischen Ansatzes, um komplexe Themen wie die Textarbeit zu analysieren. Das erarbeitete Modell kann als Kompetenzprofil der professionellen Leser(innen) verstanden werden, das eine reflektierende, intertextualisierende, soziologisch denkende und transformierende Analyse im Bezugsrahmen Text, Kontext, Community und Selbst erfordert. Somit kann Lesekompetenz als eine der Kernkompetenzen wissenschaftlichen Arbeitens im Sinne eines kontrolliert und systematisch hinterfragenden Vorgehens im Wissenschaftssystem beschrieben werden.

Als zentrales Ziel der Textarbeit wird in der Arbeit von Centeno García (2016) die Befähigung zum professionellen Handeln beschrieben. Diese wiederum führt zur Ableitung bestimmter Kompetenz- bzw. Lernziele, wobei der Erwerb einer Textsortenkompetenz für Studierende als Teil ihrer Fachkompetenz zu betrachten ist. Zentraler Bestandteil der Hochschullehre sollte dabei die bedachte Auswahl von Texten und Textsorten im Hinblick auf die Vielschichtigkeit der kommunikativen und textuellen Zusammenhänge eines Faches darstellen. Nach der Festlegung auf das Lernziel der Textarbeit, kann anschließend mit der methodischen Planung der Lehre begonnen werden.

Für die hochschuldidaktische Praxis erarbeitete Centeno García einige Professionalisierungsziele für Lehrende. Darunter fallen beispielsweise die Verbindung von fachlichem und sprachlichem Lernen, die Durchführung sprach- und kommunikationssensibler Lehre durch Transparenz und Reflektion, die Berücksichtigung größerer sprachlicher und kommunikativer Einheiten sowie die Zusammenführung von lehrbezogenen und lernbezogenen (studentischen) und (wissenschafts-) praxisbezogenen Text- und Textsortennetzen.

Die Autorin fasst zusammen, dass die Verknüpfung erkenntnistheoretischer, kommunikations- und sprachwissenschaftlicher sowie soziologischer Betrachtungsweisen für die Theoriebildung ein ergänzendes kohärentes und angemessenes Bild ergeben, ein entsprechendes didaktisches Konzept für die Weiterbildung von Lehrenden in Hinblick auf sozialisatorische Textarbeit in geisteswissenschaftlichen Fächern jedoch bisher aussteht, aber wünschenswert wäre.

Die Dissertation von Centeno García (2016) hat zum einen ein Kompetenzprofil des professionellen Lesenden erarbeitet, was über die Geisteswissenschaft hinaus hochschuldidaktisch nutzbar ist. Zum anderen sind die dargelegten Professionalisierungsziele für Lehrende eine wertvolle Grundlage für hochschuldidaktische Weiterbildungen. Zu



klären bleibt, ob das Kompetenzprofil des professionellen Lesenden auch bei formallastigen Texten gültig ist oder dafür ggf. modifiziert werden muss.

### 4.3 Fragebogen: Wirksamkeit hochschuldidaktischer Workshops

In der folgend vorgestellten Dissertation (Ulrich, 2013) wurde ein hochschuldidaktischer Workshop auf Basis des aktuellen (Hochschuldidaktik)Forschungsstandes konzipiert und mit 79 Nachwuchslehrenden durchgeführt. Seine Wirksamkeit wurde anhand der mehrfachen Befragung der Lehrenden und ihrer 3.183 Studierenden überprüft. Die Befragung erfolgte per Fragebogen, wobei nach Möglichkeit auf bereits publizierte und validierte Fragebögen zurückgegriffen wurde.

Um möglichst valide Aussagen über die Wirksamkeit des hochschuldidaktischen Workshops treffen zu können, wurde bei den Lehrenden ein experimentelles Längsschnittsdesign im Feld gewählt. Hierzu wurden alle teilnahmebereiten Lehrenden zu Beginn der Studie (Januar 2010) in zwei (Experimental-)Gruppen (EG) randomisiert. Die zweite Gruppe erhielt die Intervention in Form des hochschuldidaktischen Workshops ein Semester später, und konnte so zuvor als Kontrollgruppe fungieren. Dieses sogenannte „Mehr-Phasen-Design“ hatte somit nur weiterbildungsbereite Lehrende in beiden Gruppen, so dass Unterschiede in beiden Gruppen kausal auf die Intervention rückführbar waren.

Nach einer ersten Messung vor Beginn der Intervention im Januar 2010 erfolgte in den folgenden Semesterferien die Intervention (siehe X in Abb. 1) samt Befragung für die erste Lehrendengruppe. Im Semester darauf, im Juli 2010, wurde eine erneute Befragung vorgenommen. In den anschließenden Semesterferien im September und Oktober 2010 fand die Intervention samt Befragung für die zweite Lehrendengruppe statt. Im Semester darauf, im Januar 2011, fand die finale Befragung statt. Abbildung 1 gibt einen Überblick zu allen Messzeitpunkten.

Pro Semester wurden jeweils auch die Studierenden der Lehrenden im Rahmen der Lehrevaluation befragt. Da die Zuordnung der Studierenden zu den Lehrenden nicht randomisiert werden konnte, ist das Design auf Studierendenebene quasi-experimentell.

	Jan 10	Feb/ Mrz 10	Jul 10	Sept/ Okt 10	Jan 11
EG 1	O	X	O		O
EG 2	O		O	X	O

Abb. 3: Design der Studie von Ulrich (2013): Messzeitpunkte Studierende (O) und Lehrende (O+X)

Die untersuchten Variablen decken verschiedene Ebenen von Kirkpatrick und Kirkpatrick (2006) ab. Unmittelbar nach Abschluss der Intervention wurde diese von den Lehrenden bewertet (Ebene Reaktion). Zu allen Messzeitpunkten wurde der Lernstand der Lehrenden (Ebene Lernen) erhoben. In den Semestern nach der Intervention wurden zugleich

die Zielerreichung der im Workshop für die Lehre gesetzten Ziele erfragt (Ebene Verhalten), z.B. zur Aktivierung der Studierenden durch Einsatz neuer Lehrmethoden. Jeweils zum Semesterende wurden alle Lehrveranstaltungen evaluiert. Diese Ergebnisse werden sowohl für die Lehrenden herangezogen (Ebene Ergebnisse) als auch für die Studierenden. Die Studierenden bewerten in der Lehrevaluation die Lehrveranstaltung (Ebene Reaktion) und schätzen ihren eigenen Kompetenzerwerb ein (Ebene Lernen). Tabelle 4 gibt einen Überblick:

Tab. 4: Untersuchte Aspekte bei Ulrich (2013), differenziert nach den Ebenen von Kirkpatrick und Kirkpatrick (2006)

Ebene	Zielgruppen	
	Lehrende	Studierende
<b>Reaktion</b>	Bewertung der hochschuldidaktischen Qualifikation	Bewertung der Lehrveranstaltung (Struktur, Professionelle Beziehungsgestaltung)
<b>Lernen</b>	Wissen (über gute Lehre), Fähigkeiten (Selbstwirksamkeit, Fähigkeitsselbstkonzept als Lehrender)	Lernerfolg in der Lehrveranstaltung (Kompetenzeinschätzung: Fach-, Methoden- & Personalkompetenz)
<b>Verhalten</b>	Hochschuldidaktische Performanz in der Lehre (Zielerreichung der im Workshop für die Lehre gesetzten Ziele)	--
<b>Ergebnisse</b>	Lehrevaluationsergebnisse (Details siehe Studierende)	--

Insgesamt führte der als gut bewertete hochschuldidaktische Workshop zu höherem Wissen und besseren Fähigkeiten der Lehrenden. Zugleich setzten sich und erreichten die Lehrenden didaktische Ziele in ihrer Lehre. Abschließend erhielten Lehrende mit höherem Wissen und besseren Fähigkeiten auch bessere Lehrevaluationsergebnisse durch ihre Studierenden.

Die Dissertation von Ulrich (2013) konnte aufzeigen, dass eine hochschuldidaktische Weiterbildung, welche nach dem aktuellen Stand der Hochschuldidaktikforschung konzipiert wurde, bei Nachwuchslehrenden zu nachhaltigen höheren Kompetenzen, veränderten Lehrverhalten und besseren Lehrevaluationsergebnissen führt. Sie entkräftet damit Vorwürfe, hochschuldidaktische Weiterbildung habe keinen positiven Einfluss auf die Lehre. Zugleich zeigt sie eine systematische Validierungsmöglichkeit hochschuldidaktischer Weiterbildung. Einschränkend ist festzuhalten, dass die Aussagen nur auf freiwillig teilnehmende Nachwuchslehrende generalisierbar sind und in der Lehrendenebene auf Selbsteinschätzungen basieren. Des Weiteren bleibt noch unklar, ob die Weiterbildung auf Studierendenebene auch objektive Leistungsindikatoren (Noten, Performanztests etc.) verbessert und wie sich die Nachwuchslehrenden in der Lehre nach Abschluss der Studie entwickelt haben.

#### 4.4 Interview: Struktur akademischer Lehrkompetenz und lehrbezogene Entwicklungsprozesse

In der kumulativen Dissertation von Trautwein (2014) werden drei Forschungsfragen untersucht. Zur Frage 1 erarbeitet die Autorin zunächst ein Strukturmodell akademischer Lehrkompetenz. Anschließend werden in Frage 2 lehrbezogene Entwicklungsprozesse rekonstruiert, um veränderungssensitive Bereiche und Auslöser der Veränderung zu identifizieren. Zuletzt wurde in Frage 3 die Veränderung von Lehr- und Lernüberzeugungen durch hochschuldidaktische Weiterbildung in den Blick genommen.

An der Studie nahmen acht Studiengangsteilnehmende des Masters of Higher Education der Universität Hamburg teil. Diese wurden in 70-113 minütigen episodischen Interviews ausführlich befragt. Ergänzend zu den Interviews wurden für Forschungsfrage 2 zusätzlich die Lehrportfolios analysiert. Die Dokumentenanalyse der Lehrportfolios war nur bei zwei Teilnehmenden möglich, da nur diese es regelmäßig führten. In der Forschungsfrage 3 wurde ergänzend die Videografie der Microteaching Workshops herangezogen. Die Microteaching-Aufnahme diente u.a. dazu, die Lehrenden über die Entwicklung in der Lehre in den Interviews reflektieren zu lassen.

Um Veränderungen zu messen, wurden mehrere Messzeitpunkte realisiert (vgl. Abbildung 3). Es handelt sich somit um eine Längsschnittstudie im Feld mit einer Untersuchungsgruppe.

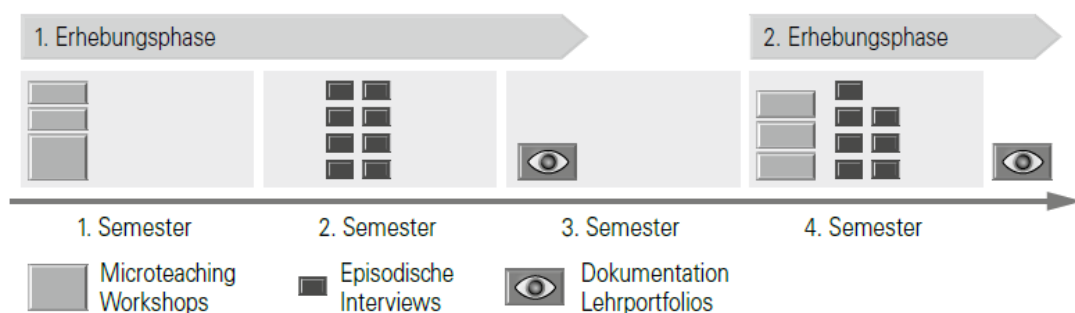


Abb. 4: Messzeitpunkte der Studie (aus Trautwein & Merkt, 2013, S. 56 und Trautwein, 2014, S. 27)

Die Interviews beinhalteten erzählgenerierende Fragen des Interviewers, welche auf das episodische Wissen der Studienteilnehmer(innen) abzielten, z.B. „Ich interessiere mich für deine Erfahrungen in der Lehre. Wenn du einmal zurück blickst auf deine ersten Lehrerfahrungen, wie waren die und wie ging es dann weiter für dich in der Lehre?“ (Trautwein & Merkt, 2013, S. 55). Ergänzend wurden Fragen gestellt, welche semantisches Wissen erfragen, z.B. „Was, würdest Du sagen, kennzeichnet gute Lehre?“ (Trautwein & Merkt, 2013, S. 55). Anschließend sollten die Befragten die Planung und Durchführung einer für sie typischen Lehrveranstaltung erläutern, Höhe- und Tiefpunkte in der eigenen Lehre benennen und persönliche Veränderungen als Lehrperson sowie ihrer Lehrhandlungen beschreiben.

Die Interviews wurden transkribiert und mit Hilfe der Grounded Theory ausgewertet (vgl. Hussy, Schreier & Echterhoff, 2013; Glaser & Strauss, 2010). Das zunächst theoretisch

erarbeitete Strukturmodell akademischer Lehrkompetenz wurde mit den Interviews abgeglichen und datenbasiert modifiziert (vgl. Abbildung 4).

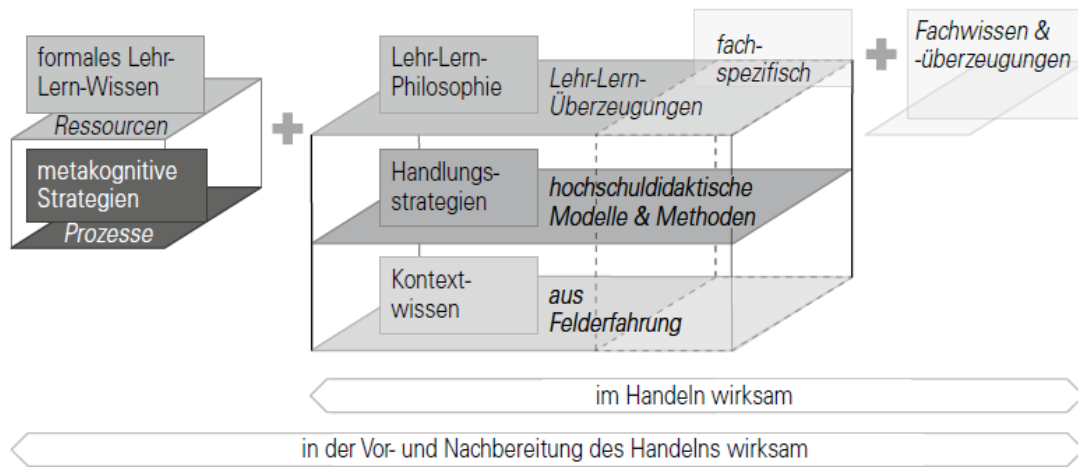


Abb. 5: Struktur akademischer Lehrkompetenz (aus Trautwein & Merkt, 2013, S. 59 und Trautwein, 2014, S. 87)

Zu lehrebezogenen Entwicklungsprozessen wurden zunächst relevante Bereiche erfasst (z.B. „Interaktion mit Studierenden“), folgend Auslöser der Entwicklungsprozesse identifiziert (z.B. „Lehrerfahrungen“) und abschließend Zusammenhänge innerhalb und zwischen beiden Bereichen in den Blick genommen.

Als dritte Forschungsfrage wurden die Veränderung von Lehr- und Lernüberzeugungen durch hochschuldidaktische Weiterbildung zwischen beiden Messzeitpunkten in einer Einzelfallanalyse untersucht (vgl. Trautwein, 2013).

Insgesamt kann man die Forschungsfragen grob der Taxonomie von Donabedian (1966) zuordnen (vgl. Tabelle 5). Forschungsfrage 1 differenziert die hochschuldidaktische Qualifikation der Lehrenden aus, was in Donabedians (1966) Taxonomie der Ebene der Struktur zuzuordnen ist. Die Forschungsfrage 2 zu lehrebezogenen Entwicklungsprozessen ist der Ebene der Prozesse zuordnungsbar, während die Forschungsfrage 3 zu Veränderungen von Lehr- und Lernüberzeugungen durch hochschuldidaktische Weiterbildung das Ergebnis dieser darstellt. Forschungsfrage 3 nutzt zur ihrer Beantwortung die Beschreibung von Lehrprozessen, welche sich aber im Ergebnis, hier einer stärkeren Studierendenorientierung, geändert haben.

Tab. 5: Untersuchte Aspekte bei Trautwein (2014), differenziert nach Donabedians (1966) Ebenen

Bereich nach Donabedian (1966)	Forschungsdesiderate
Prozess	Lehrebezogene Entwicklungsprozesse (Forschungsfrage 2)
Struktur	Strukturmodell akademischer Lehrkompetenz (Forschungsfrage 1)
Ergebnis	Veränderung von Lehr- und Lernüberzeugungen durch hochschuldidaktische Weiterbildung (Forschungsfrage 3)

Trautweins (2014) Dissertation konnte die Struktur akademischer Lehrkompetenz und deren Entwicklung durch hochschuldidaktische Weiterbildung und Lehrpraxis aufdecken und validieren. Sie zeigt im Detail auf, wie hochschuldidaktische Weiterbildung bei den Nachwuchslehrenden reflexiv wirkt und deren Lehrpraxis positiv verändert. Sie unterstreicht damit die Bedeutung reflexiver Nachbereitung hochschuldidaktischer Weiterbildungen durch die Lehrenden. Einschränkend ist festzuhalten, dass die Aussagen nur auf freiwillig teilnehmende Nachwuchslehrende generalisierbar sind und auf Selbstberichten beruhen. Unklar ist noch, wie die Änderungen in der Lehrpraxis auf die Studierenden wirken.

#### **4.5 Testverfahren: Hochschuldidaktische Tests für Lehrende und studentische Leistungsprüfung**

Hochschuldidaktische Studien mit Testverfahren können diese bei Lehrenden und/oder Studierenden einsetzen. Bei Lehrenden sind hochschuldidaktische Wissenstests vor und nach hochschuldidaktischen Weiterbildungen naheliegend, bei Studierenden bietet sich die meist ohnehin durch Prüfungsordnungen vorgegebene Leistungsprüfung an. Alternativ bestehen in einigen Fächern auch standardisierte Leistungstests für Studierende (z.B. Bildungswissenschaftlicher Wissenstest für Lehramtsstudierende und -absolvent(inn)en, vgl. BilWiss, 2016).

Zum aktuellen Zeitpunkt sind noch keine Studien publiziert, in denen hochschuldidaktische Testverfahren für Lehrende eingesetzt wurden. Eine Ursache liegt darin, dass (hochschuldidaktische) Weiterbildungen üblicherweise ohne abschließende Leistungsprüfungen stattfinden, was aufgrund der Freiwilligkeit der meisten Weiterbildungen nur schwer bei den Teilnehmenden durchsetzbar wäre.

Studentische Leistungsprüfungen finden hingegen umfangreich statt, und zu der Auswirkung hochschuldidaktisch relevanter Variablen (z.B. Einsatz aktivierender Lehrmethoden) auf die studentische Leistungsprüfung liegen umfangreiche Studien vor (z.B. Hattie, 2015). Exemplarisch soll folgend eine Studie von Deslauriers et al. (2011) vorgestellt werden.

Deslauriers et al. (2011) haben Studierende der Ingenieurwissenschaft in einer Physiklehrveranstaltung des zweiten Semesters in (letztendlich) zwei Untersuchungsgruppen (N = 538) randomisiert. In der zwölften Woche der Lehrveranstaltung wurde in der Experimentalgruppe die Lehrmethode geändert: Während zuvor in beiden Gruppen eine klassische Vorlesung (v.a. Lehrendenvortrag und Studierende rezipieren Inhalte passiv) gehalten wurde, fand nun eine interaktive Vorlesung statt.

Die interaktive Vorlesung nutzte, im mehrfachen Wechsel während des Termins, die Lehrmethoden „Audio Response System“, auch Clicker-System genannt, mit studentischer Diskussion, Feedback seitens des Lehrenden sowie studentische Kleingruppenarbeit.

Zur Messung wurden in den Wochen 10-12 die Studierenden von vier trainierten Beobachter(inne)n bzgl. des studentischen Engagements während der Stunde bewertet, zugleich wurde die Studierenden gezählt (Anwesenheitsquote). Zum Ende der 12. Woche mussten beide Gruppen eine Leistungsprüfung mit 12 Multiple-Choice-Aufgabe bestehen,

wobei das Skript der Stunde 12 in der Experimentalgruppe beiden Gruppen zum Üben bereitgestellt wurde. Die Studie nutzt demnach ein experimentelles Längsschnittdesign im Feld.

Insgesamt hatten die Studierenden der Experimentalgruppe im Leistungstest eine doppelt so hohe Punktzahl wie die Studierenden in der Kontrollgruppe. Hierbei ist zu bemerken, dass der Lehrende der Kontrollgruppenveranstaltung über höhere Lehrerfahrung und bessere Lehrevaluationen als die Lehrenden der Experimentalgruppe verfügte. Ähnlich stiegen das studentische Engagement und die studentische Anwesenheit in der Experimentalgruppe an, welches zu Beginn der Messung in beiden Gruppen identisch war. Von den 538 Studierenden nahmen aber nur 382 am Leistungstest teil. Zugleich ist „nur“ der kurzfristige Lernerfolg gemessen worden.

Die Studie von Deslauriers et al. (2011) zeigt, dass ein gutes hochschuldidaktisches Lehrkonzept zu höheren studentischen Lernerfolgen führt. Zugleich weist sie nach, dass die Lehrerfahrung des Dozierenden für den studentischen Lernerfolg eine eher untergeordnete Rolle spielt. Angesichts teils hoher Durchfallquoten in einigen Fächern mag sich der höhere zeitliche Aufwand eines guten hochschuldidaktischen Lehrkonzeptes durch die Zeiteinsparung aufgrund der geringeren Anzahl an Wiederholungsklausuren „rechnen“, wodurch auch „hochschuldidaktikferne“ Lehrende besser erreicht werden können. Zu klären ist aber die Frage, ob der höhere Lernerfolg nicht nur bei der hochschuldidaktischen Änderung einzelner, sondern aller Vorlesungstermine persistiert und auch in der Abschlussklausur nachweisbar bleibt.

## 5 Fazit

Die Breite forschungsmethodischer Möglichkeiten für gute hochschuldidaktische Studien ist aufgrund der Interdisziplinarität des Feldes sehr groß. Die vorgestellten Messebenen, Designs und Erhebungsmethoden sowie Forschungsbeispiele sollen Anregungen für die Konzeption neuer Studien geben. Aufgrund der vielen Möglichkeiten in dem immer noch zu wenig beforschten Feld der Hochschuldidaktik können Forschende mit ihren Studien oftmals wissenschaftliches Neuland betreten.

Potenziale bieten sich hierbei in der Triangulation unterschiedlicher Forschungsparadigma und -methoden: diese sind alle mit bestimmten Vorteilen verbunden und gewissen Restriktionen unterworfen, welche sich wechselseitig ergänzen (Vorteile) bzw. teilweise aufheben (Restriktionen). Das Ziel besteht dabei nicht in einem Einsatz möglichst vieler, sondern im Sinne der Forschungsfragestellung passender Methoden (vgl. auch Kap. 3.3), um ein „umfassenderes Bild des Gegenstandsbereiches“ (Kelle, 2008, S. 50) zu erhalten.

Exemplarisch hätten die Studien von Trautwein (Kap 4.4) und Ulrich (Kap. 4.3) auch vereint und aufeinander abgestimmt durchgeführt werden können (wenn sie zeitgleich an denselben Universitäten in einem gemeinsamen mixed-method Forschungsprogramm stattgefunden hätten): Beide Studien untersuchen die Wirkung hochschuldidaktischer Weiterbildung auf Lehrende. Trautweins (2014) qualitative Dissertation kann sehr detailreich die Lehrkompetenzentwicklung der Lehrenden durch die Weiterbildung darlegen, während Ulrich (2013) „nur“ Selbsteinschätzungen bei Lehrenden per Fragebogen abfragt. Ulrichs (2013) quantitative Dissertation umfasst dafür auch die Perspektive der Stu-

dierenden in der Lehrevaluation. Die Kombination beider Studien in derselben Lehrendenstichprobe hätte ggf. weitere Erkenntnisse ermöglicht (was beiden Promovenden in ihrem konkreten Fall nicht möglich war), z.B. welche Facetten der Lehrkompetenz im Detail für die besseren Lehrevaluationen verantwortlich waren.

Wenn die Maßnahmen des Qualitätspaktes Lehre (vgl. Kap. 1) in einem hochschuldidaktischen Forschungsprogramm qualitativ und quantitativ beforscht würden, wie könnte dieses aussehen? Gehen wir von einem Weiterbildungs- und Beratungsprogramm für Nachwuchslehrende aus, welches gute und neue Lehr-/Lernkonzepte umsetzen soll, und zwei Jahre beforscht wird. Das Forschungsprogramm würde die Mikro-Ebene in den Blick nehmen, da Nachwuchslehrende wenig Einfluss auf die Ausgestaltung der Studiengänge (Meso-Ebene) und der Hochschulpolitik (Makro-Ebene) haben. Da neue Nachwuchslehrenden häufig Interesse an den Weiterbildungen bekunden, könnten diese randomisiert einer früher qualifizierten Gruppe (1. Semester) und einer später qualifizierten Gruppe (2. Semester) zugeordnet werden, was ein experimentelles Design ermöglicht: Die später qualifizierte Gruppe würde im 1. Semester als Warte-Kontrollgruppe fungieren. Im folgenden Jahr würde man mit den dann neuen Nachwuchslehrenden genauso vorgehen, so dass mehrere Kohorten entstünden. Die erste Kohorte könnte mindestens zwei Jahre verfolgt und jedes Semester einmal befragt werden.

Es würden Lehrende *und* ihre Studierenden befragt, wobei randomisiert ausgewählte Personen beider Gruppen zusätzlich qualitativ im Detail befragt werden könnten.

Auf der Ebene der Strukturdimensionen (vgl. Abb. 1) würden die Kompetenzen der Lehrenden (qualitativ per Interview, quantitativ per Fragebogen) sowie Charakteristika der Lehrveranstaltung (z.B. Anzahl der Studierenden) erfasst. Auf Ebene der Prozessdimensionen könnten Lehrprozesse und die individuellen Lernaktivitäten (qualitativ per Beobachtung und Interview, quantitativ per Lehrevaluationsfragebogen) erfasst werden. Die eingesetzten Medien der Lehrenden könnten einer qualitativen Dokumentenanalyse bzgl. didaktischer Aufbereitung unterzogen werden. Auf Ebene der Ergebnisse würden die Kompetenzen der Studierenden erhoben (qualitativ per Interview, quantitativ per Lehrevaluationsfragebogen und Klausurnote).

Parallel würden die Nachwuchslehrenden, welche nicht am Weiterbildungsprogramm teilnahmen, befragt und mit den Teilnehmenden verglichen. Deren Einverständnis vorausgesetzt, würden auch deren Studierende mit untersucht.

Die Wirkung der Weiterbildung (vgl. Kirkpatrick & Kirkpatrick, 2006) würde auf diese Weise durch die Zufriedenheit der Teilnehmenden, die Lehrkompetenzen der Lehrenden, das Lehrendenverhalten sowie die Klausurergebnisse und Lehrevaluationen erhoben.

Somit wäre hier ein hochschuldidaktisches mixed-method Forschungsprogramm gegeben, welches seine Ergebnisse je Kohorte in die Hochschule zurückgeben könnte. Dank seiner vielen Erhebungsmethoden kann es die Feinheiten und Zusammenhänge der jeweiligen Faktoren (z.B. Lehrmethode X auf studentischen Lernerfolg im Fach Y) aufzeigen. Es würde damit der Lehr- und Hochschulentwicklung dienen und dadurch auch seiner eigenen Verstetigung helfen. Der Personalaufwand für die Realisierung eines solchen Projektes ist jedoch als hoch einzustufen.

Einschränkend muss festgehalten werden, dass sich der vorliegende Artikel wie auch das eben dargestellte potentielle Forschungsprogramm auf die klassische Individualperspektive fokussiert und systemische Aspekte bei der Auswahl von Forschungsbeispielen größtenteils außen vor gelassen werden. Dies lag zum einem an dem pädagogisch-psychologischen Forschungshintergrund der Autor(inn)en, der sich auf die Individualperspektive der Lehrenden und Studierenden konzentriert. Zum anderen lag dies an dem eher geringen Einfluss von Rahmenbedingungen auf den studentischen Lernerfolg (Hattie, 2015), welchen die Autor(inn)en als ein zentrales Kriterium hochschuldidaktischer Forschung ansehen. Dennoch sind Forschungsstudien zum Einfluss von Systemfaktoren (z.B. Studiendauer in Hochschule vs. Universität) ebenso wichtige Felder hochschuldidaktischer Forschung, und die Triangulation auch dieser Aspekte stellt eine weitere, anschließende Herausforderung dar.

Abschließend ist zu erwähnen, dass sich auch Replikationen bestehender Studien lohnen, ggf. ergänzt durch weitere Forschungsparadigma und -methoden: Lehre und Lernen an Hochschulen unterliegt einem stetigen Wandel – so nimmt z.B. die Bedeutung der Methodenkompetenz der Absolvent(inn)en aufgrund des heutigen einfachen Zugangs zu Informationen zu, während die Bedeutung der Fachkompetenz aufgrund der geringeren Halbwertszeit von Wissen stagniert. Zugleich steigen Qualität und Quantität der Auswertungsmethoden aufgrund der sich erweiternden IT-Möglichkeiten. Daher sind auch publizierte hochschuldidaktische Studien, gerade bei älteren Publikationsdaten, kritisch hinsichtlich der vorgestellten Designs und Ergebnisse zu reflektieren und nach Möglichkeit mit eigenen, aktuellen Studien zu prüfen und je nach Datenlage zu widerlegen oder zu modifizieren.



## Literatur

- Battaglia, S. (2010). Quo vadis, hochschuldidaktische(Hochschul)Forschung? *Journal Hochschuldidaktik*, 21(1), 28–32. Abgerufen von: [http://www.zhb.tu-dortmund.de/hd/fileadmin/JournalHD/2010\\_1/2010\\_1\\_Battaglia.pdf](http://www.zhb.tu-dortmund.de/hd/fileadmin/JournalHD/2010_1/2010_1_Battaglia.pdf)
- Bates, R. (2004). A critical analysis of evaluation practice. The Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27, 341–347.
- Berendt, B. (2005). „Academic Staff Development“ im Kontext und zur Unterstützung des Bologna-Prozesses. Stellenwert und Stand hochschuldidaktischer Weiterbildung. In B. Berendt, B. Szczyrba & J. Wildt (Hrsg.). *Neues Handbuch Hochschullehre* (Rn. L 2.2.). Berlin: Raabe.
- BilWiss. (2016). *Bildungswissenschaftliches Wissen als Teil professioneller Kompetenz in der Lehramtsausbildung*. Abgerufen von: <http://www.bilwiss.uni-frankfurt.de>
- Bos, W., & Tarnai, C. (1999). Content analysis in empirical social research. *International Journal of Educational Research*, 21, 659–671.
- Braun, E., Weiß, T., & Seidel, T. (2014). Lernumwelten in der Hochschule. In T. Seidel & A. Krapp (Hrsg.). *Pädagogische Psychologie* (6. Aufl., S. 433–454). Weinheim: Beltz.
- Bundesministerium für Bildung und Forschung. (2013). Bekanntmachung des Bundesministeriums für Bildung und Forschung von Richtlinien zur Förderung von Vorhaben der Begleitforschung zum Qualitätspakt Lehre. Abgerufen von: <https://www.bmbf.de/foerderungen/bekanntmachung.php?B=893>
- Bundesministerium für Bildung und Forschung. (2015). *Begleitforschung zum Qualitätspakt Lehre*. Abgerufen von: <http://www.hochschulforschung-bmbf.de/de/1622.php>
- Bundesministerium für Bildung und Forschung. (2016). *Qualitätspakt Lehre*. Abgerufen von: <http://www.qualitaetspakt-lehre.de/>
- Centeno García, A. (2016). *Textarbeit in der geisteswissenschaftlichen Lehre* (Dissertation, Martin-Luther-Universität Halle-Wittenberg). Berlin: Frank & Timme.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Hillsdale: L. Erlbaum Associates.
- Dany, S. (2007). *Start in die Lehre. Qualifizierung von Lehrenden für den Hochschulalltag* (Bildung – Hochschule – Innovation, Bd. 4). Berlin: LIT.
- Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, 332(6031), 862–864.
- Deutsche Gesellschaft für Hochschuldidaktik. (2016). *Nachwuchspreis hochschuldidaktische Forschung*. Abgerufen von: <http://www.dghd.de/nachwuchspreis-hochschuldidaktische-forschung.html>
- Diekmann, A. (2009). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen* (20. Aufl.). Reinbek: Rowolth.
- Donabedian, A. (1966). Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly*, 44(3), 166–203.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Berlin: Springer.

- Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., & DuBois, D. (2008). Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of vocational behavior*, 72(2), 254–267.
- Field, A. (2013). *Discovering Statistics using SPSS* (4. Aufl.). London: Sage.
- Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, 75(1), 27–61.
- Glaser, B. G., & Strauss, A. (2010). *Grounded Theory: Strategien qualitativer Forschung* (3. Aufl.). Bern: Huber.
- Hattie, J. A. (2009). Visible learning. A synthesis of over 800 meta-analyses relating to achievement. London: Routledge.
- Hattie, J. A. (2015). The applicability of visible learning to higher education. *Scholarship of Teaching and Learning in Psychology*, 1(1), 79–91.
- Heckmann, C. (eingereicht). *Lehrkompetenzentwicklung in der Hochschullehre videoanalytisch abbilden* (Unveröffentlichte Dissertation, Goethe-Universität Frankfurt).
- Helmke, A. (2009). Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze: Klett-Kallmeyer.
- Hugener, I., Pauli, C., & Reusser, K. (2006). Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis". 3. Videoanalysen (Materialien zur Bildungsforschung, Band 15). Frankfurt am Main: GFFP.
- Hussy, W., Schreier, M., & Echterhoff, G. (2013). *Forschungsmethoden in Psychologie und Sozialwissenschaften* (2. Aufl.). Berlin: Springer.
- Johannes, C., Fendler, J., Hoppert, A., & Seidel, T. (2011). *Projekt LehreLernen (2008 – 2010): Dokumentation der Erhebungsinstrumente*. Münster: Monsenstein und Vannerdat.
- Johannes, C., & Seidel, T. (2010). Professionelles Lernen von Anfängern in der Hochschullehre – Erwartungen und Vorstellungen über Hochschullehre im Rahmen des Projektes LehreLernen. *Personal- und Organisationsentwicklung in Einrichtungen der Lehre und Forschung*, 5(2+3), 31–40.
- Johannes, C., & Seidel, T. (2012). Professionalisierung von Hochschullehrenden. *Zeitschrift für Erziehungswissenschaft*, 15(2), 233–251.
- Johnson, R. N. (2000). *A qualitative study of student feedback: Lecturers' and students' perception experiences*. Dissertation, Sheffield Hallam University. Abgerufen von: <http://shura.shu.ac.uk/3158/>
- Kauffeld, S. (2016). *Nachhaltige Personalentwicklung und Weiterbildung. Betriebliche Seminare und Trainings entwickeln, Erfolge messen, Transfer sichern* (2., überarbeitete Auflage). Berlin: Springer.
- Kelle, U. (2008). *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung: Theoretische Grundlagen und methodologische Konzepte* (2. Aufl.). Wiesbaden: VS.
- Kelle, U. (2014). Forschungsdesigns für die quantitative Sozialforschung. In N. Baur & J. Blasius (Hrsg.). *Handbuch Methoden der empirischen Sozialforschung* (S. 153–166). Wiesbaden: Springer VS.

- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: the four levels* (3rd ed.). San Francisco: Berrett-Koehler.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht*. Münster: Waxman.
- Mayring, P. (2002). *Einführung in die qualitative Sozialforschung* (5. Aufl.). Weinheim: Beltz.
- Merkt, M. (2014). Hochschuldidaktik und Hochschulforschung. Eine Annäherung über Schnittmengen. *die hochschule. journal für wissenschaft und bildung*, 13(1), 92–105.
- Oser, F., Curcio, G. P., & Düggeli, A. (2007). Kompetenzmessung in der Lehrerbildung als Notwendigkeit – Fragen und Zugänge. *Beiträge zur Lehrerbildung*, 25(1), 14–26.
- Przyborski, A., & Wohlrab-Sahr, M. (2014). Forschungsdesigns für die qualitative Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 117–133). Wiesbaden: Springer VS.
- Reese, R. J., Prout, H. T., Zirkelback, E. H., & Anderson, C. R. (2010). Effectiveness of schoolbased psychotherapy: A meta-analysis of dissertation research. *Psychology in the Schools*, 47(10), 1035–1045.
- Schaeper, H., & Weiß, T. (2016). The Conceptualization, Development, and Validation of an Instrument for Measuring the Formal Learning Environment in Higher Education. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Hrsg.). *Methodological Issues of Longitudinal Surveys. The Example of the National Educational Panel Study* (S. 267–290). Wiesbaden: Springer.
- Schaper, N. (2014a). Aus- und Weiterbildung: Konzepte der Trainingsforschung. In F. Nerdinger, G. Blickle & N. Schaper (Hrsg.), *Arbeits- und Organisationspsychologie* (3. Aufl., S. 461–488). Berlin: Springer.
- Schaper, N. (2014b). Forschung in der Hochschulbildung. In J. Kohler, P. Pohlenz & U. Schmidt (Hrsg.). *Handbuch Qualität in Studium und Lehre* (Bd. 48, S. 69–96, Rn. D 2.4-1). Stuttgart: Raabe.
- Schmidt, U., Heinzemann, S., Schulze, K., Walkemeyer, J., & Faaß, M. (2015). Evaluation des Bund-Länder-Programms für bessere Studienbedingungen und mehr Qualität in der Lehre (Qualitätspakt Lehre) – Zweiter Zwischenbericht –. Bonn: Bundesministerium für Bildung und Forschung. Abgerufen von: [http://www.qualitaetspakt-lehre.de/\\_media/Zweiter%20Zwischenbericht.pdf](http://www.qualitaetspakt-lehre.de/_media/Zweiter%20Zwischenbericht.pdf)
- Seidel, T., & Hoppert, A. (2011). Merkmale von Lehre an der Hochschule. Ergebnisse zur Gestaltung von Hochschulseminaren mittels Videoanalysen. *Unterrichtswissenschaft*, 39(2), 154–172.
- Seidel, T., & Prenzel, M. (2006). Stability of teaching patterns in physics instructions: findings from a video study. *Learning and Instruction*, 16(3), 228–240.
- Seidel, T., Prenzel, M., Duit, R., & Lehrke, M. (2003). Technischer Bericht zur Videostudie Lehr-Lern-Prozesse im Physikunterricht. Kiel: IPN.
- Stein, P. (2014). Forschungsdesigns für die quantitative Sozialforschung. In N. Baur & J. Blasius (Hrsg.). *Handbuch Methoden der empirischen Sozialforschung* (S. 135–151). Wiesbaden: Springer VS.
- Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order metaanalysis and validation study. *Review of Educational Research*, 81(1), 4–28.

- Trautwein, C. (2013). Lauft es schief oder lauft es von selbst?! – Abgabe von Kontrolle im Wandel von einer lehr- zu einer lernzentrierten Lehre. In J. Wildt & M. Heiner (Hrsg.). *Professionalisierung der Lehre. Perspektiven formeller und informeller Entwicklung von Lehrkompetenz im Kontext der Hochschulbildung* (S. 139–149). Blickpunkt Hochschuldidaktik (Bd. 123). Bielefeldt: Bertelsmann.
- Trautwein, C. (2014). *Struktur akademischer Lehrkompetenz und lehrebezogene Entwicklungsprozesse* (Unveroffentlichte Dissertation, Albert-Ludwig-Universitat Freiburg).
- Trautwein, C., & Merkt, M. (2013). Akademische Lehrkompetenz und Entwicklungsprozesse Lehrender. *Beitrage zur Hochschulforschung*, 35(3), 50–77. Abgerufen von: <http://www.bzh.bayern.de/uploads/media/3-2013-Trautwein-Merkt.pdf>
- Ulrich, I. (2013). *Strategisches Qualitatsmanagement in der Hochschullehre: Theoriegeleitete Workshops fur Lehrende zur Forderung einer kompetenzorientierten Lehre* (Dissertation, Freie Universitat Berlin). Wiesbaden: Springer VS.
- Ulrich, I. (2016). *Gute Lehre in der Hochschule. Praxistipps zur Planung und Gestaltung von Lehrveranstaltungen*. Berlin: Springer.
- Ulrich, I. & Heckmann, C. (2013). Wirksamkeitsmessung von Hochschuldidaktik: Messmoglichkeiten und Anwendungsbeispiele hochschuldidaktischer Wirksamkeitsmessung. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba & J. Wildt (Hrsg.). *Neues Handbuch Hochschullehre* (Rn. I 4.4). Berlin: Raabe.
- Wibbeke, G. (2015). *Evaluation einer hochschuldidaktischen Weiterbildung an der Medizinischen Fakultat Heidelberg* (Unveroffentlichte Dissertation, Ruprecht-Karls-Universitat Heidelberg).
- Wildt, J., & Jahnke, I. (2010). Konturen und Strukturen hochschuldidaktischer Hochschulforschung – ein Rahmenmodell. *Journal Hochschuldidaktik*, 1, 4–8.
- Wolbring, T. (2013). *Fallstricke der Lehrevaluation. Moglichkeiten und Grenzen der Messbarkeit von Lehrqualitat*. Frankfurt am Main/New York: Campus.
- Wolbring, T. (2014). Wie valide sind studentische Lehrveranstaltungsbewertungen? Sachfremde Einflusse, studentische Urteilerstandards, Selektionseffekte. *Qualitat in der Wissenschaft*, 8(2+3), 56 –60.

## Autor/-innen

Dr. Immanuel Ulrich, Goethe-Universitat Frankfurt, Institut fur Psychologie, Arbeitseinheit Lehren und Lernen im Erwachsenenalter, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main; Email: [ulrich@psych.uni-frankfurt.de](mailto:ulrich@psych.uni-frankfurt.de)

Dipl.-Psych. Carmen Heckmann, Goethe-Universitat Frankfurt, Institut fur Psychologie, Arbeitseinheit Lehren und Lernen im Erwachsenenalter, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main; Email: [heckmann@paed.psych.uni-frankfurt.de](mailto:heckmann@paed.psych.uni-frankfurt.de)



**Zitiervorschlag:** Ulrich, I. & Heckmann, C. (2017). Taxonomien hochschuldidaktischer Designs und Methoden aus pädagogisch-psychologischer Sicht samt Musterbeispielen aus der aktuellen Forschung. *die hochschullehre*, 3. Abgerufen unter: [www.hochschullehre.org](http://www.hochschullehre.org)