

Birgit SPINATH¹ & Eva SEIFRIED (Heidelberg)

Was brauchen wir, um solide empirische Erkenntnisse über gute Hochschullehre zu erhalten?

Zusammenfassung

Aufgrund hochschulpolitischer Entwicklungen der letzten Jahre besteht eine starke Nachfrage nach empirischen Befunden über gute Hochschullehre. Dieser Nachfrage stehen jedoch überraschend wenige gesicherte Befunde gegenüber. In diesem Beitrag soll erörtert werden, warum es so schwierig ist, solide Erkenntnisse über die Qualität von Hochschullehre zu generieren. Es werden sechs Probleme der Forschung über die Qualität von Hochschullehre vorgestellt und diskutiert. Dabei wird auch jeweils dargelegt, wie diesen Problemen begegnet werden könnte und was damit benötigt wird, um gesicherte Erkenntnisse zu generieren.

Schlüsselwörter

Hochschullehre, empirische Bildungsforschung, Probleme, Desiderata

¹ E-Mail: birgit.spinath@psychologie.uni-heidelberg.de



What is needed to gain solid empirical insights into good teaching and learning in higher education?

Abstract

Due to current political developments in the tertiary sector, there is a growing demand for empirical evidence about good teaching and learning in higher education. However, there are surprisingly few solid empirical findings. In this paper, we will discuss problems faced by research in higher education, which can be divided into six key areas: criteria, predictors, specificity, causality, samples, and measurement instruments. We will also explain what is needed to address these problems and thereby generate valid evidence (e.g., definitions of good teaching and learning at universities, well-established specific measurement instruments).

Keywords

Higher education, empirical educational research, problems, desiderata

1 Gesteigerte Nachfrage nach empirischen Befunden über gute Hochschullehre

Die Qualität hochschulischer Bildung spielte lange Zeit sowohl national als auch international eine eher untergeordnete Rolle in der empirischen Bildungsforschung. Deren Fokus lag und liegt stattdessen vor allem auf dem schulischen Bereich. In jüngerer Zeit rückte die Hochschullehre jedoch zunehmend in das Blickfeld der empirischen Bildungsforschung. Dafür sind verschiedene Entwicklungen verantwortlich, wie z. B. die Notwendigkeit zur Akkreditierung von Studienprogrammen inklusive einer systematischen Qualitätssicherung der Hochschullehre, die sich diversifizierende Hochschullandschaft mit zunehmend privaten Anbietern sowie das Bereitstellen von Fördermitteln für Forschung und Entwicklung rund um die Qualität der Hochschullehre.

Jedoch steht der starken Nachfrage nach gut gesicherten Befunden darüber, was gute Hochschullehre ist, ein überraschender Mangel an solchen entgegen. Unter gesicherten Befunden über gute Hochschullehre verstehen wir im Folgenden die Ergebnisse von Studien, die durch systematische Variation von Lehr-Lernbedingungen und bestmöglichem Ausschluss alternativer Erklärungen nachgewiesen haben, dass eine bestimmte Methode einer anderen hinsichtlich des Erzielens zentraler hochschulischer Bildungsergebnisse (allem voran fachlicher und überfachlicher Lernerfolg, aber auch Motivation, Selbstregulationskompetenzen etc.) überlegen ist. Der Goldstandard für diese Zwecke sind Randomized Controlled Trials (RCT). Darüber hinaus leisten selbstverständlich auch andere Forschungsansätze wichtige Beiträge zu der zugrundeliegenden Frage nach guter Hochschullehre. Beispielsweise müssen verlässliche Messinstrumente entwickelt werden, um aus RCT-Studien angemessene Schlüsse ziehen zu können. Auch sind korrelative Designs wichtig, um aufzuzeigen, ob und wo Zusammenhänge vorliegen und damit als Entscheidungsgrundlage dafür, wo Interventionen ansetzen sollten. Schlussendlich kann die Frage danach, was gute Hochschule ausmacht und wie sie weiter verbessert werden kann, empirisch jedoch nur durch experimentelles oder daran angelehntes Vorgehen beantwortet werden.

Um die These zu untermauern, dass es zu der Frage, was gute Hochschullehre ausmacht, wenig empirische Evidenz gibt, greifen wir auf ein aktuelles Review zurück, in dem – wie bei HATTIE (2009) – auf der Ebene von Metaanalysen das verfügbare Wissen über effektives Lehren und Lernen zusammengetragen wurde. SCHNEIDER & PRECKEL (2017) haben 38 Metaanalysen zu 105 Korrelaten von hochschulischen Leistungen zusammengefasst (ca. 3330 Effektstärken, ca. 2 Millionen Lernende). Die 105 untersuchten Variablen sind in elf Kategorien unterteilt (z. B. soziale Interaktionen, Intelligenz und Vorleistung), die wiederum den Bereichen instruktionsbezogene Variablen und lernendenbezogene Variablen zugeordnet sind. Sucht man aus dieser Zusammenstellung diejenigen Befunde heraus, die in unserem Sinne gesicherte Schlüsse über gute Hochschullehre zulassen, so stellt sich heraus, dass dies nur für einen kleinen Teil der betrachteten Variablen zutrifft. Zum einen sind nur Studien über instruktionsbezogene Variablen relevant für die

Frage, wie effektive Hochschullehre aussehen könnte (die Studien zu studentischen Lernvoraussetzungen und deren Lernaktivitäten gehören nicht im engeren Sinne dazu). Diese machen jedoch den kleineren Teil der Studien aus. Zudem lässt ein Mangel an kontrollierten experimentellen Studien häufig keine kausalen Schlüsse über die berichteten Zusammenhänge zu. Solche Ursache-Wirkungs-Befunde wären jedoch für Ableitungen zur Verbesserung der Hochschullehre nötig. Und schließlich sind – wie auch bei HATTIE (2009) – zahlreiche Studien einbezogen worden, die gar nicht untersucht haben, ob sich eine bestimmte Lehr-Lernmethode auf Leistungen auswirkt. So z. B. wurden zahlreiche Studien inkludiert, bei denen es rein um das Feststellen eines korrelativen Zusammenhangs zwischen etwa studentischen Lehrbeurteilungen oder Peer-Beurteilungen von Leistungen mit Leistungen der Studierenden ging. Es muss also festgehalten werden, dass diese Zusammenfassung an Forschungsbefunden eine Fülle an Befunden über effektive Hochschullehre und eine Sicherheit an Wissen suggeriert, die in diesem Maße nicht gegeben sind.

Zusammenfassend lässt sich sagen: Nachdem sich in den letzten Jahren im Bereich des schulischen Lehrens und Lernens der Erkenntnisstand explosionsartig erweitert hat, gewinnt nun zunehmend auch der Hochschulbereich an Bedeutung in der Forschung und an Interesse für die Öffentlichkeit. Doch auch wenn die Nachfrage und auch die Anzahl der Befunde zu guter Lehre im Hochschulbereich steigen, sind bei deren Interpretation einige Aspekte kritisch zu betrachten.

2 Probleme der Forschung über die Qualität von Hochschullehre und mögliche Verbesserungsansätze

Die Tatsache, dass es wenige gut gesicherte Befunde über wirksames Lehren in der Hochschule gibt, lässt sich damit erklären, dass diese Art der Forschung mit mehreren Problemen konfrontiert ist. In diesem Abschnitt werden sechs Probleme der Forschung über die Qualität von Hochschullehre vorgestellt und diskutiert, nämlich

Probleme der Kriterien, Probleme der Prädiktoren, Probleme der Spezifität, Probleme der Kausalität, Probleme der Stichproben sowie Probleme der Messinstrumente.

2.1 Das Kriterien-Problem

Zunächst stellt sich das Kriterien-Problem oder die Frage „Was ist gute Hochschullehre?“. Empirische Studien benötigen als Grundlage eine Definition dessen, was die Ziele der Hochschullehre sind und wie die Zielerreichung in messbaren Kriterien operationalisiert werden kann. Die Ziele der Hochschullehre sind jedoch vielfältig und reichen von der Vermittlung basaler fachlicher Standards über Schlüsselkompetenzen bis hin zur Befähigung zu eigener Forschung. Diese unterschiedlichen Ziele sind nicht immer im Einklang und können zum Teil am besten durch unterschiedliche Lehrmethoden erreicht werden. Studien zur Wirksamkeit guten Lehrens sollten diese Vielfalt möglicher Ziele abbilden und nicht eine Verkürzung auf den reinen Erwerb von Fachwissen vornehmen (vgl. HELMKE & SCHRAIDER, 2010).

Entgegen der obigen Empfehlung konzentrieren sich viele empirische Studien jedoch auf nur eines der Kriterien. So findet man Studien, die eher Voraussetzungen für gute Lehre (z. B. Fach- oder Diagnosekompetenzen) oder gezeigte Lehraspekte fokussieren (z. B. Setzung passender Lernziele und Anforderungen), bis hin zu Studien, die den Lernerfolg von Studierenden betrachten. Auch innerhalb dieser Kriterien finden sich Variationen; zum Beispiel kann der Lernerfolg der Studierenden durch objektive Tests oder aber durch Studierenden-Urteile operationalisiert werden. Im Review von SCHNEIDER & PRECKEL (2017) werden standardisierte Leistungstests, ad hoc konstruierte Tests, von Lehrenden vergebene Noten sowie „andere“ (z. B. Abschlüsse) als abhängige Variablen (gemeinsam) betrachtet.

Theoretisch und auch in empirischen Studien finden sich somit sehr unterschiedliche Kriterien für erfolgreiche Hochschullehre. Dadurch werden eine Vergleichbarkeit der Studien sowie die Erstellung eines Gesamtbildes guter Hochschullehre

erschwert. Was daher nötig ist, sind klare Definitionen guter Hochschullehre. Erst diese ermöglichen eine umfassende Beforschung durch empirische Studien.

Aus empirischer Sicht gibt es eine Reihe von Methoden zur Bestimmung von Kriterien für gute Hochschullehre: theoretische Ableitungen (z. B. Theorien der Arbeitszufriedenheit, Grundsätze der Instruktionspsychologie), Analysen von Lehrzielen (z. B. Dokumentenanalysen von Gesetzestexten kombiniert mit Lehrendenbefragungen), Befragungen von Lehrenden und Lernenden (z. B. „Was ist für Sie gute Lehre?“), Inspektionen von Lehrevaluationsbögen (z. B. Zusammenfassen der sich empirisch zeigenden Dimensionen) oder Analysen von Vorhersagekriterien für wichtige Ziele der Lehre (z. B. Prädiktoren von Lernerfolg, Studierendenzufriedenheit) (vgl. SPINATH & STEHLE, 2011). Allerdings wird sich die Frage, welche Ziele gute Hochschullehre erreichen soll, nicht allein empirisch beantworten lassen. Stattdessen bedarf es dafür auch eines diskursiv herbeigeführten Konsenses, der auch die Gegebenheiten der jeweiligen Fachkulturen zu berücksichtigen hat.

HELMKE & SCHRADER (2010) benennen in ihrer Definition der Ziele einer guten Hochschulausbildung Fachwissen und Expertise, überfachliche kognitive Kompetenzen sowie volitionale, emotionale und soziale Schlüsselqualifikationen. Auch andere Definitionen guter Hochschullehre sind sehr breit angelegt und beinhalten umfassende Bereiche, zum Beispiel die Ausbildung mündiger Bürger/innen (vgl. ULRICH, 2016). Prinzipiell können Kriterien guter Hochschullehre im kognitiven Bereich (z. B. Lernerfolg), im motivationalen Bereich (z. B. Motivation zum Lernen auch langfristig), im affektiven Bereich (z. B. Zufriedenheit der Lehrenden und Studierenden) sowie im behavioralen Bereich angesiedelt werden (z. B. niedrige Abbruchquoten). Ein allgemein anerkanntes und umfassendes Modell guter Hochschullehre würde die Erforschung einzelner Komponenten, aber vor allem auch des Konstrukts als Ganzem ermöglichen.

In vielen empirischen Studien wird jedoch nur ein Kriterium guter Hochschullehre als abhängige Variable untersucht. Häufig ist dies der Lernerfolg der Lernenden (s. a. HATTIE, 2009; SCHNEIDER & PRECKEL, 2017). Wir plädieren dafür, mehrere Kriterien in einer Untersuchung zu berücksichtigen (sowohl dem Inhalt als auch

der Erhebungsmethode nach), um dadurch ein Gesamtbild guter Hochschullehre zu definieren und untersuchen zu können. Erst in einer Zusammenschau der Auswirkungen bestimmter Maßnahmen auf verschiedene Kriterien lassen sich solide Erkenntnisse gewinnen. Dabei sollten Lernzuwachs und Zufriedenheit von Studierenden und Lehrenden als übergeordnete Kriterien guter Lehre gelten. Jedoch sind diese Kriterien nicht nur von der Qualität der Lehre abhängig, sondern von individuellen Voraussetzungen der Studierenden (und Rahmenbedingungen). Auch dieser Tatsache sollte in empirischen Studien Rechnung getragen werden.

2.2 Das Prädiktoren-Problem

Sozusagen auf der anderen Seite der Lehr-Lern-Gleichung, das heißt komplementär zum Kriterien-Problem, steht das Prädiktoren-Problem oder die Frage „Welches sind vielversprechende Prädiktoren guter Hochschullehre?“. An dieser Stelle soll zunächst ein Problem eingeführt werden, das Kriterien und Prädiktoren in gleichem Maße betrifft: Zum Teil werden die relevanten Maße in Studien im Hochschulkontext eher nach praktischen oder ökonomischen Gesichtspunkten gewählt: Häufig werden beispielweise die ohnehin an einer Universität eingesetzten Evaluationsbögen als eine abhängige Variable genutzt. Inwiefern diese Bögen theoriegeleitet entstanden sind und sich für die Beantwortung der spezifischen Forschungsfrage einer bestimmten Studie eignen, bleibt offen bzw. ist fraglich. Auch wenn beispielsweise Selbstauskunfts-Fragebögen zur Messung erworbener Kompetenzen genutzt werden, stehen vermutlich teilweise eher ökonomische als theoretische Gründe im Vordergrund dieser Operationalisierung, da für Aussagen über das eigentlich interessierende Konstrukt (hier: erworbene Kompetenzen) aus theoretischer Sicht andere Maße relevant wären (z. B. objektive Kompetenztests, die theoretisch begründet sind). Unseres Erachtens müssen in empirischen Studien sowohl die Kriterien als auch die Prädiktoren theoriegeleitet operationalisiert werden, um aus den Studien sinnvolle Implikationen ableiten zu können.

Zudem wurden wie bei den Kriterien guter Hochschullehre auch Prädiktoren auf sehr unterschiedlichen Ebenen untersucht. Zum Teil werden überdauernde oder eher (situations-/lehr-)spezifische Merkmale von Dozierenden und Studierenden

betrachtet (z. B. deren Persönlichkeit oder Motivation), zum Teil Unterrichtsmerkmale (z. B. die Redezeiten von Lehrenden und Studierenden, die Art und Häufigkeit von Feedback von Lehrenden an Lernende und v. v.); selten werden beide Ebenen gemeinsam betrachtet (vgl. SCHNEIDER & PRECKEL, 2017). Interessante und vielversprechende Faktoren zur Verbesserung der Lehre sind prinzipiell solche, die nahe an der Lehre und potentiell veränderbar sind. Somit scheint vor allem die Beforschung relevanter Unterrichtsmerkmale aussichtsreich statt etwa die der Persönlichkeitsmerkmale von Lehrenden. Denn für die einzelne handelnde Person sind besonders die Faktoren relevant, auf die sie einen direkten Einfluss hat (z. B. Lehr- und Präsentationsmodi). Dies ist jedoch unter anderem aufgrund der möglichen Bandbreite relevanter Variablen als auch aufgrund deren Spezifität gleichzeitig besonders schwierig. HELMKE & SCHRADER (2010) weisen darauf, dass kurzschlüssige Interpretationen über ‚gute Hochschullehre‘ drohen, wenn nicht berücksichtigt wird, dass die Nutzung und die Wirkungen von Hochschullehre von vielfältigen Faktoren abhängen und daher nicht mit einfachen und direkten Effekten einzelner Prädiktoren zu rechnen ist. Insgesamt sollte die Forschung daher ihr Augenmerk stärker auf instruktionsbezogene Variablen legen, jedoch die Bedeutsamkeit der lernendenbezogenen Variablen nicht unberücksichtigt lassen (z. B. um mögliche Moderations- oder Interaktionseffekte nicht zu übersehen oder um spezifischere Wirkungsaussagen machen zu können).

2.3 Spezifitäts-Problem

Ist die Lehr-Lern-Gleichung mit Kriterien und Prädiktoren prinzipiell aufgestellt, stellt sich das Spezifitäts-Problem oder die Frage „Sind Lehr-Lernprinzipien spezifisch für die Hochschule versus andere Orte und sind sie spezifisch für Fächer, Veranstaltungsformen, Gruppen von Lehrenden oder generell?“. Es darf als sicher gelten, dass Prinzipien, die sich für schulisches Lernen als wirksam erwiesen haben, nicht ohne Weiteres auf den Hochschulkontext übertragbar sind. Gleichmaßen sind die Fachkulturen an Hochschulen so unterschiedlich, dass Befunde in einem Fach nicht unbedingt für ein anderes gelten. Darüber hinaus muss für Vorlesungen nicht gelten, was in Seminaren gilt und für Lehrende (und Lernende) mit

bestimmten Voraussetzungen mag es wiederum anders sein als für andere. Kurzum, am Problem der Spezifität zeigt sich sehr gut, was BERLINER (2002, S. 18) in seinem Kommentar „Educational research: The hardest science of all.“ so formulierte:

In my estimation, we have the hardest-to-do science of them all! We do our science under conditions that physical scientists find intolerable. We face particular problems and must deal with local conditions that limit generalizations and theory building – problems that are different from those faced by the easier-to-do sciences. Let me explain this by using a set of related examples: The power of context, the ubiquity of interactions, and the problem of “decade by findings” interactions.

Unter anderem aufgrund der unterschiedlichen Kontexte bzw. Kontextbedingungen von Hochschullehre, die sich zudem in ständigem Wandel befinden, ist die Beforschung von Hochschullehre schwierig. HELMKE & SCHRADER (2010) verweisen darauf, dass die Ziel- und Kontextabhängigkeit guter Lehre von der Hochschuldidaktik häufig nicht berücksichtigt wird und dies zum schlechten Ruf und der geringen Akzeptanz hochschuldidaktischer Maßnahmen beigetragen haben könnte. Auch hier gilt es (wie auch in der Forschung), mögliche Spezifitäten zu berücksichtigen.

Empirische Studien variieren stark in Hinblick auf Aspekte der Spezifität. Viele Forschungsergebnisse erwachsen aus Studien, die das eigene Lehren beforschen (vgl. Scholarship of Teaching and Learning, SoTL; national: HUBER, 2011; HUBER, PILNIOK, SETHE, SZCZYRBA & VOGEL, 2014; SPINATH & SEIFRIED, 2012 bzw. international: BOYER, 1990; HUTCHINGS, HUBER & CICCONE, 2011). Dieser Ansatz ist prinzipiell wünschenswert, jedoch beziehen sich die meisten Forschungsergebnisse auf einzelne Kurse, Fächer oder Personengruppen. Somit ist unklar, inwiefern die Ergebnisse entsprechender Studien generalisierbar sind. HELMKE & SCHRADER (2010) verweisen darauf, dass sich „[g]ute Lehre [...] angesichts der Einbettung des individuellen Lernprozesses in ein komplexes Geflecht personeller und situativer Faktoren und wegen der Zielabhängig-

keit aller Unterrichtsbemühungen nicht auf einige allgemein gehaltene Präskriptionen reduzieren“ lässt (S. 274f.).

Die empirische Bildungsforschung hat bereits viele Befunde für den schulischen Bereich erbracht. Prinzipiell könnten sich Theorien und Prinzipien aus diesem Bereich auf den hochschulischen Sektor übertragen lassen, denn einige Studien weisen darauf hin, dass sich bestimmte Befundmuster in beiden Bereichen finden (vgl. SCHNEIDER & PRECKEL, 2017 bzw. im Vergleich zu HATTIE, 2009). Allerdings müssen für die Hochschule spezifische Aspekte berücksichtigt werden, so dass auch eigene theoretische Ansätze zu formulieren sind (vgl. folgender Abschnitt). Beispielsweise spielt in der Hochschule im Vergleich zur Schule das selbstregulierte Lernen eine deutlich größere Rolle. Auch sind die Ziele der Hochschule häufig generischer in ihrer Art als diejenigen des schulischen Lernens. Allgemeine Modelle wie das Angebots-Nutzungs-Modell des Unterrichts (HELMKE, 2009) bilden eine erste Grundlage zur Beschreibung der Bedingungen guter Hochschullehre, die jedoch kontextabhängig zu spezifizieren sind (vgl. HELMKE & SCHRADER, 2010).

Auch wenn sich schulische und hochschulische Lehre in manchen Aspekten ähneln und Theorien und Prinzipien ggf. übertragen werden könnten, so gibt es auch wichtige Unterschiede zwischen den beiden Bereichen: SCHNEIDER & PRECKEL (2017) weisen zum Beispiel darauf hin, dass die Lernenden in der Hochschule eine deutlich selektiertere Gruppe als in Schulen darstellen (höhere kognitive Fähigkeiten, bessere Lernstrategien), dass an der Hochschule häufig größere Gruppen unterrichtet werden müssen und dass die Ziele der hochschulischen Bildung andere als die der schulischen Bildung sind (vgl. dazu auch die Definition von HELMKE & SCHRADER, 2010). Diese und weitere Aspekte müssen bei der Entwicklung hochschulspezifischer Wirkungsmodelle berücksichtigt werden. Zwar können allgemeine Erkenntnisse zum Lehren und Lernen bzw. zu Unterrichtsqualität auch der Hochschullehre nutzen, dennoch müssen fachliche und didaktische Aspekte berücksichtigt werden (vgl. HELMKE & SCHRADER, 2010). Das Rahmenmodell zur Wirkungsanalyse des Hochschulunterrichts von HELMKE & SCHRADER (2010) könnte als Grundlage für systematische Forschung dienen.

Um solide empirische Erkenntnisse über gute Hochschullehre zu erhalten, bedarf es unseres Erachtens einer Analyse von Gemeinsamkeiten und Unterschieden zwischen a) Hochschullehre und anderen Lehr-Lernkontexten, b) Fächern, c) unterschiedlichen Gruppen von Lehrenden und Lernenden und d) Veranstaltungsformen. Die Forschung muss zeigen, inwiefern vorgefundene Erkenntnisse spezifisch für die Hochschullehre sind (vs. allgemein für Lehr-Lernkontexte; vgl. auch voriger Abschnitt) und inwiefern bestimmte Erkenntnisse auf andere Fächer oder andere Lehrende und Lernende sowie Veranstaltungsformen übertragbar sind. Die aktuelle Bewegung zum SoTL ist hierfür sicherlich fruchtbar, allerdings müssen auch einige Aspekte berücksichtigt werden, damit SoTL den Standards wissenschaftlicher Forschung genügen kann: Um Ergebnisse generalisieren zu können, muss SoTL zum einen auf Theorien aufbauen, zum anderen müssen alle Phasen und Kontextfaktoren sorgfältig dokumentiert werden, um die Wahrscheinlichkeit einer Replizierbarkeit zu erhöhen. Des Weiteren muss die Generalisierbarkeit von Ergebnissen geprüft werden, indem die Designs in andere Kontexte übertragen werden. Und zuletzt müssen die Ergebnisse publiziert werden. Erst dann können Analyse von Gemeinsamkeiten und Unterschieden vollzogen werden. Daher plädieren wir dafür, die spezifischen Kontextbedingungen zu berücksichtigen und zu berichten, so dass dann eine differenzierte Einordnung einzelner Befunde möglich ist. Wenn verschiedene Studien trotz unterschiedlicher Kontextbedingungen zu gleichen Ergebnissen kommen, kann eine gewisse Generalisierbarkeit dieser Befunde angenommen werden.

2.4 Kausalitäts-Problem

Ein weiteres und für Schlussfolgerungen kritisches Problem ist das Kausalitäts-Problem oder die Frage „Wie können Wirkmechanismen identifiziert werden?“. Obwohl die meisten Studien auf Kausalaussagen abzielen, werden überwiegend Designs gewählt, die keine Kausalschlüsse zulassen. Als Beispiel sei hier die Frage danach zu nennen, ob Anwesenheit von Studierenden zu einem höheren Lernerfolg führt. SCHULMEISTER (2015) weist darauf hin, dass die meisten Studien zu dieser Frage korrelative Designs aufweisen und dass aus diesen keine kausalen

Schlussfolgerungen abgeleitet werden können; die Richtung des Zusammenhangs zwischen Anwesenheit und Lernerfolg sei nicht klar. Anwesenheit könnte unabhängig von anderen Faktoren auf den Lernerfolg wirken oder aber eine Mediator- oder Moderatorvariable darstellen – oder es könnten andere Variablen für den Zusammenhang zwischen Anwesenheit und Lernerfolg kausal verantwortlich sein. CREDÉ, ROCH & KIESZCZYNSKA (2010) haben in ihrem metaanalytischen Review versucht, sich der Kausalitätsfrage mithilfe von Regressionsanalysen zu nähern; doch auch dieses Vorgehen kann letztlich keine Kausalität nachweisen. Für einen solchen Nachweis wären experimentelle Studien nötig. Bei der Umsetzung experimenteller Studien kommt es jedoch zu einigen Schwierigkeiten: Am Beispiel der Frage nach den Wirkungen von Anwesenheit wird deutlich, wie schwer das Finden einer adäquaten Kontrollgruppe ist. Selbst wenn sich in einer Studie dann ein positiver Effekt von Anwesenheit auf den Lernerfolg zeigen würde, steht die Identifikation der dafür relevanten Aspekte aus. Zum Beispiel wäre zu klären, ob in der Veranstaltung Tipps für die Prüfung gegeben oder Inhalte vermittelt wurden, zu denen die Lernenden sonst keinen Zugang gehabt hätten. Wenn dies der Fall ist, wäre der positive Zusammenhang von Anwesenheit und Lernerfolg ein Artefakt. Zwar ist die ausschließlich in der Veranstaltung erfolgende Vermittlung spezifischer Inhalte relativ leicht zu kontrollieren, aber eine Erhebung/Kontrolle *aller* relevanten Variablen ist in einem realen Setting kaum möglich. Zudem müssen auch ethische Aspekte bei der Versuchsplanung berücksichtigt werden, zum Beispiel wäre es bedenklich und auch rechtlich nicht möglich, eine Gruppe zu einer Anwesenheit – oder gar Abwesenheit – zu zwingen. Obwohl also viele korrelative Studien auf einen positiven Zusammenhang zwischen Anwesenheit und Lernerfolg hinweisen, bleibt die Frage nach der Richtung des Zusammenhangs bzw. den Ursachen und Wirkungen ungeklärt. Dieses Kausalitäts-Problem trifft auch auf viele weitere Forschungsfragen zu, zu denen keine kontrollierten experimentellen Studien vorliegen.

Eine Lösung könnte in einer stärkeren Anlehnung der Studien an experimentelles Vorgehen bestehen. Beispielsweise könnten bestimmte – möglicherweise die Studienergebnisse beeinflussende – Variablen miterfasst und dadurch kontrolliert wer-

den oder es könnte mit balancierten Designs gearbeitet werden (z. B. abwechselnde Reihenfolge der Nutzung verschiedener Methoden in verschiedenen (Teil-)Gruppen, um niemanden zu bevor- oder benachteiligen). Zudem besteht ein Vorschlag zur Verbesserung unseres Erachtens vor allem auch in einer besseren Kommunikation von Forschungsergebnissen, indem z. B. die Grenzen der jeweiligen Forschung klarer kommuniziert werden.

2.5 Stichproben-Problem

Ein praktisches Problem betrifft das Stichproben-Problem oder die Frage „Gelingt es, hinreichend große und repräsentative Stichproben zu gewinnen?“. Dieses Problem ist eng mit dem Spezifitäts-Problem verknüpft: Es werden häufig nur kleine Stichproben erhoben, die allenfalls einen Teil der Lehrenden und Studierenden repräsentieren. Häufig ist die Teilnahme an einer Studie freiwillig, so dass allein dadurch eine (positive Selbst-)Selektion stattfindet. Zudem entstammen die Stichproben häufig psychologischen, bildungswissenschaftlichen und/oder eigenen Lehrveranstaltungen (vgl. SoTL) oder zumindest dem eigenen universitären Kontext. Im Gegensatz zum schulischen Bereich, in dem bundesweite oder gar international repräsentative Stichproben erhoben werden (z. B. PISA) lassen sich fach- oder hochschulübergreifende Studien an großen und zufällig ausgewählten Stichproben bislang kaum durchführen. Um dies zu ermöglichen, bedürfte es eines ähnlichen Konsenses für die Bildungsmonitoring-Strategie an Hochschulen, wie dies über die von der Kultusministerkonferenz vereinbarte Strategie für die Schulen der Fall ist. Auch ist zur Gewinnung großer, unselektierter, nach Möglichkeit längsschnittlich begleiteter Stichproben das Commitment von Hochschulleitungen und die Kooperation von Datenschutz-Beauftragten erforderlich.

2.6 Messinstrumente-Problem

Das letzte hier zu diskutierende Problem betrifft das Messinstrumente-Problem oder die Frage „Gibt es Messinstrumente, die Prädiktoren und Kriterien zuverlässig und valide erfassen?“. Für den schulischen Bereich finden sich zur Erfassung von

verschiedensten lehr- und lernrelevanten Konstrukten vielfältige valide und objektive Instrumente (z. B. Kompetenztests, Instrumente zu motivationalen Konstrukten), die regelmäßig und an großen und repräsentativen Stichproben normiert werden. Für den Hochschulbereich liegen derartige Maße nur punktuell vor. Es besteht vor allem ein Mangel an Maßen, die objektive Kriterien, wie z. B. Lernerfolg erfassen. Meist werden in der hochschulischen Forschung Selbstberichte von Lehrenden und Studierenden genutzt. Wünschenswert wären neben objektiven Leistungskriterien auch Verhaltensmaße, so dass etwa festgestellt werden kann, ob Intentionen auch tatsächlich in Verhalten umgesetzt werden.

3 Fazit

Die Qualität der Hochschullehre ist in den letzten Jahren zunehmend in das Blickfeld der Öffentlichkeit gerückt. Daher besteht inzwischen eine erhöhte Nachfrage nach soliden empirischen Erkenntnissen über gute Hochschullehre. Die empirische Bildungsforschung ist bei der Gewinnung dieser Erkenntnisse jedoch mit verschiedenen Problemen konfrontiert. In diesem Beitrag haben wir sechs dieser Probleme diskutiert und einige Vorschläge gemacht, was wir für die Gewinnung solider empirischer Erkenntnisse benötigen. Beispielsweise sollte dem Kriterien-Problem mit Definitionen guter Hochschullehre und damit zusammenhängend Untersuchungsdesigns mit verschiedenen Kriterien für gute Hochschullehre begegnet werden. Das Prädiktoren-Problem sollte durch eine stärkere Fokussierung der Forschung auf solche Faktoren begegnet werden, die nah an der Hochschullehre und potentiell veränderbar sind. Einer Lösung beider genannten Probleme wäre zudem eine theoriegeleitete Operationalisierung zuträglich. Dem Spezifitäts-Problem könnten ein Transfer von Theorien und Prinzipien aus gut erforschten Bereichen, die Entwicklung hochschulspezifischer Wirkungsmodelle und die Analyse von Gemeinsamkeiten und Unterschieden entgegenstehen. Das Kausalitäts-Problem könnte zumindest abgeschwächt bzw. deutlich gemacht werden durch eine stärkere Anlehnung an experimentelle Vorgehensweisen und die realistische Verdeutlichung der Grenzen durchgeführter Studien. In Bezug auf das Stichproben-Problem wären eine bun-

desweite Bildungsmonitoring-Strategie sowie das Commitment von Hochschulleitungen sowie die Kooperationsbereitschaft von Datenschutz-Beauftragten hilfreich. Das Messinstrumente-Problem wird sich in dem Maße lösen lassen, wie die Forschung sich über Selbstberichtmaße hinaus auch objektiven Maßen für Lernerfolgen, Verhalten und weiteren objektiven Kriterien für gute Hochschullehre zuwendet. Die Umsetzung der meisten dieser Vorschläge würde von einer Kooperation von Forscherinnen und Forschern verschiedener Universitäten profitieren (z. B. um Normierungen an repräsentativen Stichproben zu ermöglichen). Wir hoffen daher, dass unser Beitrag nicht nur theoretisch rezipiert, sondern auch in konkretes Handeln bzw. Forschen umgesetzt wird.

4 Literaturverzeichnis

Berliner, D. C. (2002). Comment: Educational research: The hardest science of all. *Educational Researcher*, 31, 18-20.

Boyer, E. L. (1990). *Scholarship reconsidered. Priorities of the professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Credé, M., Roch, S. G. & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, 80, 272-295.

Hattie, J. A. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London u. a.: Routledge.

Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.

Helmke, A. & Schrader, F.-W. (2010). Hochschuldidaktik. In D. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 273-279). Weinheim: Beltz.

Huber, L. (2011). Forschen über (eigenes) Lehren und studentisches Lernen – Scholarship of Teaching and Learning (SoTL). Ein Thema auch hierzulande? *Das Hochschulwesen*, 59, 118-124.

Huber, L., Pilniok, A., Sethe, R., Szczyrba, B. & Vogel, M. (Hrsg.) (2014). *Forschendes Lehren im eigenen Fach. Scholarship of Teaching and Learning in Beispielen*. Bielefeld: W. Bertelsmann Verlag.

Hutchings, P., Huber, M. T. & Ciccone, A. (2011). *The scholarship of teaching and learning reconsidered. Institutional integration and impact*. San Francisco, CA: Jossey-Bass.

Schneider, M. & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143, 565-600.

Schulmeister, R. (2015). *Studien zur Anwesenheit in Lehrveranstaltungen. Ein nur scheinbar triviales Problem. Eine Meta-Studie von 300 empirischen Arbeiten*. Hamburg.

Spinath, B. & Seifried, E. (2012). Forschendes Lehren: Kontinuierliche Verbesserung einer Vorlesung. In M. Krämer, S. Dutke & J. Barenberg (Hrsg.), *Psychologiedidaktik und Evaluation IX* (S. 171-180). Aachen: Shaker.

Spinath, B. & Stehle, S. (2011). Evaluation von Hochschullehre. In L. Hornke & M. Amelang (Hrsg.), *Enzyklopädie der Psychologie, Grundfragen und Anwendungsfelder psychologischer Diagnostik: Psychologische Diagnostik 1* (S. 617-667). Göttingen: Hogrefe.

Ulrich, I. (2016). Professionalisierung der Hochschuldidaktik: Modelle guter Hochschullehre als gemeinsames Fundament. In M. Heiner, B. Baumert, S. Dany, T. Haertel, M. Quellmelz & C. Terkowsky (Hrsg.), *Was ist gute Lehre? Perspektiven der Hochschuldidaktik* (Blickpunkt Hochschuldidaktik, Bd. 129, S. 31-46). Bielefeld: Wilhelm-Bertelsmann.

Autorinnen



Prof. Dr. Birgit SPINATH || Ruprecht-Karls-Universität Heidelberg, Psychologisches Institut || Hauptstraße 47-51, D-69117 Heidelberg

www.psychologie.uni-heidelberg.de/ae/paeps/personen/spinath.html

birgit.spinath@psychologie.uni-heidelberg.de



Dr. Eva SEIFRIED || Ruprecht-Karls-Universität Heidelberg, Psychologisches Institut || Hauptstraße 47-51, D-69117 Heidelberg

www.psychologie.uni-heidelberg.de/ae/paeps/personen/seifried.html

eva.seifried@psychologie.uni-heidelberg.de