

Wirkungsanalyse eines Lehrentwicklungsprojekts in zwei Jahren: Geht das?

René Krempkow, Angela Weißköppel

Abstract *Wirkungsanalysen und Evaluationen stehen im Spannungsverhältnis zwischen dem Einsatz wissenschaftlich-empirischer Methoden zur Datengewinnung und dem Anspruch an eine gewisse methodische Strenge einerseits und dem Erfordernis der Bereitschaft zu Kompromissen sowie pragmatischer Herangehensweise aufgrund kurzer Projektlaufzeit und weiterer Rahmenbedingungen andererseits. In diesem Beitrag wird der Umgang mit gegenläufigen Ansprüchen anhand der Vorstellung eines Evaluationskonzepts (dessen Entwicklung etwa ein Jahr nach Projektbeginn startete) und ausgewählter Ergebnisse reflektiert und diskutiert. Ziel der Evaluation und Wirkungsanalyse ist es u.a., digitales Lehren und Lernen an der HTW Berlin stärker als bisher evidenzbasiert weiterzuentwickeln. Hierbei sollen Maßnahmen und Prozesse formativ auf ihre Wirkungsorientierung hin reflektiert und überprüft werden.*

Impact analyses and evaluations are torn between the use of scientific-empirical methods for data collection and the demand for a certain methodological rigour on the one hand, and the need for a willingness to compromise and a pragmatic approach due to the short project duration and other framework conditions on the other. In this article, the handling of conflicting demands is reflected and discussed on the basis of the presentation of an evaluation concept (the development of which began about a year after the start of the project) and selected results. One of the aims of the evaluation and impact analysis is to further develop digital teaching and learning at HTW Berlin in a more evidence-based way than before. In doing so, measures and processes are to be formatively reflected upon and reviewed in terms of their impact orientation.

1. Einleitung

Die im Titel aufgeworfene Frage ist keine rhetorische. Sie stellt sich in dieser Form in einem Projekt im Feld der Lehrentwicklung, für das eine Evaluation und Wirkungsanalyse erfolgen soll. Dies steht im Spannungsverhältnis zwischen dem Einsatz wissenschaftlich-empirischer Methoden zur Datengewinnung und dem Anspruch an eine gewisse methodische Strenge (insbesondere für Wirkungsanalysen) einerseits und dem Erfordernis der Bereitschaft zu Kompromissen sowie pragmatischer Herangehensweise aufgrund kurzer Projektlaufzeit und weiterer Rahmenbedingungen andererseits.

In diesem Beitrag soll der Umgang mit solchen gegenläufigen Ansprüchen anhand der Vorstellung des Evaluationskonzepts für das Curriculum Innovation Hub und erster Ergebnisse reflektiert und diskutiert werden. Die Entwicklung des Konzepts für das Projekt mit dreijähriger Laufzeit startete aufgrund personeller Engpässe erst etwa ein Jahr nach Projektbeginn. Dabei stellen sich Evaluationskonzept und Erhebungsdesign auch dem Anspruch, Ansätze zur Erfassung von Wirkungszusammenhängen und zur Reflexion von Wirkungsvermutungen zu formulieren.

Die Evaluation und Wirkungsanalyse soll u.a. helfen, digitales Lehren und Lernen an der HTW Berlin stärker als bisher evidenzbasiert weiterzuentwickeln. Hierbei sollen Maßnahmen und Prozesse des Curriculum Innovation Hub formativ auf ihre Wirkungsorientierung hin reflektiert und überprüft werden. Dabei werden sowohl die Ebene des Gesamtprojekts als auch die Ebene der Teilprojekte bzw. Arbeitspakete einbezogen. Die formative Wirkungsüberprüfung soll das Curriculum Innovation Hub bzw. dessen Maßnahmen und Teilprojekte während der Durchführung konzeptionell begleiten und unterstützen sowie Rahmenbedingungen schaffen, die die Wirksamkeit der Maßnahmen und Teilprojekte wahrscheinlicher machen. Der Fokus der Wirkungsüberprüfung soll dabei insgesamt auf Zielen, Ergebnissen und Nachhaltigkeit liegen. Im vorliegenden Beitrag liegt der Fokus auf Zielen und der daraus abgeleiteten Entwicklung des Evaluationskonzepts sowie ersten Erfahrungen in der Umsetzung von Evaluationsaktivitäten.

Abschließend soll in diesem Beitrag diskutiert werden, inwieweit gewon-
nene erste Erfahrungen auch auf ähnliche Maßnahmen andernorts übertrag-
bar wären.¹

2. Projektziele

Im Folgenden werden – bezugnehmend auf die Zielklärung als einen der ersten Schritte der durchgeführten Evaluationsaktivitäten – zunächst die übergreifenden Ziele des Curriculum Innovation Hub insgesamt (Abschnitt 2.1) und in den darauffolgenden Abschnitten 2.2, 2.3 und 2.4 die Ziele von drei ausgewählten Teilprojekten des Curriculum Innovation Hub vorgestellt. Hierbei lehnt sich deren Formulierung einerseits naturgemäß stark an den bewilligten Projektantrag an, andererseits wurden aber überall dort, wo es möglich war, resultierend aus den im Rahmen der Evaluation durchgeführten Gesprächen mit den Vertreter*innen der Teilprojekte, die Ziele noch konkretisiert und zum Teil angepasst.

Aus den Teilprojektbeschreibungen bzw. im Gespräch mit den Teilprojekt-
leitungen und -mitarbeiter*innen wurden potenziell operationalisierbare Zie-
le herausgearbeitet, damit diese dann eine Basis für mögliche Evaluations-
maßnahmen sein können. Soweit möglich, werden in den nachfolgenden Text-
abschnitten zu den jeweiligen Teilprojekten auch bereits erfolgte Evaluations-
maßnahmen und/oder das geplante weitere Vorgehen bezüglich Evaluation
und ggf. Ansätzen für Wirkungsanalysen kurz beschrieben.

2.1 Ziele des Curriculum Innovation Hub

Durch die pandemiebedingte Ad-hoc-Digitalisierung analoger Formate ist die Schaffung wissenschaftlich fundierter, didaktisch optimierter, hochschulwei-
ter Konzepte in den Hintergrund getreten. Stattdessen fanden Anpassungen
bzw. Veränderungen von Lehrformaten als »emergency remote teaching«
(Hodges et al., 2020) statt. Das Curriculum Innovation Hub hat zum Ziel,
diese unter Zeitdruck entstandenen Konzepte sowie die übergreifenden Ver-
änderungen durch die anhaltende Online-Lehre unter Pandemiebedingungen

¹ Dieser Beitrag stellt eine erweiterte und aktualisierte Fassung einer früheren Ausar-
beitung dar, die für die Frühjahrstagung 2023 des AK Hochschulen der DeGEval – Ge-
sellschaft für Evaluation erstellt wurde.

aufzuarbeiten und fundiert konzeptionell weiterzuentwickeln, um nachhaltige Strukturen zu etablieren und so langfristig Studienerfolg zu gewährleisten. Das Projekt Curriculum Innovation Hub soll seinen Fokus auf der Ebene der Studienprogramme haben und durch eine didaktisch durchdachte, technisch und räumlich weiter untermauerte Basis an den Schnittstellen der Fachbereiche den Weg für zukunftsfähige blended Curricula bereiten. Auch nach der Pandemie soll das Präsenzstudium an der HTW Berlin anders aussehen. Innovative Konzepte sollen dafür im Curriculum Innovation Hub entwickelt und umgesetzt werden. Nach erfolgter Evaluation und Weiterentwicklung sollen diese Bestandteile dann in Studienprogrammen fest verankert werden.

Die geplanten Evaluationsaktivitäten setzen auf mehreren Ebenen an: Während das übergreifende Ziel konzeptioneller Weiterentwicklung anhand der entstandenen Konzepte überprüft werden kann, ist dies für die weiteren übergreifenden Ziele, nachhaltige Strukturen zu etablieren und langfristig Studienerfolg zu gewährleisten, deutlich schwieriger. Immerhin könnte zum Abschluss der Projektlaufzeit dargestellt (und somit z.B. nach Art einer Checkliste ggf. qualitativ überprüft) werden, ob es und wenn ja, welche Strukturveränderungen es aufgrund von Projektaktivitäten gab. Inwiefern diese nachhaltig sind, kann bei Strukturveränderungen allerdings nur schwer eingeschätzt werden, da selbst an Hochschulen Strukturen einem steten Wandel unterliegen. Dies kann daher frühestens einige Jahre nach Projektabschluss wirklich eingeschätzt werden.

2.2 Innovative digitale MINT-Curricula

Im Curriculum Innovation Hub sollen an der Schnittstelle der beiden ingenieurwissenschaftlichen Fachbereiche innovative digitale MINT-Curricula entwickelt werden. Dies bedeutet vor allem eine Weiterentwicklung von einem derzeit eher lehrendenzentrierten Fokus hin zu mehr Studierendenzentrierung und -beteiligung. Wichtige curriculare Innovationen, die dies befördern sollen, sind die Gestaltung durchdachter Flipped-Classroom-Modelle insbesondere in den Grundlagenfächern. Hier wurde geplant, dass die entwickelten digitalen und hybriden Lehr- und Prüfungsformate konkret z.B. durch Auswertung von Prüfungsdaten und gruppenspezifische Interviews sowie Befragungen evaluiert werden (siehe Gliederungspunkt 3.2 zu Erhebungsinstrumenten). So ist vorgesehen, für die Einführung digitalisierter Mathematik-Übungsaufgaben mit automatisierter Auswertung (vgl. hierzu Stromeyer & Zeiser, 2023) Wirkungsaussagen abzuleiten, indem die Studie-

renden zufällig einer Treatment- und einer Kontrollgruppe zugewiesen werden. So soll ermöglicht werden, die späteren Mathematik-Klausurergebnisse beider Gruppen zu vergleichen und in Gruppeninterviews subjektive Einschätzungen und Weiterentwicklungsvorschläge hierzu zu erfassen, sowohl vor als auch nach o.g. Maßnahme. Darüber hinaus soll die Entwicklung fachübergreifender und interdisziplinärer Kompetenzen als zentrale Idee des Curriculum Innovation Hub über die verschiedenen ingenieurwissenschaftlichen Studiengänge hinweg adressiert werden. Hierzu ist angedacht, dies mittels (ohnehin durchgeführter) Absolvent*innenbefragungen zu erheben (siehe Erhebungsinstrumente, Abschnitt 3.2).

2.3 Innovatives Lernraumdesign

Das Ziel des Teilprojekts Innovatives Lernraumdesign ist es, innovative Strategien zur räumlichen Gestaltung und hochschulweiten Umsetzung von Blended-Learning-Formaten für die Kombination von Präsenz- und Online-Lehre zu entwickeln. Dabei sollen didaktische, räumliche, technologische und auch organisatorische Perspektiven zusammengeführt werden und von Anfang an strategische Maßnahmen sowie relevante Organisationsstrukturen und -prozesse zur Skalierung und damit hochschulweiten Integration von nachhaltigen Raumarchitekturen und Ausstattungsszenarien für hybride Lehr-Lern-Formate mitgedacht werden (vgl. auch Geister et al., 2023). In einem mehrstufigen Verfahren werden innovative Settings für hybride Lehr-/Lernumgebungen entwickelt, modelliert, implementiert und evaluiert. Die beiden beteiligten Fachbereiche nehmen dazu Bedarfe und Anforderungen differenzierter Nutzer*innengruppen und Akteur*innengruppen zur Entwicklung und Modellierung physischer Lernumgebungen im Kontext digitaler Lehr-Lern-Formate auf, implementieren und evaluieren auf dieser Basis vier Modell-Lernräume mit unterschiedlichen räumlich-technologischen Settings sowie jeweils drei Arten von Lehr-/Lernszenarien. Zur Aufnahme von Bedürfnissen von Nutzer*innen (Lehrende und Studierende) bei hybriden und studierendenzentrierten Lernumgebungen erfolgen Leitfadeninterviews, um die Konzeptionierung der Modellräume nicht nur über den aktuellen Forschungsstand, sondern auch über spezifische Bedarfe an der HTW Berlin aufzubauen. Darüber hinaus finden quantitative und qualitative Nutzer*innenbefragungen (Lehrende und Studierende) statt, um die Modellräume zu evaluieren (siehe hierzu ausführlicher Abschnitt 4.2 in diesem Beitrag).

2.4 Wirkungsanalysen und Evaluation

Um digitales Lehren und Lernen stärker als bisher evidenzbasiert weiterzuentwickeln und um zugleich die Vorhaben auf ihre Wirkungsweisen hin systematisch zu untersuchen, soll im Rahmen des Curriculum Innovation Hub das Thema Wirkungsanalysen und Evaluation in einem weiteren Teilprojekt an der Schnittstelle des Zentralen Referats Hochschulentwicklung und Qualitätsmanagement (HE & QM) und des Lehrenden-Service-Centers intensiv konzeptuell weiterentwickelt werden. Die entwickelten digitalen und hybriden Lehr- und Prüfungsformate sollen durch flächendeckende Befragungen und gruppenspezifische Interviews evaluiert werden. Darüber hinaus ist die Weiterentwicklung der Evaluation des digitalen und hybriden modulbezogenen Lehr- und Prüfungserfolgs an der HTW Berlin unerlässlich.

3. Evaluationsdesign und potenziell einsetzbare Erhebungsinstrumente

3.1 Zum Evaluationsdesign

Ideal für eine Evaluation im Sinne von Wirkungsanalysen wäre ein Vorher-Nachher-Vergleich (Prä-Post-Design) mit Treatment- und Kontrollgruppen, was hierzu als eine Art Goldstandard gelten kann. Dies ist jedoch aufgrund der verfügbaren Ressourcen, datenschutzrechtlicher Rahmenbedingungen und des beschränkten zeitlichen Horizonts des Projekts an der HTW Berlin nur eingeschränkt und nur für bestimmte Teilprojekte des Curriculum Innovation Hub und deren konkrete Ziele umsetzbar. Allerdings lassen sich auch durch Kohortenvergleiche unter Umständen Wirkungsaussagen ableiten, z.B. bei gleicher Zusammensetzung der Kohorten bzw. bei statistischer Kontrolle der Zusammensetzung und ansonsten gleichen oder ähnlichen Kontextbedingungen (vgl. z.B. Döring & Bortz, 2016; Krempkow et al., 2018). Es soll daher geprüft werden, ob beispielsweise mittels Kohortenvergleichen oder Vergleichen von Treatment- vs. Kontrollsituation auf der Basis von Daten des Studiengangsmonitoring und/oder des Studienqualitätsmonitorings bzw. aggregierter Daten der Lehrevaluation, eigener Studierendenbefragungen oder Items in Absolvent*innenbefragungen Wirkungsaussagen abgeleitet werden können.

3.2 Mögliche Erhebungsinstrumente/Datenquellen

Idealerweise können für eine Evaluation verschiedene Erhebungsinstrumente so eingesetzt werden, dass ein Methodenmix möglich ist und sich die Erhebungsinstrumente (z.B. quantitative und qualitative) in ihren Stärken und Schwächen ergänzen (vgl. z.B. Kelle, 2019; Kuckartz, 2014; Rathke et al., 2023). Daher erfolgte für die Konzeptentwicklung zunächst eine Bestandsaufnahme der an der HTW Berlin für die vorgenannten Ziele verfügbaren und damit potenziell einsetz-/nutzbaren Erhebungsinstrumente bzw. Datenquellen, die nachfolgend in einer Übersicht aufgelistet werden. Die Erhebungsinstrumente bzw. Datenquellen decken zudem verschiedene Ebenen ab, weshalb ihre jeweilige Aussagekraft bei der Konzeption und Durchführung der Evaluation zu beachten ist (vgl. ausführlich Krempkow, 2007). Daher werden sie in der folgenden Tabelle in diese Ebenen eingordnet.

Tabelle 1: Einordnung der Erhebungsinstrumente/Datenquellen in eine Ebenen-Matrix

Mikroebene (v.a. einzelne Lehrende, ggf. Module)	Mesoebene (Studiengang/-fach, ggf. Fachbereich)	Makroebene (ganze Hochschule und ggf. Kontext)
–	Lehrendenbefragung	Lehrendenbefragung
–	Studiengangsmonitoring	Studiengangsmonitoring
–	Studienqualitätsmonitoring	Studienqualitätsmonitoring
Studierendenbefragung in Lehrveranstaltungen/ Lehrveranstaltungsevaluation	(falls möglich und sinn- voll aggregierte Daten von Studierendenbefragung/ Lehrveranstaltungs- evaluation)	(falls möglich und sinn- voll aggregierte Daten von Studierendenbefragung/ Lehrveranstaltungs- evaluation)
–	Absolvent*innenbefragung	Absolvent*innenbefragung
Leitfadeninterviews, z.T. Fokusgruppeninterviews	Fokusgruppeninterviews	–

Eigene Darstellung.

3.3 Auswahl der Erhebungsinstrumente/Datenquellen

Nachdem zunächst mögliche Erhebungsinstrumente bzw. Datenquellen aufgelistet und den entsprechenden Ebenen zugeordnet wurden, soll nun die Vorgehensweise bei der Auswahl für die Evaluation vorgestellt werden. Hierzu werden jeweils die konkreten Erhebungsmöglichkeiten sowie ihre Vor- und Nachteile vorgestellt und die daraus abgeleitete (geplante) weitere Vorgehensweise beschrieben.

Befragung von Lehrenden: Lehrendenbefragungen können auf verschiedenen Ebenen eingesetzt werden. An der HTW Berlin findet dies a) hochschulweit statt, wobei mehrere Fragenvorschläge aus dem Curriculum Innovation Hub einbezogen sind (so u.a. zur Einschätzung der Raumsituation); sowie b) als spezifische eigene Curriculum-Innovation-Hub-Befragung der Lehrenden in innovativen Lern-/Lehrräumen. Die hochschulweite Befragung (a) war bereits seit längerem geplant, sodass es als glücklicher Umstand gesehen werden kann, dass noch mehrere Fragen aus dem Curriculum Innovation Hub integriert werden konnten. Allerdings mussten diese dementsprechend knapp ausfallen und konnten nicht detaillierter auf Spezifika einzelner Projektmaßnahmen eingehen. Daher bietet sich hierfür eine spezifische eigene Curriculum-Innovation-Hub-Befragung der Lehrenden an (b), wie sie im Teilprojekt zu innovativen Lernräumen erfolgt. Hierbei kann detailliert auf Maßnahmen eingegangen werden; ein repräsentatives Bild der Sicht aller Lehrenden an der HTW Berlin kann so aber wiederum nicht erstellt werden. Deshalb ergänzen sich die auf den verschiedenen Ebenen als Erhebungsinstrumente eingesetzten Lehrendenbefragungen gut. Da die hochschulweite Lehrendenbefragung durch eine externe Organisation durchgeführt wurde, entfällt hier der Erhebungsaufwand für das Curriculum Innovation Hub, sodass beider Einsatz als durchführbar eingeschätzt werden konnte. Allerdings entfällt hierbei nicht der Aufwand für die Auswertung, Aufbereitung und Verbreitung der entsprechenden Ergebnisse.

Aufbereitung Studiengangsmonitoring: Hierfür ist via Datencockpit »Cognos« der Zugang zu Berichten/aggregierten Daten (v.a. Hochschulstatistik-Kennzahlen) der HTW Berlin möglich, sodass beispielsweise der Prüfungserfolg von Studiengängen bzw. -fächern semesterweise verfolgt und verglichen werden könnte. Sofern es möglich ist, Curriculum-Innovation-Hub-Maßnahmen einzelnen Fächern und Semestern zuzuordnen, wäre es eventuell möglich (unter Berücksichtigung weiterer Randbedingungen), den Studienerfolg mit diesen Maßnahmen in Zusammenhang zu bringen und so deren Effekte zu un-

tersuchen oder zumindest plausible Wirkungsvermutungen abzuleiten. Darüber hinaus wäre es wünschenswert, Prüfungsergebnisse auch auf Individualebene der Studierenden mit bestimmten Projektaktivitäten in Zusammenhang zu bringen, beispielsweise für das Teilprojekt Innovative Digitale MINT-Curricula. Dies würde aber voraussetzen, dass eine datenschutzkonforme Lösung hierfür auf Individualebene gefunden wird. Während Letzteres – auch aufgrund der begrenzten Projektlaufzeit – bisherigen Erfahrungen zufolge als eher unwahrscheinlich gelten muss, könnte das Erstgenannte durchaus möglich sein und wird daher für den Einsatz als Erhebungsinstrument weiter vorgesehen.

Aufbereitung Studienqualitätsmonitoring/(DZHW-)Studierendenbefragung: Die bundesweite (DZHW-)Studierendenbefragung findet üblicherweise alle vier Jahre statt. Sie hat 2020 und – außer der Reihe – 2021 (als Corona-Semesterbefragung) stattgefunden, ist also erst wieder für 2026 vorgesehen. Daher ist hier eine Integration von Fragenvorschlägen aus dem Curriculum Innovation Hub nicht möglich. Hier wären höchstens ggf. aggregierte Daten als Hintergrundinfo nutzbar – z.B. zur Frage, wie sich die Ergebnisse im Zeitverlauf veränderten (vgl. für die zu prüfenden Randbedingungen für einen solchen Ergebnisvergleich Krempkow, 2003), z.B. im Zuge der Pandemie. Da auf diese Weise aber kein Zusammenhang mit Maßnahmen des Curriculum Innovation Hub hergestellt werden kann, wird dies nicht weiterverfolgt.

Lehrevaluation bzw. Studierendenbefragung in Lehrveranstaltungen: Hier wäre es erstens möglich, für einzelne Lehrende individuelle Lehrevaluations-(Zusatz-)Fragen vorzuschlagen. Zweitens können aggregierte Daten als Hintergrundinfo genutzt werden: Beispielsweise könnte – da bei ausschließlicher Durchführung von Lehrendenbefragungen oft ein Außenkriterium zur Einordnung ihrer Ergebnisse fehlt – eine Gegenüberstellung mit Ergebnissen von Studierendenbefragungen erfolgen. Auf diese Weise können verschiedene Perspektiven auf die Lehre erfasst und ein umfassenderes Gesamtbild der Perspektiven auf die Lehre ermöglicht werden. Darüber hinaus ist es drittens möglich, spezifische eigene Curriculum-Innovation-Hub-Befragungen z.B. zu ausgewählten Lehrveranstaltungen in (innovativen) Lern-/Lehrräumen zur Einschätzung dieser Räume durchzuführen. Aufgrund des hohen Koordinationsaufwandes wird die erstgenannte Möglichkeit zurückgestellt und nicht prioritär verfolgt. Dagegen können letztere beide Möglichkeiten aufgrund der bereits erfolgten Datenerhebung bzw. aufgrund der in Kooperation mit dem entsprechenden Teilprojekt erfolgenden Datenerhebung relativ ressourcenschonend umgesetzt werden und wurden daher angegangen.

Absolvent*innenbefragung: Diese finden turnusmäßig an der HTW Berlin alle zwei Jahre statt. Zuletzt hat sie Anfang 2024 stattgefunden; d.h. es wird zeitlich eine Herausforderung, aber bei entsprechender Vorbereitung nicht unmöglich, dazu einige Items bezüglich Curriculum Innovation Hub(-Zielerreichung) auszuwerten. Auf diese Weise können – anders als mittels Studierendenbefragungen möglich – ggf. auch einzelne längerfristige Effekte des Studiums (vgl. Krempkow, 2007; Krempkow, 2009) bzw. Projektmaßnahmen in den Blick genommen werden. Für darüber hinausgehende langfristige Effekte können die Daten aus der Absolvent*innenbefragung allerdings voraussichtlich erst nach Projektende zur Verfügung gestellt werden. Deshalb wurde hierzu mit der durchführenden Einrichtung (HE & QM) der HTW Berlin besprochen, dass die Ergebnisse ggf. im Rahmen der standardisierten Auswertung in den Folgejahren mit ausgewertet und zur Verfügung gestellt werden sollen.

Leitfaden-/Fokusgruppeninterviews: Geplant ist hierzu sowohl die (Nach-)Nutzung der Auswertungen bereits durchgeföhrter Interviews in den Teilprojekten wie z.B. Innovatives Lernraumdesign (auch um eine Überlastung der Beteiligten durch übermäßig viele Erhebungen zu vermeiden) als auch ggf. ergänzende eigene Interviews, z.B. als leitfadengestützte Expert*inneninterviews (nach Meuser & Nagel, 1994; 2003; mit themenzentrierter Auswertung angelehnt an Witzel, 1995). Bei solchen qualitativen Interviews kann viel stärker auf die konkrete Situation eingegangen werden und ggf. Verständnisfragen direkt im Interview geklärt werden – anders als in standardisierten quantitativen Erhebungen (wobei oft, z.B. in Online-Befragungen, Ratingskalen verwendet werden, die generell das Problem einer geringeren Tiefenschärfe mit sich bringen). Leitfadeninterviews können bei relativ freiem Erzählfluss eher die Nähe zum Untersuchungsgegenstand gewährleisten und bleiben durch den Leitfaden bzw. gleichlautende Stimuli dennoch grundlegend vergleichbar. Allerdings können meist nur standardisierte (Online-)Befragungen so erhebungsökonomisch durchgeführt werden, dass repräsentative Aussagen zur Grundgesamtheit möglich sind. Qualitative Interviews und standardisierte Befragungen werden hier mit dem Ziel gegenseitiger Ergänzung eingesetzt. Leitfadengestützte Fokusgruppeninterviews kommen zum Einsatz, wenn es als vorteilhaft anzusehen ist, dass sich die Interviewten in einer relativ natürlichen Gesprächssituation austauschen (vgl. z.B. Ruddat, 2012; Tausch & Menold, 2015), z.B. im Teilprojekt Innovative Digitale MINT-Curricula (siehe hierzu auch nachfolgende Ausführungen zu ersten Erfahrungen).

Studienabschlussbefragung: Diese war ursprünglich angedacht als ein Erhebungsinstrument, das zu einem noch früheren Zeitpunkt einen Rückblick auf das gesamte Studium ermöglichen kann (als Absolvent*innenbefragungen, die standardmäßig das erste Mal ein bis zwei Jahre nach Abschluss erfolgen – vgl. z.B. Krempkow, 2007). Allerdings ist die Durchführung einer Studienabschlussbefragung an der HTW Berlin nicht möglich und wurde daher frühzeitig verworfen.

Sounding Board: Ein zusätzlich vorgesehenes Instrument für die Curriculum-Innovation-Hub-Evaluation (allerdings kein Erhebungsinstrument, daher nicht in Tabelle 1 aufgeführt) ist das eigens dafür geschaffene Sounding Board. Dies ist ein beratendes Gremium für die fachliche und strukturelle Weiterentwicklung des Curriculum Innovation Hub, das hierfür aktiv Meinungen von Betroffenen und Beteiligten einholt und Entwicklungsimpulse setzt. Als Größe für eine möglichst gute Arbeitsfähigkeit wurden acht Personen festgelegt. Diese Zusammensetzung erfolgte überwiegend aus HTW-internen Personen, nämlich mit fünf Personen aus den fünf HTW-Fachbereichen für eine Stärkung der Verankerung in diesen, sowie einer*r Studierenden aus dem AStA (der Studierendenvertretung). Hinzu kamen zwei HTW-externe Expert*innen für die Einbindung externer Perspektiven und Impulse. Die konkrete Arbeit sieht hierbei wie folgt aus: Im Vorfeld der etwa jährlich stattfindenden Sounding Board Meetings werden alle Teilprojekte (neben der Vorbereitung eines kurzen Impulses zum aktuellen Stand) darum gebeten, jeweils zwei bis drei Sätze zu einer Frage oder einem Problem zu übersenden, die/das den jeweiligen Teilprojekten auf den Nägeln brennt. Die Mitglieder des Sounding Board erhalten dies ca. zwei Wochen vorher mit der Bitte, hierfür Ideen, Anregungen oder ggf. Rückfragen vorzubereiten, die im Meeting diskutiert werden können. Im online durchgeführten Meeting stellt jedes Teilprojekt in einem festen Zeitslot von maximal zehn Minuten die wichtigsten Projektergebnisse und den Stand im Zeitplan vor und erläutert kurz, welche Probleme ggf. aufgetreten sind und welche eigenen Lösungsideen es eventuell bereits gibt.

4. Erfahrungen und Ergebnisse

4.1 Anforderungen und Realisierungsmöglichkeiten)

Ein Evaluationsdesign im Sinne von Wirkungsanalysen mit Treatment- und Kontrollgruppe sowie als Prä-Post-Design lässt sich aus den eingangs be-

reits erwähnten Gründen in diesem wie in ähnlich gelagerten Projekten nur für bestimmte Ziele umsetzen (vgl. zu Anforderungen und Realisierungsmöglichkeiten von Wirkungsanalysen auch Standards der DeGEval, 2016; Schmidt & Jongmanns, 2021). Dies findet eine Begründung auch im Kriterium der »Durchführbarkeit« von Evaluationen (DeGEval, 2016, S. 22), wonach standardmäßig zu prüfen ist, dass »der Aufwand für die Beteiligten und Betroffenen in einem adäquaten Verhältnis zum intendierten Nutzen der Evaluation gehalten wird.«

Hier sollen nur exemplarisch einige Evaluationsmaßnahmen vorgestellt werden, in denen Lehrformate konkret evaluiert werden, z.B. durch spezifische digitale Test-Übungsaufgaben und Fokusgruppeninterviews sowie Befragungen.

So ist im Teilprojekt »Innovative digitale MINT-Curricula« geplant, für die Einführung digitalisierter Mathematik-Übungsaufgaben Wirkungsaussagen abzuleiten, indem die Studierenden zufällig einer Treatment- und einer Kontrollgruppe zugeteilt werden und unterschiedliche Test-Übungsaufgaben digital bearbeiten. So soll ermöglicht werden, die Ergebnisse beider Gruppen zu vergleichen und zusätzlich in Fokusgruppeninterviews subjektive Einschätzungen und Weiterentwicklungsvorschläge hierzu zu erfassen. Ein erster Durchlauf der Erhebungen zur Evaluation hierzu hat in einer Mathematik-Lehrveranstaltung für Ingenieure bereits stattgefunden und die Erfahrungen sind positiv. Daher ist dies ähnlich auch für die Informatik geplant. Ein (wegen der weiteren geplanten Evaluationsaktivitäten noch vorläufiges) Ergebnis ist in diesem konkreten Fall die daraus abgeleitete Überlegung, bei digitalen Mathematik- (bzw. Informatik-)Übungsaufgaben an den jeweils bewältigten Schwierigkeitsgrad angepasstes Feedback zu geben und auch die Testaufgaben noch entsprechend weiterzuentwickeln, um den Lern-/Lehrerfolg bestmöglich zu fördern. So könnte eine potenzielle Stärke von digitalen Übungsaufgaben mittels entsprechenden Feedbacks noch stärker zum Tragen kommen, nämlich dass sie in besonderer Weise auch zum Selbstlernen geeignet sind. Zur Überprüfung auch von möglichen Effekten einer solchen Selbstlernförderung war ursprünglich eine Verknüpfung der Ergebnisse der digital bearbeiteten Test-Übungsaufgaben mit den Ergebnissen der später ohnehin standardmäßig im Studienablauf durchgeführten Mathematikklausur angedacht (z.B. pseudonymisiert anhand der Matrikelnummer). So sollte es möglich sein, auch längerfristige Effekte in der Treatment- und der Kontrollgruppe zu erfassen, dies scheiterte allerdings u.a. an Datenschutzbedenken.

Im Teilprojekt »Innovatives Lernraumdesign« kommen stärker befragungsbasierte Erhebungsinstrumente zum Einsatz. So haben quantitative und qualitative Nutzer*innenbefragungen (Lehrende und Studierende) stattgefunden, um die Modellräume zu evaluieren. Die Durchführung war entsprechend den personellen Ressourcen und technischen Unterstützungs möglichkeiten (Survey-Tool) nicht nur zwischen dem Teilprojekt »Innovatives Lernraumdesign« und dem Teilprojekt »Wirkungsanalysen und Evaluation« abzustimmen, sondern auch mit dem Zentralen Referat Hochschulentwicklung und Qualitätsmanagement (HE & QM). Da ein Bedarf zur Unterstützung insbesondere für quantitative Nutzer*innenbefragungen mittels Survey-Tool bestand, lag der Fokus der Wirkungsanalysen und Evaluation hier auf den quantitativen Lehrenden- und Studierendenbefragungen und auf der Abstimmung mit HE & QM. Hierzu soll im folgenden Abschnitt 4.2 die Vorgehensweise vorgestellt werden, die – wiederum exemplarisch – auch Verknüpfungsmöglichkeiten von Erhebungsmethoden und Datenquellen aufzeigt und die eventuell einen Mehrwert auch für andere Hochschulen darstellen kann.

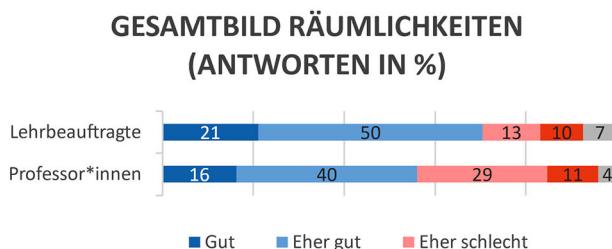
Zunächst sollen aber noch kurz die bisherigen Erfahrungen mit dem Instrument Sounding Board vorgestellt werden: Sie sind als positiv einzuschätzen, denn es konnten mehrere von den Meeting-Teilnehmenden als konstruktiv empfundene Anregungen für die weitere Arbeit entwickelt werden. Dies entspricht den eingangs genannten Zielen der (überwiegend formativen) Evaluation. Durch die online erfolgte Durchführung sowie die Begrenzung auf insgesamt drei Stunden konnte auch ein günstiges Aufwand-Nutzen-Verhältnis realisiert werden. Als besonders wichtig für das Gelingen erwies sich dabei (neben der Aufbereitung der Präsentationen so, dass auch Fachfremde alles nachvollziehen und ein plastisches Bild bekommen können) – in diesem wie auch in ähnlich gelagerten Gremien – die Meetingmoderation.

4.2 Erste Ergebnisse

In einem ersten Schritt wurden für eine hochschulweite Lehrendenbefragung im Herbst 2022 einige Fragen als Zusatzfragen gemeinsam mit dem Teilprojekt »Innovatives Lernraumdesign« entwickelt und im Winter 2022/23 erfolgreich eingesetzt (so z.B. zur Raumsituation an der HTW). Ein interessantes Ergebnis hierzu ist, dass die Einschätzung der Raumsituation an der Hochschule insgesamt zwischen den Fachbereichen nur wenig variiert und es keine signifikanten Differenzen zwischen den Fachbereichen gibt. Hierzu ist als

Hintergrund nützlich zu wissen: Die Fachbereiche sind in der Befragung mit nur geringfügig von der Grundgesamtheit abweichenden Anteilen gut repräsentiert, sodass die Befragung diesbezüglich als repräsentativ einzuschätzen ist. Auch die Gesamtfallzahl war mit 306 Befragten und einer Rücklaufquote von 30 % für eine Online-Befragung zufriedenstellend. Allerdings sind Lehrbeauftragte mit 22 %-Punkten Differenz (im Vergleich zum Anteil in der Grundgesamtheit) deutlich unterrepräsentiert. Das heißt, die Befragungsergebnisse spiegeln, obwohl je etwa hundert Befragte der jeweiligen Statusgruppe zuzurechnen sind, überproportional die Professor*innen-Sicht wider. Dies ist auch wichtig für die Einordnung eines weiteren interessanten Ergebnisses, denn es zeigt sich: Während Lehrbeauftragte zu 71 % die Raumsituation insgesamt als gut oder eher gut einschätzen, trifft dies für nur 56 % der Professor*innen und damit signifikant seltener zu (s. Abb. 1).

Abbildung 1: Bewertung der Raumsituation an der HTW Berlin



Eigene Darstellung, Datenbasis: N=306, HTW-Lehrendenbefragung, 2023.

Es kann vermutet werden, dass dies z. B. daran liegt, dass den Lehrbeauftragten die Raumsituation aufgrund seltenerer Anwesenheit weniger wichtig ist. Dies lässt sich im Nachhinein, da nicht erfragt, jedoch nicht empirisch prüfen. Das heißt, dass die Ausgangssituation von den Lehrenden zwar überwiegend positiv eingeschätzt wird, erhebliche Anteile – insbesondere der Professor*innen – aber noch Verbesserungspotenzial sehen.

In jedem Fall wurde eine mögliche Schlussfolgerung aus diesen Ergebnissen, bei großer Nachfrage nach bestimmten Räumen – z. B. neu eingerichteten innovativen Lern-/Lehrräumen – deutlich häufiger vor Ort befindliche Profes-

sor*innen bei der Raumvergabe zu priorisieren, als solche zur Diskussion gestellt.

Neben standardisierten Ratingfragen bietet eine Online-Befragung aber den Vorteil, Freitextantworten zu erheben, um Verbesserungspotenziale zu erfassen. So erfolgte dies auch zur Raumsituation. Hierbei sprachen Unzufriedene am häufigsten eine zu starke Ausrichtung auf Frontallehre an, z.B.: »praktisch keine Flexibilität für Gruppenarbeiten, Fishbowl-Formate etc.« »i. W. muss Möblierung mobiler/flexibler werden, um Projektarbeiten in Kleingruppen zu fördern. Als Vorbild können viele Co-Working-Spaces in Berlin dienen.« »Flexiblere Raum-/Tischgestaltung, bessere Buchbarkeit passender Räume«, »flexible Bestuhlung und bessere Möglichkeiten für hybride Lehrveranstaltungen«, »Modulare Räume für Frontalelemente und Gruppenarbeit in einer Räumlichkeit«, »offenere Raumkonzepte«, um typische Antworten hierzu aufzuzeigen. Diese Ergebnisse bestätigen – nun auch mittels der breit angelegten hochschulweiten Lehrendenbefragung – im Wesentlichen die grundsätzliche Ausrichtung der innovativen Lernraumgestaltung auf stärker interaktive Möglichkeiten der Lehre, wie sie auch bereits aufgrund des aktuellen Forschungsstandes und der Leitfadeninterviews zur Bedarfserhebung inzwischen durch das Teilprojekt Innovative Lernraumgestaltung mit der Einrichtung von vier Modell-Lernräumen angegangen wurde.

Während die in die hochschulweite Lehrendenbefragung integrierten zusätzlichen Fragen zur Raumsituation vor allem der Erfassung der Ausgangssituation dienten, liegt der Hauptfokus bei der Evaluation der neu eingerichteten Modellräume auf der Erfassung ihrer Einschätzung mittels Nutzer*innenbefragungen (Studierende und Lehrende). Hierzu wurden mittels der Befragungssoftware Evasys entsprechende spezifische Fragebögen erstellt, die bewusst als Paper-Pencil-Befragung konzipiert und eingesetzt wurden – mit dem Ziel, die Rücklaufquote möglichst hochzuhalten. Tatsächlich konnte mit 280 ausgefüllten Fragebögen eine erfreulich hohe Fallzahl realisiert werden. Auch für die speziell auf die Modellräume zugeschnittene eigene Lehrendenbefragung liegen mit 47 ausgefüllten Fragebögen ausreichend hohe Fallzahlen vor, um Häufigkeitsverteilungen und Mittelwerte berechnen und so auch die Perspektive der Lehrenden darstellen zu können (vgl. z.B. Döring & Bortz, 2016; Döring, 2023).

Darüber hinaus sollen ggf. noch ergänzend qualitative Lehrendeninterviews genutzt und für die Evaluation herangezogen werden, um bestimmte

Aspekte zu erfassen, die nicht in die Lehrendenbefragung integriert werden konnten.

5. Mögliche Schlussfolgerungen und Übertragbarkeit der Erfahrungen

Was lässt sich nun aus den ersten Erfahrungen und Ergebnissen an möglichen Schlussfolgerungen ableiten? Unseres Erachtens lässt sich Folgendes festhalten, was sich gegebenenfalls auch auf ähnliche Projekte an anderen Hochschulen bzw. ähnliche Evaluationen übertragen ließe:

Auch wenn kein Vorher-Nachher-Vergleich (im Prä-Post-Design) mit Treatment- und Kontrollgruppen im Sinne eines Goldstandards möglich ist, sind dennoch Vergleiche von Treatment- vs. Kontrollsituation mittels Befragungen durchführbar.

Für eine breite(re) Einschätzbarkeit von Bedarfen auch unabhängig vom Projektzusammenhang bzw. darüber hinaus hat sich der Einsatz von einigen zusätzlichen Fragen in einer hochschulweiten (Lehrenden-)Befragung bewährt. Dies hatte für die mit der Projektevaluation betraute (Teilzeit-)Stelle zudem den Vorteil, dass dadurch weniger Projektressourcen gebunden wurden, als bei (weiteren) eigens durchgeführten Befragungen notwendig wäre.

Eine Bestandsaufnahme der für die Ziele der Evaluation verfügbaren und damit potenziell einsetz-/nutzbaren Erhebungsinstrumente bzw. Datenquellen hat sich bei der Konzeptentwicklung bewährt (wie auch für den Einsatz einiger zusätzlicher Fragen): Idealerweise können sie im Sinne eines Methodenmix so kombiniert werden, dass sich z.B. quantitativ und qualitativen Erhebungsinstrumente in ihren Stärken und Schwächen ergänzen. Dies hat sich teilweise bereits schon bei der Interpretation von Ergebnissen als nützlich erwiesen, obwohl noch nicht alle Daten für eine umfassendere Gesamtbetrachtung über alle Datenquellen vorliegen.

Als hinderlich, um auch längerfristige Effekte zu erfassen (z.B. mittels Zusammenführen eigens erhobener Evaluationsdaten mit Daten von Prüfungen), erwiesen sich neben Zeit- und Ressourcenknappheit insbesondere Datenschutzbedenken (trotz geplanter Pseudonymisierung). Hier wäre es hilfreich, wenn Datenschutzbeauftragte künftig in stärkerem Maße auch konkrete Lösungsvorschläge für erkannte Datenschutzprobleme bzw. -bedenken mitteilen würden. Dies könnte z.B. auch mittels entsprechender

Handreichungen geschehen, die dann ggf. für ähnliche Anwendungsfälle in Projekten an anderen Hochschulen bzw. ähnlichen Evaluationen adaptierbar sein könnten.

Literaturverzeichnis

- DeGEval (2016). *Standards für Evaluation (DeGEval-Standards). Erste Revision auf Basis der Fassung 2002. Verabschiedet durch die Mitgliederversammlung der DeGEval e. V. am 21. September 2016 (Langfassung)*. https://www.degeval.org/fileadmin/Publikationen/DeGEval_Standards_fuer_Evaluation_-_Erste_Revision__2016_.pdf
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer VS.
- Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer VS.
- Geister, S., Ninnemann, K. & Schneidt, J. (2023). Bist du angekommen an deiner Hochschule? Die Bedeutung informeller Lernumgebungen für die soziale Integration und das Well-Being von Studierenden. In K. Sachse & S. Kurzenhäuser-Carstens (Hg.), *Krisen meistern, Ressourcen aufbauen, Nachhaltigkeit stärken. Abstracts zur 26. Fachtagung der Gesellschaft für angewandte Wirtschaftspsychologie (GWPs)* (S. 26). Pabst Science Publishers.
- Hodges, C., Moore, S., Lockee, B., Trust, T. & Bond, A. (2020). *The Difference Between Emergency Remote Teaching and Online Learning*. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>
- Kelle, U. (2019). Mixed Methods? In N. Baur & J. Blasius (Hg.). *Handbuch Methoden der empirischen Sozialforschung* (S. 159–172). Springer.
- Kloke, K. & Krücken, G. (2012). Sind Universitäten noch lose gekoppelte Organisationen? Wahrnehmung und Umgang mit Zielkonflikten an deutschen Hochschulen aus der Organisationsperspektive unter besonderer Berücksichtigung der akademischen Lehre. In F. G. Becker, G. Krücken & E. Wild (Hg.), *Gute Lehre in der Hochschule. Wirkungen von Anreizen, Kontextbedingungen und Reformen* (S. 13–29). Bertelsmann.
- Krempkow, R. (2009). Von Zielen zu Indikatoren – Versuch einer Operationalisierung für Lehre und Studium im Rahmen eines Quality Audit. *Qualität in der Wissenschaft (QiW)*, 3(1), 44–53.

- Krempkow, R. (2007). *Leistungsbewertung, Leistungsanreize und die Qualität der Hochschullehre. Konzepte, Kriterien und ihre Akzeptanz*. Universitätsverlag Webler.
- Krempkow, R. (2003). Bessere Didaktik trotz schlechterer Rahmenbedingungen. Ergebnisse zweier Befragungen zur Situation von Lehre und Studium an der TU Dresden. *Zeitschrift für Evaluation – ZfEv* 2(2), 257–277.
- Krempkow, R., Mayweg, E. & Reinecke, K. (2018). *Improving Teaching and Learning and the Evaluation of Effects. Contribution to the EAIR 40th Annual Forum Budapest 2018 (Hungary), 26.-29.08.2018*. <https://www.researchgate.net/publication/327581604>
- Kuckartz, U. (2014). *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*. Springer.
- Meuser, M. & Nagel, U. (1994). Expertenwissen und Experteninterview. In R. Hitzler, A. Honer & C. Maeder (Hg.), *Expertenwissen. Die institutionalisierte Kompetenz zur Konstruktion von Wirklichkeit* (S. 180–192). Westdeutscher Verlag.
- Meuser, M. & Nagel, U. (2003). Das Experteninterview – Wissenssoziologische Voraussetzungen und methodische Durchführung. In B. Friebertshäuser & A. Prengel (Hg.), *Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft* (S. 481–491). Juventa.
- Rathke, J., Krempkow, R. & Janson, K. (2023). Kompetenzen im Wissenschaftsmanagement. In R. Krempkow, E. Höhle & K. Janson (Hg.), *Karriere im Wissenschaftsmanagement?* Universitätsverlag Webler. <https://doi.org/10.53183/9783946017318>
- Ruddat, M. (2012). Auswertung von Fokusgruppen mittels Zusammenfassung zentraler Diskussionsaspekte. In M. Schulz, B. Mack & O. Renn (Hg.), *Fokusgruppen in der empirischen Sozialwissenschaft. Von der Konzeption bis zur Auswertung* (S. 195–206). VS Verlag für Sozialwissenschaften.
- Schmidt, U. & Jongmanns, G. (2021). Intendierte und evidente Wirkungen. Förderprogramme für die Qualität in Studium und Lehre evaluieren. In R. Kordts-Freudinger, N. Schaper, A. Scholkmann & B. Szczyrba (Hg.), *Handbuch Hochschuldidaktik* (S. 615–628). utb/wbv Media.
- Stromeyer, L. & Zeiser, A. (2023). Von den mathematischen Fertigkeiten zur Anwendung: digitale Aufgaben und digitale Labore. In E. Liebscher, R. Hübl, J. Mecker & B. Wacker (Hg.), *Digitale Lehre im Rahmen der Grundausbildung in MINT-Fächern an Hochschulen: Didaktische Integration von digitalen Medien und E-Learningsystemen in Lehrveranstaltungen. Tagungsband 2022* (S. 159–177). Merseburg. <https://doi.org/10.25673/103431>

- Tausch, A. & Menold, N. (2015). Methodische Aspekte der Durchführung von Fokusgruppen in der Gesundheitsforschung. Welche Anforderungen ergeben sich aufgrund der besonderen Zielgruppen und Fragestellungen? In GESIS – Leibniz-Institut für Sozialwissenschaften (Hg.), *GESIS Papers 2015/12*.
- Witzel, A. (1995). Auswertung problemzentrierter Interviews: Grundlagen und Erfahrungen. In R. Strobl & A. Böttger (Hg.), *Wahre Geschichten? Zur Theorie und Praxis qualitativer Interviews* (S. 49–75). Nomos.

