

Evaluation forschenden Lernens – Dokumentation des Forschungskolloquiums

Eileen Lübcke, Peter Tresp, Johannes Polzin, Sylvia Lepp, Jan Hense, Mareike Bartels, Mandy Schiefner-Rohs, Anna-Christin Herrmann, Jennifer Preiß

Zum Hintergrund

Im Rahmen des Qualitätspakt Lehre (QPL) sind viele BMBF-geförderte Projekte entstanden, die einen Schwerpunkt auf forschendes Lernen (FL) legen. Diese Projekte sind nun seit einigen Semestern aktiv, haben ihre Einstiegsprobleme erfolgreich bearbeitet und daher nun einerseits den Raum, sich mit der Evaluation ihres Programms auseinanderzusetzen. Andererseits besteht die Notwendigkeit, zu belegen, dass ihr FL-Projekt erhaltenswürdig ist, da der Qualitätspakt Lehre Ende 2020 ausläuft und sich somit die Frage nach der Verstetigung immer dringender stellt.

Das Projekt „FideS“ („Forschungsorientierung in der Studieneingangsphase“) war ein BMBF-gefördertes Begleitforschungsprojekt zum QPL und ist nun als solches unter dem Namen „FideS-Transfer“ verlängert worden. Ziel von „FideS-Transfer“ ist es, die QPL-Partner-Projekte, die in der Studieneingangsphase forschendes Lernen durchführen, in ihren Bedarfen zu unterstützen. So basiert diese Unterstützung auf der Aufarbeitung und Bereitstellung der Forschungsergebnisse aus „FideS“: Es geht dabei um den Ausbau der Vernetzung, Erstellung von Materialien zu Herausforderungen forschenden Lernens, Weiterentwicklung digitaler Unterstützungstools sowie Perspektiven auf die Evaluation forschenden Lernens zu legen.

Der Besuch des Gastwissenschaftlers Prof. Dr. Peter Tresp, der gemeinsam mit Thomas Hildbrand das „Zürcher Framework“ für forschendes Lernen entwickelte (Tresp/Hildbrand 2012), bot schließlich eine gute Gelegenheit, ein Forschungskolloquium dem Thema „Evaluation forschenden Lernens“ zu widmen und dazu auch Projektpartner_innen einzuladen, um gemeinsam über diese Herausforderung nachzudenken. Teil nahmen schließlich etwa 25 Personen – Repräsentant_innen unterschiedlicher Projekte aus ganz Deutschland, Mitarbeitende der „Servicestelle Evaluation“ der Universität Hamburg, das FideS-Projektteam aus Hamburg und Kaiserslautern, Peter Tresp, sowie Prof. Dr. Jan Hense, der selbst zu Lehrevaluationen forscht und derzeit Vorsitzender der DeGEval (Gesellschaft für Evaluation) ist.

Im Rahmen des Forschungskolloquiums wurden ganz unterschiedliche Perspektiven auf die Evaluation von forschendem Lernen deutlich. Wie Jan Hense zu Beginn seines resümierenden Vortrags betonte: In der Evaluation kommen sehr viele unterschiedliche Ebenen zum Tragen, die schnell und leicht vermischt werden. Wir wollten mit dem Forschungskolloquium einen Beitrag dazu leisten, diese verschiedenen Perspektiven auf die Evaluation von forschendem Lernen zu explizieren und zu reflektieren (z.B. Evaluationsebenen, -methoden oder Ziele). Wir haben versucht, uns dem Thema über unterschiedliche Zugänge zu nähern: Zum einen haben wir Berichte aus der Begleitforschung, zum anderen aber auch direkte Projektbeispiele und Überblicksbeiträge integriert.

Die Zielsysteme forschenden Lernens – Forschungsergebnisse aus FideS

Um sich über Evaluation Gedanken zu machen, ist es notwendig, erst einmal zu explizieren, welche Ziele man mit forschendem Lernen in der Studieneingangsphase überhaupt verfolgt. Um dies zu explizieren, konnten wir auf die Interviewdaten aus dem Projekt FideS zurückgreifen. Dort wurden mit 40 Beteiligten aus insgesamt 19 Projekten Interviews geführt.

Ein Aspekt der Analyse der Transkripte war die Herausarbeitung der Ziele forschenden Lernens unserer Projektpartner_innen und damit verbunden auch die Darstellung von Wirkannahmen zum forschenden Lernen. Die Untersuchungsebene in diesem Kontext ist folglich nicht unbedingt ein bestimmtes Projekt oder eine bestimmte Lehrveranstaltung, sondern das methodisch-didaktische Prinzip des forschenden Lernens selbst. Aus insgesamt 15 Interviews konnten zusammenfassend vier Zielsysteme der Akteure erschlossen werden, die in folgendem Schaubild dargestellt werden¹:

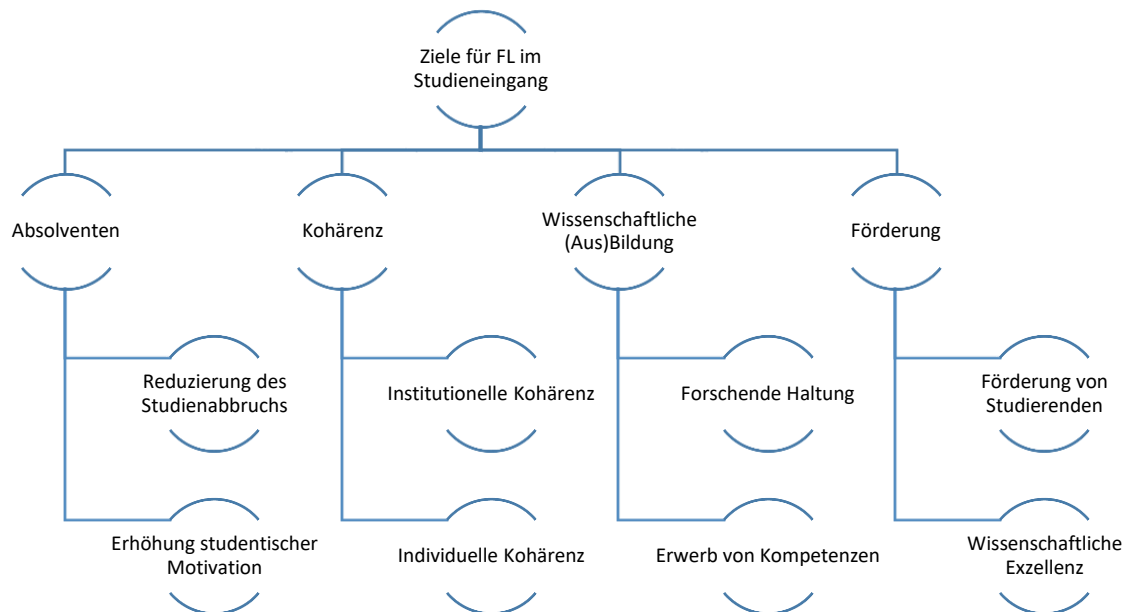


Abb. 1: Hauptziele für forschendes Lernen in der Studieneingangsphase. Quelle: Präsentation FideS, Folie 7

Das Ziele und Wirkannahmen keine trennscharfen Kategorien sind, sieht man am ersten Zielsystem der Erhöhung der Absolventenquote. Unterziele für dieses Hauptziel ist die Reduzierung des Studienabbruchs. Da Motivation ein entscheidender Faktor für den Verbleib in einem Studiengang ist, ist die Erhöhung der Motivation ebenfalls ein Unterziel, welches mit forschendem Lernen erreicht werden soll. Gleichzeitig ist es eine Wirkannahme, dass mit Erhöhung der Motivation die Wahrscheinlichkeit des Studienabbruchs reduziert wird. Wie detailreich die im Rahmen von FideS herausgearbeiteten Ziele sind, zeigt das Beispiel „Wissenschaftliche (Aus-)Bildung“:

¹ Die Ergebnisse dieser Forschung wurden außerdem publiziert in Lübcke, E. Heudorfer, A. (in Druck 2019): Die Ziele forschenden Lernens. Eine empirische Analyse im Rahmen der QPL-Begleitforschung. In: Reinmann, G. Lübcke, E. Heudorfer, A.: Forchendes Lernen in der Studieneingangsphase - Empirische Befunde, Fallbeispiele und individuelle Perspektiven. Springer VS. S. 17-48.

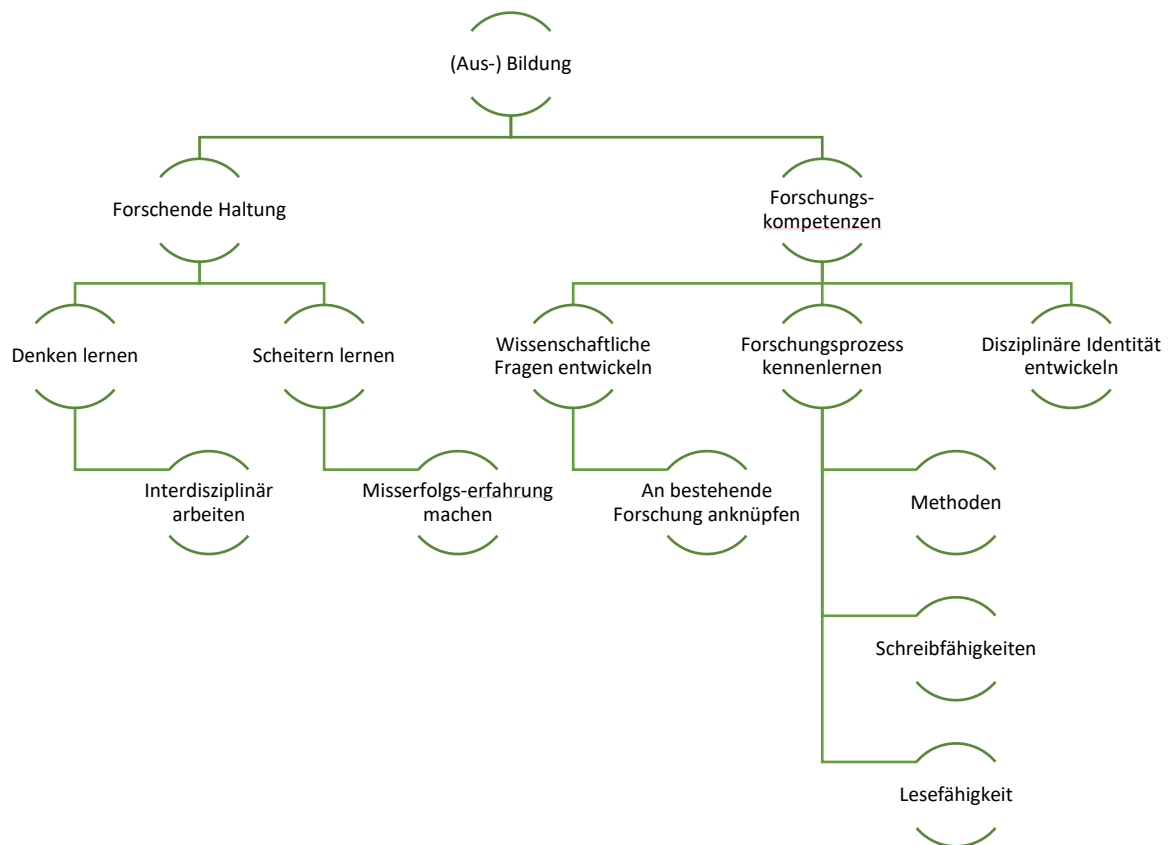


Abb. 2.: Ziele forschenden Lernens mit dem Hauptziel der (Aus-)Bildung. Quelle: Präsentation FideS, Folie 8

Zu betonen ist hier, dass Wirkannahmen keine Wirkungen sind, sondern gewissermaßen Arbeitshypothesen der Projektakteure. Diese Arbeitshypothesen – in diesem Fall die Zielsysteme forschenden Lernens - können als Kategorien genutzt werden, um bei einer Evaluation mögliche Felder einzugrenzen bzw. Foki zu legen, die bei einer Evaluation forschenden Lernens untersucht werden können. Zugleich legen die Zielsysteme offen, dass methodisch aus unterschiedlichen Bereichen Zugänge gewählt werden müssen: So sind in einem Fall Absolventenstudien angemessen, während im anderen Fall eher Lehrveranstaltungsevaluationen im Fokus stehen.

Eine Evaluation auf Basis der beschriebenen Zielsysteme könnte somit Unterschiedliches leisten: Es könnten die aktuell eher impliziten und in der Basis individuellen Annahmen zu den Wirkungen des Einsatzes forschenden Lernens expliziert und auf Haltbarkeit überprüft werden. Somit wäre ein erster Schritt zur Evidenzbasierung des forschenden Lernens geleistet. Zudem haben die einzelnen Unterziele Relevanz sowohl für die konkrete Umsetzung von Lehrveranstaltungen forschenden Lernens, wie auch für die strategische Ausrichtung von Studiengängen bis hin zur gesamten Hochschule – sei es mit Blick auf die wissenschaftliche Nachwuchsförderung oder der „Arbeitsmarktreife“ von Absolventinnen und Absolventen. Eine Präzisierung der Ziele für das eigene Projekt könnte damit Aufschluss über die angemessene Form der Evaluation bieten.

Die in FideS erschlossenen Zielsysteme bieten somit eine Grundlage für Untersuchungen auf der Makro-, Meso- und Mikroebene. Offen bleibt die Frage nach dem *Wie* und welche Instrumentarien für welche Zielsysteme überhaupt angemessen sind.

Peter Tremp – Evaluation und Forschungsorientierung in Studium und Lehre

Peter Tremp beschäftigte sich in seinem Beitrag zunächst damit, in welchem Rahmen Lernen stattfindet, nämlich innerhalb der Universität als Institution. Anhand historischer Konzeptionen von Universität, dargestellt in Bildern, präsentierte er die Entwicklung des Verständnisses dieser Bildungseinrichtung – ausgehend von einem Turm, in dem über Lernen eines bestehenden Wissenskataloges verschiedene „Etagen“ erklimmen werden konnten, hin zu einem Segelschiff, das ins Unbekannte aufbricht: Die Crew hat die nötigen Fertigkeiten das Schiff zu manövrieren, verfügt über Karten, doch weiß niemand, was auf der Reise zu erwarten ist. Anhand weiterer Bilder erläuterte er, wie sich die Lerneinrichtung bestehenden Wissens schließlich zu einer Forschungseinrichtung entwickelte.

Im Zürcher Framework steht nun genau diese Frage im Zentrum: Was heißt Verknüpfung von Forschung und Lehre für die hochschuldidaktische Gestaltung von Hochschule? Dafür steht die Frage im Mittelpunkt, welche Forschungstätigkeiten Studierende im Rahmen des Studiums durchführen (Klammer in der Mitte). Dem zugeordnet wurde untersucht, wie diese Etappen dokumentiert werden (davon werden dann Formen von Leistungsnachweisen abgeleitet) und welcher universitäre Lehrrahmen dafür geboten wird (vgl. Abb. 3).

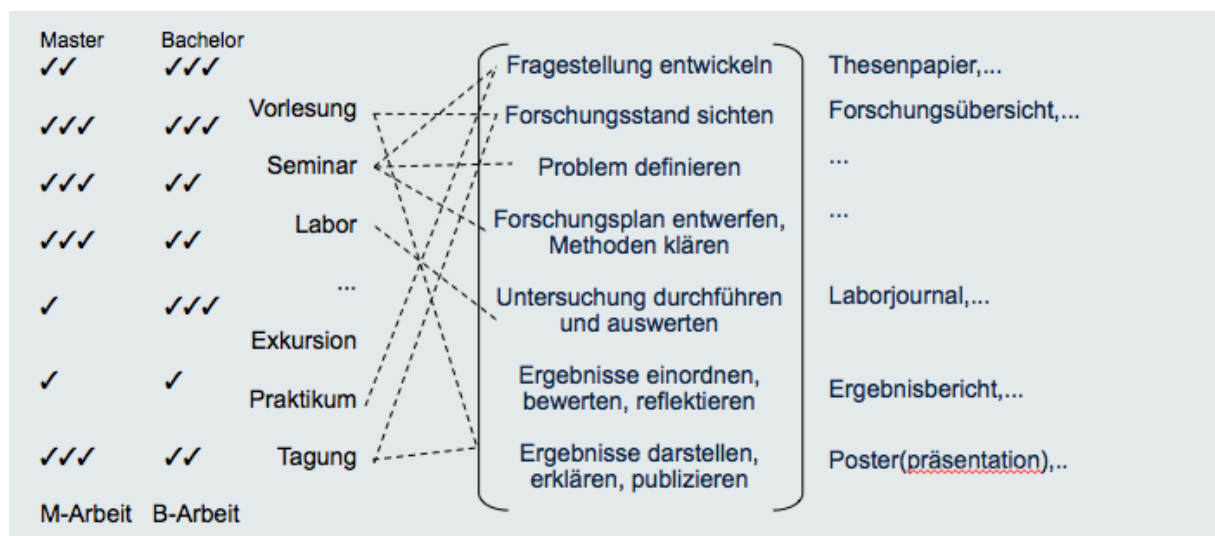


Abb. 3: Studienstufen mit unterschiedlichen Schwerpunktsetzungen, Quelle: Tremp & Hildbrand, 2012

Dabei wird deutlich, dass die Verantwortung für die Gestaltung der Verknüpfung von Forschung und Lehre auf unterschiedlichen Ebenen liegt: Auf der Ebene der didaktischen Gestaltung von Lehrveranstaltungen in der Hand der Lehrenden, auf der Ebene der Studienganggestaltung in der Hand von Modul- resp. Studiengangsverantwortlichen und schließlich auch auf der Ebene der Hochschule in der Hand von strategischen Entscheidungen auf Hochschulebene. Tremp schloss daraus die offene Frage, dass untersucht werden müsste, wie diese unterschiedlichen Verantwortungsebenen gegenstandadäquat evaluiert werden. Damit ist die in Tremps Beitrag betrachtete Perspektive auf die Evaluation forschenden Lernens die Betrachtung von der Bedingung, einen Rahmen für forschendes Lernen zu schaffen. Auch hier zeigt sich, wie im vorangegangenen Beitrag, dass Evaluation forschenden Lernens nicht alleine auf Veranstaltungsebene erfolgen kann, sondern ein Thema der Mikro- Meso- und Makro-Ebene von Hochschule ist.

Tremp betonte weiter die Relevanz von Handlungsraum für Studierende im Universitätskontext. Mit einer Parallele zur Kindergartenstufe und Zitaten von Fröbel und Seele beschrieb er die Relevanz der Aktivität „Spielen“ und wies damit auf den Bedarf hin, einen Rahmen für Studierende zu schaffen, in dem sie selbst forschend tätig werden können. Er beschrieb, dass in der Universität dafür der Grundmodus des Forschens herrscht, als strukturierte Grundidee des Lernens. Dozierende stellen dementsprechend die Personifizierung

dieses Prinzips dar, eine Bibliothek wäre daran angelehnt beispielsweise eine strukturierte Lernumgebung.

Tremp kritisierte, dass bestehende Lehrevaluationsbögen aktuell zumeist diesen Grundmodus des Forschens übergehen, womöglich, weil Lehrevaluationen meist an einer Verwaltungslogik orientiert sind. Leitende Überlegungen einer Evaluation müssten zunächst sein, die Ebenen in ihrer Differenz aufzunehmen. Tremp beschrieb zwei Ebenen: So gibt es eine strukturelle Differenz der unterschiedlichen Rollen (Lehrende und Studierende) und eine pädagogische Differenz zwischen Zeigen und Lernen. Beide Formen müssen in einer Evaluation zueinander ins Verhältnis gesetzt werden.

Lehre ist gemäß Huber (2018) in drei unterschiedlichen Lehrqualitätsstufen zu messen: „Standardgemäß“ als solides Konzept, das die Mindestansprüche erfüllt; „gut“ als Konzept, das flexibel auf Situationen reagiert und variieren kann und „exzellent“ als ein herausragendes Konzept, das „über sich hinaus wirkt“.

Unabhängig von den Qualitätsstufen der Lehre präsentiert Peter Tremp als kreative Alternative zur klassischen Lehrevaluation die traditionelle Idee einer Lehrreflexion, bei der es nicht um Bewertung geht, sondern darum, miteinander ins Gespräch zu kommen; er nahm dabei Bezug auf einen studentischen Vorschlag der 1960er Jahre, wonach Studierende Seminar- oder Vorlesungsrezensionen an Universität anfertigen könnten, in Analogie zu Buchrezensionen. Einerseits stellen sie damit eine Verbindung zwischen sich und ihren Erfahrungen in der Veranstaltung her, andererseits besteht die Möglichkeit, dass diese Rezensionen auf Mesoebene zur Veranstaltungsverbesserung aufgegriffen werden. Damit würde man die forschungsorientierten Handlungen, die im Zürcher Framework eingeführt wurden auch im Rahmen der Evaluation adressieren, schließlich kennt man Rezensionen als Forschungstätigkeiten, indem Artikel oder Bücher begutachtet werden. Lehrrezensionen wären dementsprechend ein pädagogisches Pendant dazu, welche aber Forschungsorientierung noch in sich tragen. Daraus ergeben sich wiederum viele neue, noch offene Fragen, beispielsweise danach, was der Inhalt einer solchen Rezension sein könnte. Einerseits können sie natürlich fachliche Inhalte reflektieren, aber auch didaktische Qualität der Lehrveranstaltung wäre als Inhalt eben als Evaluationsalternative denkbar. Dabei würden daraus dann jedoch weitere Fragen entstehen: Wer kann überhaupt in welchem Studienabschnitt und welchem Fach über die Qualität von Lehre urteilen? Wie kann man erfassen, was nachwirkt und nicht nur die Performanz des Lehrenden während der Lehrveranstaltung? All diese und weitere Fragen wurden im Plenum diskutiert, ohne dabei eindeutige Antworten zu generieren.

Zusammenfassend lässt sich festhalten, dass Tremp einige Kernfragen thematisierte, die im Kontext der Frage nach der Evaluation von forschendem Lernen aufkommen:

- In was für einer Institution findet forschendes Lernen statt und wie zeigen sich deren Werte in den Handlungspraktiken? Welches Selbstverständnis hat die Hochschule und welche Erwartungen entstehen daraus?
- Welche Akteure der Universität können über welche Bereiche urteilen? Ist es angemessen, Studierenden die Evaluation zu überlassen, oder gibt es nicht auch Bereiche, in denen Lehrende angemessenere Einschätzungen treffen können?
- Was soll gemessen werden? Geht es um die Erfassung der Lerngelegenheit (beispielsweise die durchgeführten Forschungstätigkeiten) oder geht es um das daraus (und möglicherweise auch aus anderem) generierte Können der Studierenden (das beispielsweise über Prüfungsleistungen erfasst werden kann)?

Die anschließende Diskussion widmete sich zunächst der offenen Frage der Evaluationsinstrumente, da die Möglichkeit der Erstellung eines einzigen Instruments, das für unterschiedlichste Projekte nutzbar wäre, angezweifelt wurde. Als Beispiel für ein gut geeignetes Evaluationsinstrument, welches keine klassische standardisierte Befragung ist,

wurde von Fokusgruppen und den damit verbundenen positiven Erfahrungen berichtet. Anschließend wurde überlegt, ob bereits die Artefakte der Studierenden, wie beispielsweise Forschungsberichte, möglicherweise in Kombination mit einem Forschungstagebuch oder Memos, eine gute Evaluationsgrundlage darstellen könnten.

Danach wurde wieder auf die Frage eingegangen, was evaluiert werden soll. Neben Rahmenbedingungen, Tätigkeiten und der Nutzung durch Studierende wurde außerdem hervorgehoben, dass es sinnvoll sei, die Gelegenheiten für Selbstwirksamkeitserfahrungen zu erfassen. Daraus entstand die grundsätzliche Frage, ob bereits das Durchführen von Forschungstätigkeiten forschendes Lernen sei oder ob es dafür eine innere Verbindung der Studierenden mit dem Forschungsgegenstand bedarf (Ist es auch Forschung, wenn mich das Thema nicht interessiert? Bzw. ist das ohne das andere überhaupt möglich?). All diese Fragen werfen wieder die offene Frage danach auf, mithilfe welcher Evaluationsinstrumente forschendes Lernen untersucht werden kann. Diese Frage stand dann insbesondere im nächsten Beitrag im Fokus.

Entwicklung eines Instruments zur Evaluation an der Hochschule für Wirtschaft und Umwelt Nürtingen-Geislingen

Hintergrund des Beitrages von Sylvia Lepp und Johannes Polzin von der Hochschule Nürtingen-Geislingen ist die Erarbeitung eines Evaluationsbogens für Lehrveranstaltungen und Projekte im Format (Projekte) forschenden Lernens. Im Forschungskolloquium präsentierten sie den aktuellen Stand ihres Prozesses.

Auslöser für die Neuerstellung der Evaluation war die Feststellung, dass aktuelle Bögen den Rahmen und die enthaltenen Entwicklungen eines Projektes forschenden Lernens nicht erfassen können. Da die Hochschule für Wirtschaft und Umwelt einen großen Wert auf Forschung und Forschungsorientierung im Studium legt, gibt es folglich einen erhöhten Bedarf an angemessener Evaluation.

Die Evaluation soll dabei Zwecke auf unterschiedlichen Ebenen erfüllen:

- Auf Ebene der Studiengangentwicklung soll das Erreichen der Studienziele überprüft werden.
- Für Lehrende soll die Evaluation ein Diagnose- und Feedbackinstrument darstellen.

Der Anspruch von Lepp und Polzin ist dabei, dass Lehrende den Evaluationsbogen nicht mehr als Bewertungsbogen wahrnehmen. Damit zeigen sich Parallelen zu den Ideen von Peter Tremp, Evaluation als Gespräch oder als Reflexion zu sehen und nicht als Bewertung.

Aktuell erstellen sie gemeinsam mit den Lehrenden einen Bogen bei dem die Lehrende Mitsprache darüber haben, zu welchen Themen sie Feedback wünschen und benötigen.

Zunächst wurde dafür im Rahmen eines Workshops ein Itemkatalog erstellt und besprochen.

Die folgende Grafik stellt die Aspekte dar, welche in der Abfrage erfasst werden sollten:

Strukturdaten	Prozessdaten Lehrende	Prozessdaten Studierende	Lernergebnisse
Rahmenbedingungen - Räume - Bibliothek - Semestergröße - Technikausstattung	Lehrverhalten - Zielklärung - Struktur - Anregung zum Mitdenken - Interaktion - Praxisbezug	Lernaktivitäten - Vor- Nachbereitung - Diskussionen - Kritisches Hinterfragen - Recherchieren	Kompetenzzuwachs - Fachkompetenz - Methodenkomp. - Sozialkompetenz - Selbstkompetenz
Lehrende - Didaktische Kompetenz, - Fachkompetenz, - Motivation	Motivierung - Begeisterung - Freundlichkeit - Kooperativität	Lernformen - Lerngruppen	- Digitale Kompetenz - Professionalität (Nachhaltigkeit, Wissenschaftlichk.)
Studierende - Vorwissen - Interesse - Motivation	Produkte - Schriftliche Ausarbeitungen - Präsentationen - Projekte		Produkte - Schriftliche Ausarbeitungen - Präsentationen - Projekte

Abb. 4: Abfrageitemsbereiche. Quelle: Präsentation Lepp/Polzin, Folie 4

Bei der Erstellung der Items sollten die Lehrenden alle Bereiche und Ebenen der Bedingungsfaktoren von Lehrqualität berücksichtigen. Strukturdaten, Lehrverhalten, Lernverhalten und Lernergebnisse.

Zum Zeitpunkt des Forschungskolloquiums befand sich der Itempool in einer Feedbackschleife mit den Lehrenden, die Projekte im Format des forschenden Lernens durchführen, um die 30 Items mit der größten Relevanz zu bestimmen.

In der an die Präsentation der Vorgehensweise anschließenden Diskussion wurde zunächst die Frage behandelt, ob eine partizipative Erstellung des Fragebogens sinnvoll ist. Gemeinsam wurde festgestellt, dass eine solche Handhabung die Gefahr birgt, dass Lehrende die Fragebögen so gestalten, dass sie „gut dabei weg kommen“. Andererseits sind sie nah am Lehrgeschehen und können die stattfindenden Prozesse aufgrund der Erfahrungen erster Hand gut beurteilen. Außerdem können sie tatsächlich ein Feedbackinstrument erstellen, das sie darin unterstützen kann, die eigene Lehre zu verbessern. Voraussetzung dafür ist, dass sie offen dafür sind, eigene Schwächen betrachten zu wollen, aber auch, dass sie keine Angst vor Konsequenzen aufgrund negativer Feedbacks haben müssen.

Anschließend bewegte sich die Diskussion um den Zeitpunkt, zu dem die Evaluation durchgeführt wird – beim Forschenden Lernen ein entscheidender Faktor, denn bisherige Erfahrungen in den untersuchten Projekten deuten hier unterschiedliche Evaluationsergebnisse an. So profitieren zeitlich verlagerte Evaluationen, etwa nach Studienabschluss, davon, dass Studierende die Gesamtheit ihres Studiums rückblickend betrachten und so gegebenenfalls wahrnehmen können, dass sie von den beim forschenden Lernen erworbenen Fähigkeiten im Studienverlauf profitieren konnten – eine Einschätzung, zu der sie mangels Erfahrungswerte nicht gelangen können, wenn direkt nach Abschluss der FL-Lehrveranstaltung evaluiert wird. Gleichzeitig ist zu fragen, ob die Studierenden sich retrospektiv noch an Veranstaltungsformate erinnern können. Bei einer sehr nah an der Lehrveranstaltung terminierten Evaluation könnten sich zudem Faktoren, die als belastend wahrgenommen werden (wie bspw. der erhöhte Zeitaufwand und Arbeitseinsatz), negativ auf die Evaluation auswirken.

Die Durchführung der Evaluation in der laufenden Vorlesungszeit, etwa zur Mitte des Semesters, könnte wiederum den Vorteil bieten, dass das Feedback der Studierenden zur Prozesssteuerung genutzt werden kann. Da Schwierigkeiten beim forschenden Lernen oftmals mit dem hohen Arbeitspensum und der weniger reproduktiven Art des Lernens zusammenhängen (vgl. bspw. Huber 2014), gäbe es hier eine Möglichkeit der didaktischen Nachjustierung im laufenden Prozess. Hier würde die Evaluation weniger einen Charakter der abschließenden Bewertung annehmen und vielmehr unmittelbar zur Verbesserung der Lernsituation eingesetzt werden können – ganz im Sinne eines formativen Evaluationsansatzes.

Der Beitrag von Lepp und Polzin präsentierte eine Herangehensweise der Erstellung eines Fragebogens zur Evaluation forschenden Lernens. Spannend ist, dass die relevanten Merkmale zur Evaluation von den partizipierenden Lehrenden über ihre Items bestimmt werden. Insbesondere diese Form der Partizipation Lehrender an der Erstellung ist dabei ein inspirierendes Novum. Deutlich wurde bei dem Beitrag auch, dass die Form der Hochschule Auswirkungen auf die Evaluationserstellung hat – von der Beauftragung bis hin zur Entwicklung des Instruments. Aus dem Beitrag von Lepp und Polzin lässt sich folglich die Frageperspektive generieren, wer die Entscheidungsgewalt über die Form der Evaluation haben darf.

Welchem Zweck dient eine Evaluation forschenden Lernens?

Haben sich die bisherigen Beiträge mit der konkreten Frage nach dem „Wie“ der Evaluation forschenden Lernens beschäftigt, stellt der letzten Beitrag von Jan Hense noch einmal die Frage, warum und wozu eigentlich eine Evaluation forschenden Lernens dienen soll. Er präsentiert dafür generelle Grundlagen von Evaluation und stellt dieses abschließend den besonderen Merkmalen des forschenden Lernens gegenüber.

Jan Hense von der Universität Gießen und derzeitiger Vorsitzender der DeGEval (Gesellschaft für Evaluation) argumentierte in seinem resümierenden Beitrag, dass viele im Kolloquium diskutierte Fragen mit grundsätzlichen Entscheidungen betreffend einer Evaluation zu tun haben. Zunächst präsentierte er die von der DeGEval definierten Standards von Evaluation: Eine Evaluation soll (1) nützlich, (2) durchführbar, (3) fair und (4) genau sein. Evaluationen können dabei unterschiedliche Zwecke verfolgen. Um der Forderung nach Nützlichkeit Rechnung tragen zu können, ist es wichtig, zu fragen, welche dieser Zwecke im konkreten Fall verfolgt werden sollen. Sie können als Entscheidungshilfe dienen, Verbesserungshinweise geben, zu Rechenschafts- bzw. Legitimationszwecken herangezogen werden und generelles Lernen über den Evaluationsgegenstand fördern. Eine Evaluation ist immer eine Intervention, die etwas verändern möchte.

Er betonte, dass ohne zuvor den Zweck einer Evaluation definiert zu haben, die Konzeption einer Evaluation nicht zielorientiert möglich sei. Insbesondere die während des Kolloquiums diskutierten Fragen nach dem Verfahren, Datenquellen, Instrumenten, Zeitpunkt, Inhalt, Generalisierbarkeit und weiterem, ließen sich ohne die vorherige Festlegung des Zwecks nicht sinnvoll diskutieren. Hilfreich bei der Frage nach dem Zweck ist die Frage nach den „intended users“, also den Nutzer_innen der Evaluation. Mit unterschiedlichen Nutzergruppen sind meist auch unterschiedliche Zwecke der Evaluation verbunden. So sind beispielsweise Evaluationen für Projektverantwortliche zumeist formativ und dienen der Verbesserung des Projektes. Evaluationen für Mittelgeber oder der Öffentlichkeit dienen eher der Rechenschaft, Evaluationen für Lehrende oder Hochschuldidaktiker sollen Lerngrundlagen legen, bei Evaluationen für die Scientific Community wird meist ein verallgemeinerbarer Erkenntnisgewinn angestrebt.

Jan Hense wies darüber hinaus auf die Risiken hin, die aufkommen, wenn bestehende Evaluationsinstrumente unreflektiert übernommen werden, da die dahinterliegenden Theorien

dabei unbewusst „miteingekauft“ werden. Stattdessen plädierte er für eine Reflexion der Gegenstände der geplanten Evaluation, um dann anhand dessen in bestehenden Instrumenten nach übertragbaren Ansätzen zu suchen und falls nicht fündig eigene, gegenstandsangemessene Instrumente zu entwickeln.

Dabei dürfe Evaluation und Forschung nicht synonym verstanden werden: Evaluation und Forschung ähneln sich im Prozess, verfolgen jedoch unterschiedliche Ziele: Evaluation möchte letztlich erfahren, ob etwas „gut“ oder „erfolgreich“ war, also bewerten, während Forschung im Grunde an der Differenz „wahr/unwahr“ oder „richtig/unrichtig“ interessiert ist. Um die Logik des Bewertens im Rahmen einer Evaluation zu verstehen, muss sich die Frage gestellt werden, worauf eine Bewertung grundsätzlich beruht („Logik der Evaluation“): Auf welchen Kriterien (1), auf welchen Zielwerten (2), auf welche Art gemessen wird (3) und wie unterschiedliche Kriterien zu einer übergreifenden Wertzuweisung verknüpft werden (4). Auf der Ebene der Lehrenden besteht beispielweise das Risiko, Lehrende anhand von Lehrveranstaltungsevaluationen vergleichen zu wollen (wie es beispielsweise in Berufungssituationen geschieht). Dies würde dem Prinzip der Fairness widersprechen, da Studierende aufgrund ihrer eigenen Affinität zum Thema gebiast sind und daher die Evaluation ‚unbeliebter‘ Veranstaltungen zwangsläufig für Lehrende schlechtere Ergebnisse liefern würde. Dabei bleibt immer die Frage offen, ob tatsächlich die evaluierte Maßnahme den untersuchten Effekt hervorgerufen hat oder ob andere Faktoren diese Veränderung hervorgerufen haben.

Aktuell sind Lehrveranstaltungsevaluationen die dominante Form der Evaluation im Kontext von Studium und Lehre. Eine Übertragung von gewöhnlichen Lehrveranstaltungsevaluationen auf Formate des forschenden Lernens ist jedoch ungenügend, da es, wie bereits erwähnt, die unterschiedlichen Zwecke der Evaluation gibt und der Gegenstand ein spezifischerer ist. Forschendes Lernen hat spezifischere Ziele und Begründungen als Lehre allgemein.

Fragen, deren Bearbeitung erst dann möglich sind, wenn diese Spezifika in der Evaluation berücksichtigt werden sind beispielsweise, was forschendes Lernen besonders wirksam macht, auf welche Bedingungen es angewiesen ist, oder welche vermittelnden Faktoren eine Rolle spielen. Dazu ist es wichtig, im Sinne der programmtheoriebasierten Evaluation nach den (theoretischen) Wirkannahmen zu fragen, die dem Ansatz des forschenden Lernens zugrunde liegen.

Am Beispiel der Lehrveranstaltungsebene und der Studiengangsebene skizziert Jan Hense mögliche Fragen, die eine Evaluation adressieren kann. Diese werden in der folgenden Grafik dargestellt, die aus seiner Präsentation übernommen wurde:

	Input	Prozess	Outcomes
Lehrveranstaltung	Voraussetzungen - der Lernenden - der Lehrenden - im Kontext	Forschungsprozess (Zürcher Modell) Fragestellung entwickeln Forschungsstand sichten Problem definieren Forschungsplan entwerfen, Methoden klären Untersuchung durchführen und auswerten Ergebnisse einordnen, bewerten, reflektieren Ergebnisse darstellen, erklären, publizieren	Ziele forschenden Lernens auf Ebene der Studierenden : - Kompetenzerwerb etc. - <u>spezifische Ziele</u>
Curriculare Verankerung	Voraussetzungen - Studiengang - Lehrende - Studierende - Ressourcen - ...	Implementierungsaktivitäten - Informationen? - Weiterbildung? - QS? ...	Ziele forschenden Lernens auf Ebene des Curriculums : Erfolgreiche Einführung und regelhafte Umsetzung...

Abb. 5: Mögliche Kriterien auf Lehrveranstaltungs- und Curricularer Ebene bei der Evaluation forschenden Lernens. Quelle: Präsentation Jan Hense.

Aus der Zielanalyse von FideS ergibt sich, dass unter Umständen auch noch die Hochschulebene einbezogen werden muss.

Knappes Resümee zur Veranstaltung

Zusammenfassend lässt sich festhalten, dass die Evaluation von forschendem Lernen besondere Schwierigkeiten beinhaltet. Forschendes Lernen ist ein Konzept, welches in seiner Ausprägung immer als Projekt oder Veranstaltung sichtbar wird, folglich kann nur dessen konkrete Erscheinung evaluiert werden. Im Rahmen der Veranstaltung wurde die Vielschichtigkeit des Konzeptes weiter deutlich und damit auch, auf welchen unterschiedlichen Ebenen was evaluiert werden kann. Auf die Frage nach dem ‚Wie‘ wurden selbstverständlich keine eindeutigen Antworten gefunden. Dennoch hat diese Veranstaltung viele Impulse gegeben, Ziele, Zwecke und Interessen eigener Evaluationsvorhaben auf unterschiedlichen Ebenen zu reflektieren, um schließlich ein eigenes Evaluationsinstrument zu entwickeln. So können die Zielsysteme, die aus den Arbeiten von FideS entstanden sind, als Grundlage dienen, für das eigene Projekt eine eigene Programmtheorie zu entwickeln. Mit dem Konzept Programmtheorien liegt einer der zentralen Ansätze in der Evaluationsforschung vor, bei der es um die Herausarbeitung von Wirkannahmen der Projekt- bzw. Programmakteure geht, die die Projekte gestalten und umsetzen (Chen 1990, Rogers et al. 2000). Diese individuellen Programmtheorien (Giel 2016) können dann die anstehende Evaluation leiten.

Mit dem Zürcher Framework steht eine Beschreibung des auf Handlungspraktiken fokussierten curricularen Rahmens zur Verfügung, der ebenfalls wieder in ein Evaluationsinstrument überführt werden kann. Diesmal mit der Frage, ob die verschiedenen Formate auf unterschiedlichem Studienniveau auch die entsprechenden Forschungstätigkeiten beinhalten und wie diese in das entsprechende Produkt überführt werden. Um die Angemessenheit dieser Fragen zu beantworten, können dafür aber auch dialogische Konzepte der Evaluation aufgenommen werden, wie bspw. die von Peter Treppe angeregten Reflexionsgespräche über die Lehre.

Und mit dem Beitrag von Frau Lepp und Herrn Polzin wurde die klassische itembasierte Evaluation diskutiert, hier mit dem interessanten Twist, die Lehrenden über die Verwendung der Items entscheiden zu lassen. Solche Itembatterien existieren bereits an recht vielen Universitäten (vgl. Lübcke, Favella, Schiefner-Rohs 2018)

Denn dies ist das einzige feststehende Ergebnis der Veranstaltung: Ein standardisiertes Instrument für alle Projekte forschenden Lernens an allen Hochschulen kann es nicht geben.

Forschendes Lernen ist extrem positiv mit dem Selbstverständnis der Hochschulen und der Hochschuldidaktiker_innen verwoben. Wir beschäftigen uns mit Fragen der Evaluation, um Ideen zur Verbesserung zu erhalten und unser Handeln zu legitimieren. Aber schon die Frage: „Vor wem rechtfertigen und für wen rechtfertigen?“ wird in der Konzept- und Perspektivendiskussion nicht gestellt.

Daher ist die Einnahme einer Metaperspektive über Evaluation umso wichtiger: Erfüllt die Evaluation die vier Grundstandards? Wer sind die intendend users? Welche Zwecke werden mit der Evaluation forschenden Lernens verfolgt?

Literatur:

Chen, T. (1990). *Theory-Driven Evaluations*. Newbury Park, Calif, London: Sage.

Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsí, T. A. (2000). Program theory evaluation: Practice, promise and problems. *New Directions for Evaluation*, 87, 5-13.

Giel, S. (2016). Programmtheorie in der Bildungsevaluation. In H. Marburger, C. Griese & T. Müller (Hrsg.), *Bildungs- und Bildungsorganisationsevaluation* (S. 105-122). Berlin: De Gruyter.

Lübcke, E. Heudorfer, A. (in Druck 2019): Die Ziele forschenden Lernens. Eine empirische Analyse im Rahmen der QPL-Begleitforschung. In: Reinmann, G. Lübcke, E. Heudorfer, A.: *Forschendes Lernen in der Studieneingangsphase - Empirische Befunde, Fallbeispiele und individuelle Perspektiven*. Springer VS. S. 17-48.

Lübcke, E., Favella, G., Schiefner-Rohs, M. (2018): Dokumentation der Sekundäranalyse im Projekt FideS.

Huber, L. (2014). Forschungsbasiertes, Forschungsorientiertes, Forschendes Lernen: Alles dasselbe? *Hochschulforschung*, 1+2, 22-29

Huber, Ludwig (2018): Was soll heißen „Exzellenz (in) der Lehre“? In: *Das Hochschulwesen*, 3+4/2018, S. 105-113

Tremp, Peter; Hildbrand, Thomas (2012): Forschungsorientiertes Studium – universitäre Lehre: Das «Zürcher Framework» zur Verknüpfung von Lehre und Forschung. In: Tobina Brinker & Peter Tremp (Hrsg.): *Einführung in die Studiengangentwicklung*. (= *Blickpunkt Hochschuldidaktik* 122). Bielefeld: Bertelsmann 2012, S. 101-116