



Black Box Bewertung. Gängige Bewertungspraktiken von Hochschullehrenden bei der Beurteilung wissenschaftlicher Arbeiten

KATHARINA RESCH & ANNA SCHOPF

Zusammenfassung

Bewertungen sind in der Gesellschaft allgegenwärtig und durchziehen auch den Hochschulsektor, etwa in Rankings oder Begutachtungsverfahren. Über die Bewertungspraktiken im Zuge von wissenschaftlichen Abschlussarbeiten aus der Perspektive der Lehrenden an Fachhochschulen und Universitäten ist hochschulübergreifend noch wenig bekannt. Bewertungspraktiken werden oft als „intime Praxis“ verstanden und nicht hochschulübergreifend diskutiert (Black Box Bewertung). Der vorliegende Beitrag rückt das wissenschaftliche Schreiben aus der Perspektive gängiger Bewertungspraktiken von Hochschullehrenden in den Blick. Der Beitrag geht der Frage nach, wie gängige Bewertungspraktiken von Hochschullehrenden an österreichischen Fachhochschulen und Universitäten differenziert werden können. Hierzu werden vier Differenzlinien zwischen den Hochschultypen Fachhochschule und Universität aufgegriffen: (1) interne und externe Betreuung der wissenschaftlichen Arbeit, (2) prozessorientierte und ergebnisorientierte Bewertung der wissenschaftlichen Arbeit, (3) offene und geschlossene Gutachten und (4) inhaltliche und formale Bewertungskriterien. Der Beitrag leistet einen Beitrag zur Weiterentwicklung der Hochschuldidaktik und bietet einen Ausblick auf künftige Forschungsfragen sowie Ansätze zur Organisationsentwicklung im Bereich des wissenschaftlichen Arbeitens.

Schlüsselwörter: Hochschuldidaktik; Bewertungspraktiken; Soziologie der Bewertung; wissenschaftliche Abschlussarbeiten; Gutachten

Black Box Assessment. Current thesis assessment practices of lecturers in higher education

Abstract

Assessments are omnipresent in today's society and they also affect the higher education sector, as for example in rankings or assessment practices. Furthermore, not much is known about the assessment practices of lecturers at universities and universities of applied sciences from a comparative perspective. Assessment practices are oftentimes treated as “intimate practice” and are not discussed on an institutional level (Black Box Assessment). The following contribution sheds light on scientific writing from the perspective of current assessment practices of lecturers in higher education. This

contribution raises the question of how these different assessment practices can be described comparing universities and universities of applied sciences. For this reason, we discuss four main differences: (1) internal and external supervision, (2) process-oriented and result-oriented supervision, (3) open and criteria-based reports, and (4) content-related and formal assessment criteria. This paper contributes to higher education didactics and offers potential for organisational development and future research questions.

Keywords: Higher education didactics; assessment practices; sociology of assessment; bachelor and master theses; reporting

1 Einleitung

Wissenschaftliches Arbeiten zu lehren, bedeutet für Lehrende, einerseits den Prozess des wissenschaftlichen Denkens und Schreibens anzuleiten und andererseits am Ende die Bewertung dieser wissenschaftlichen Arbeiten vorzunehmen (Peters & Dörfler, 2019). Der Umfang dieser Aufgaben, die der hochschulischen Lehre zuzuordnen sind, hat seit der Umstellung der Studienpläne auf das Bologna-System zugenommen. Studierende verfassen je nach Hochschulsektor mehrere wissenschaftliche Abschlussarbeiten: ein bis zwei Bachelorarbeiten, eine Masterarbeit und eine Dissertation. Die Prüfungs- und Bewertungspraxis für wissenschaftliche Arbeiten ist indes kaum standardisiert (Schlichte & Sievers, 2015). Es gibt deutliche Unterschiede in der Bewertungspraxis der Lehrenden in Bezug auf deren Rollenverständnis, die Vorgaben durch die jeweilige Prüfungsordnung und die Standards wissenschaftlichen Arbeitens (*consistency of practice*, Macfadyen et al., 2019). Empirische Studien, die Einblicke in die Bewertungspraxis der Hochschullehrenden zulassen, fehlen weitgehend. So sparen Handbücher der Hochschuldidaktik dieses Thema in der Regel aus, wie etwa Zumbach (2016) oder Kauffeld und Othmer (2019), die keinen Bezug zu Abschlussarbeiten herstellen. Die Bewertungspraxis an Hochschulen kann daher als „Black Box“ verstanden werden. Wenngleich es einige Ratgeber für Studierende gibt, die über Betreuerwahl, Themenfindung und den Schreibprozess aufklären, so haben viele Studierende keinerlei Informationen über die konkreten Bewertungspraktiken, die ihre Betreuenden für ihre wissenschaftlichen Arbeiten anwenden: „The examination process is still hidden behind closed doors.“ (Golding et al., 2014, S. 563). Um wissenschaftliches Arbeiten hochschuldidaktisch voranzubringen, ist daher eine Auseinandersetzung mit den Bewertungspraktiken der Hochschullehrenden entscheidend.

Bewertungspraktiken durchziehen den gesamten Bildungsbereich auf der Makro-Ebene der Bildungssysteme (z. B. PISA), auf der Meso-Ebene der Hochschulen (z. B. Universitätsrankings), auf der Mikro-Ebene der Lehrenden und Forschenden (z. B. Qualifizierungsvereinbarungen, Beurteilung wissenschaftlicher Arbeiten) und letztlich auch auf der Ebene der Studierenden (z. B. Noten) (Espeland & Sauder, 2007; Heintz, 2010). Diese Bewertungen sind Bestandteile der sozialen und akademischen Ordnung der Hochschulen (Lamont, 2012). Aus soziologischer Perspektive erfolgen Bewertungen und Beurteilungen immer als Folge institutioneller Handlungs- und Kontrolllogiken. Die Etablierung der Valuation Studies (Nicolae, 2018) seit etwa 2010 spiegelt die Vielfalt und Allgegenwärtigkeit von Bewertungspraktiken in der Gesellschaft wider sowie die gesteigerte Aufmerksamkeit gegenüber Quantifizierungen und Soziometrie. Hierzu zählen etwa neue Steuerungslogiken im Hochschulsektor wie die Zahl an Studierenden mit einer Mindeststudienleistung. Auch wissenschaftliche Arbeiten unterliegen quantifizierbaren Parametern, wie etwa einer Mindestseitenanzahl oder einer Mindestanzahl an wissenschaftlichen Quellen. So berichten Lehrende etwa, dass Bachelorarbeiten zeitlich und ressourcentechnisch auf zum Beispiel fünf Wochen Bearbeitungszeit und maximal 40 Seiten beschränkt sein sollen (Bohlinger et al., 2018). Einen erforschten Überblick über die verschiedenen Bewertungsstandards für wissenschaftliche Arbeiten an österreichischen Hochschulen gibt es bis dato nicht. Hinzu kommt, dass „die Bewertungskriterien in Prüfungsordnungen kaum spezifiziert (werden)“ (Zembylas, 2018, S. 180). Die Gründe dafür könnten in der (rechtlichen) Selbstorganisation der verschiedenen Institutionen oder im disziplinären Fachver-

ständnis liegen. Aufgrund verschiedener Prüfungsordnungen entwickeln Institutionen auch eigene Beurteilungsabläufe, in die sowohl interne als auch externe Lehrende des Faches miteinbezogen werden können. Die Prüfungsordnung einer Hochschule kann zum Beispiel einen oder zwei Gutachter:innen für eine Masterarbeit vorsehen (Peters & Dörfler, 2019).

Der erfolgreiche Abschluss des Studiums ist unter anderem von der wissenschaftlichen Abschlussarbeit und den individuellen Faktoren rund um die Betreuung dieser Arbeit abhängig, wie etwa Betreuungsstil, kulturelle Nähe zwischen Studierenden und Lehrenden oder Passung von Thema und Betreuer:in (Hu et al., 2016). Es ergibt sich daher eine gewisse Komplexität und Diversität, wenn man den wissenschaftlichen Habitus von internen und externen Lehrenden, die Unterschiede in den Formvorgaben der Prüfungsordnungen, Betreuungsstilen und Beurteilungskriterien in Betracht zieht. Lehrende wenden in der Bewertung wissenschaftlicher Texte zudem eine Mischung aus inhaltlichen und formalen Kriterien an, um zu einer nachvollziehbaren Bewertung der wissenschaftlichen Arbeiten zu gelangen. Es kommen dabei unterschiedliche Kriterien zur Anwendung, wie etwa Aufbau der Arbeit, Sprache, Stil, Eignung der Forschungsfrage, Methodik oder Eigenständigkeit in der Bearbeitung des Themas. Obwohl einige empirische Berichte von Hochschullehrenden zu deren Bewertungspraktiken vorliegen (etwa Böhlinger et al., 2018), ist das Bewerten von wissenschaftlichen Arbeiten dennoch eine intime und noch wenig diskutierte Praxis in der Hochschuldidaktik. Gute Beispiele hochschuldidaktischer Betreuung oder Schreibberatung könnten durch eine Auflösung der „Black Box“ Bewertung zugänglich gemacht werden, wenn hochschulinterne und hochschulübergreifende Diskussionen darüber angeregt werden.

Der österreichische Hochschulkontext ist nach Hochschultypus in Universitäten, Fachhochschulen, Pädagogische Hochschulen und Kunsthochschulen zu unterteilen. Dabei nehmen die Autorinnen in diesem Beitrag im Speziellen Universitäten und Fachhochschulen auf Basis ihrer persönlichen Lehrerfahrung vergleichend in den Blick. Der vorliegende Beitrag geht der Frage nach, wie gängige Bewertungspraktiken von Hochschullehrenden an österreichischen (Fach)Hochschulen anhand von vier Differenzlinien rekonstruiert werden können: *(1) interne und externe Betreuung der wissenschaftlichen Arbeit, (2) prozessorientierte und ergebnisorientierte Bewertung der wissenschaftlichen Arbeit, (3) offene und geschlossene Gutachten und (4) inhaltliche und formale Bewertungskriterien*. Dabei erfolgt die Erarbeitung der Forschungsfrage mittels Literaturanalyse und mittels Abgleichs dieser mit den eigenen Erfahrungen der Autorinnen als Betreuerinnen wissenschaftlicher Abschlussarbeiten an unterschiedlichen Hochschulen. Die Literaturanalyse verortet sich in der Hochschulforschung (Hamann et al., 2018), der Schreibdidaktik (Kruse & Ruhrmann, 2014) sowie den soziologischen Bewertungsstudien (Nicolae, 2018).

2 Interne und externe Betreuung

Der erfolgreiche Abschluss des Studiums ist unter anderem von der wissenschaftlichen Abschlussarbeit und dementsprechend von der Themen- und Betreuerwahl abhängig (Hu et al., 2016; Schlichte & Sievers, 2015). An Universitäten ist die Betreuung von wissenschaftlichen Arbeiten ausschließlich Personen mit fortgeschrittener wissenschaftlicher Karriere zugeteilt (post docs, Ass.-Professorinnen und -Professoren, Assoz.Professorinnen und -Professoren, Professorinnen und Professoren), die entweder ein Doktorat oder eine Habilitation abgeschlossen haben. Im Betreuungsprozess spiegeln sich die universitären Hierarchien – promovierte oder habilitierte Wissenschaftler:innen gelten als Expertinnen und Experten, Studierende als Laiinnen und Laien in der wissenschaftlichen Expertise, aber auch in den sozialen Rollen und im Status in der wissenschaftlichen Kommunikation (Drognitz et al., 2021). Die Betreuung von wissenschaftlichen Abschlussarbeiten durch interne (oftmals nicht-habilitierte oder promovierte) und externe (nicht-habilitierte oder promovierte) Lehrende ist die gängigste Form der Betreuungspraxis an Fachhochschulen. Die Hochschulkultur an Fachhochschulen orientiert sich eher an der Praxis und arbeitet disziplinenübergreifend. Die Organisationsstruktur weicht von der an Universitäten ab und die Lehre wird in hohem Maße von externen Lehrenden

getragen. Dadurch wird *assessing across the curriculum* die logische Konsequenz, d. h. Lehrende außerhalb des Faches betreuen wissenschaftliche Arbeiten in einer generischen Art und Weise – im Unterschied zu den Universitäten, die in der Regel wissenschaftliche Abschlussarbeiten durch das interne Personal des Mittelbaus oder Professorinnen und Professoren aus dem Fach und weniger durch externe Lehrende außerhalb des Faches betreuen lassen. Universitäten decken damit mit internem Personal die Betreuungstätigkeiten ab, während Fachhochschulen diese outsourcen müssen, da zum Beispiel das interne wissenschaftliche Personal knapp bemessen ist oder externe Betreuende bei der Wahl des Themas erste Wege in ein bestimmtes Berufsfeld ebnen (Schlichte & Sievers, 2015). Wenn nicht-habilitierte Personen laut Prüfungsordnung keine Prüfungen abnehmen, d. h. keine Abschlussarbeiten beurteilen dürfen, weil ihnen die Befugnis dazu (Habilitation) fehlt, dann führt dies in der Regel zu einer Trennung von Beurteilung und Betreuung (Peters & Dörfler, 2019).

Rollenkonflikte scheinen bei der Betreuung von Abschlussarbeiten gut dokumentiert und zudem vorprogrammiert: Das Rollenspektrum reicht hier von Fachexpertentum, Berater:in, Motivator:in, Mentor:in, qualitätssichernde Instanz, Prüfer:in (Gutachter:in) bis hin zur Projektleitung (Drognitz et al., 2021; Hu et al., 2016). Die Ausgestaltung der Rolle als interne Betreuungsperson einer wissenschaftlichen Arbeit (mit Forschungsaufgaben) oder externe Betreuungsperson (aus der Praxis) bringt weitere Varianz in den Prozess. Viele Lehrende beider Hochschultypen äußern Zweifel an ihrer Rolle und ihrem Selbstverständnis als Betreuer:innen von wissenschaftlichen Arbeiten (Macfadyen et al., 2019). Jene, die selbst nicht in der Forschung tätig sind, sind oft unsicher, welche wissenschaftlichen Standards anzuwenden sind, etwa bei der Formulierung einer Forschungsfrage, der Auswahl eines methodischen Designs oder der Zitation von Quellen. Eine weitere Schwierigkeit besteht darin, einen sicheren Rahmen für Studierende zu schaffen, in dem diese ihre wissenschaftlichen Fähigkeiten abseits einer bestimmten Bewertungslogik entwickeln können. Hierzu zählen Empathie, Verständnis und laufende Begleitung der Studierenden (*responsive supervisory expertise*). Ein Erfolg in der Betreuung liegt ebenso in der grundlegenden Kenntnis der Lehrenden über die Lebenssituation der Studierenden (Studienmotivation, Erwerbstätigkeit, Kinderbetreuung, Nähe des Themas zu eigener Berufstätigkeit etc.) sowie in der Adaptivität der Lehrenden, die jeweilige Rolle nach Bedarf wechseln zu können (Buff Keller & Jorissen, 2015; Renske et al., 2015). Externe (nicht-habilitierte oder promovierte) Lehrende an Fachhochschulen haben häufiger einen praxisorientierten Hintergrund und Zugang zum Fach bzw. Feld und verstehen sich weniger als Forschende, die akademische Standards wahren. Interne Lehrende an Universitäten, die selbst wissenschaftlich tätig sind, vertreten in der Regel einen wissenschaftlichen Habitus.

3 Prozessorientierte und ergebnisorientierte Bewertung der Arbeit

Die internen und externen Lehrenden können sich zudem in ihrem Selbstverständnis der Betreuung der wissenschaftlichen Arbeit unterscheiden: entweder als Prozessbegleiter:innen oder reine Gutachter:innen der wissenschaftlichen Arbeiten. Im ersten Fall gibt es mehrere Treffen, Coachinggespräche, Textbesprechungen oder Rückmeldungen zwischen Studierenden und Betreuenden sowie ein prozessorientiertes Verständnis davon, wie eine wissenschaftliche Arbeit entsteht (Fragen und Antworten, Feedback, laufende Betreuung). Im zweiten Fall herrscht ein ergebnisorientiertes Verständnis von wissenschaftlichem Arbeiten vor, in dem Betreuende die wissenschaftliche Arbeit in der Regel nur einmal lesen und dann in Form eines Gutachtens beurteilen. Beide Rollenanforderungen werden an die Hochschullehrenden gestellt: das Begleiten und das Bewerten (Ulmi, Bürki, Verhein-Jarren & Marti, 2017).

Dabei stellt sich generell die Frage, ob Fortschritte der Studierenden überhaupt in den Bewertungsprozess einfließen sollen (*prozessorientiertes versus ergebnisorientiertes Bewerten*). Hinweise darauf kann die jeweilige Prüfungsordnung der Hochschule enthalten. Dies würde voraussetzen, dass die betreuende Person einem prozessorientierten Selbstverständnis folgt und diesen Prozess dokumentiert hat und wiedergeben kann.

Aus der Perspektive der hochschulischen Schreibdidaktik, die sich im letzten Jahrzehnt an österreichischen Hochschulen institutionell verankert hat, ist der Schreibprozess im Rahmen der Betreuung wissenschaftlicher Arbeiten das vorrangige Handlungsfeld von Lehrenden. Diese Verschiebung vom Ergebnis zum Prozess kann der *prozessorientierten Schreibdidaktik* zugeschrieben werden (Kruse & Ruhrmann, 2014). Damit geht ein Wandel des Selbstverständnisses der Lehrenden sowie ein Wandel bei den Erwartungen der Studierenden einher. Wo es früher rein um das Korrigieren von wissenschaftlichen Texten ging, geht es nun verstärkt um die Begleitung des Schreibprozesses, die mündliche und schriftliche Vergabe von fachlichem und schreibdidaktischem Feedback und das In-Beziehung-Treten zwischen Schreibenden, Betreuenden und dem Text (Mitchell, 2010). Die Schreibberatung als didaktische Aufgabe kommt zur Rollenanforderung an Hochschullehrende hinzu. In einer begleitenden und prozessorientierten Rolle unterstützen Lehrende die Studierenden dabei, ihren Text zu erarbeiten, und suchen gemeinsam mit ihnen nach Lösungen, ihre Ziele zu erreichen. In der beurteilenden und ergebnisorientierten Rolle bewerten sie das Produkt nach den Standards des Faches oder ggf. nach Beurteilungskriterien, geben Auflagen zur Überarbeitung oder haben Sanktionsmöglichkeiten, den Text nicht anzunehmen, sollte dieser nicht den Kriterien entsprechen (Ulmi, Bürki, Verhein-Jarren & Marti, 2017). An Fachhochschulen dominiert eine kollektive Logik des zeitgleichen Studienabschlusses eines Jahrgangs, der durch klare Abgabefristen und -regeln geordnet ist und die Betreuungspraxis damit viel stärker reguliert und strukturiert als dies an Universitäten der Fall ist.

Liegt nun eine Abschlussarbeit vor, wechseln Hochschullehrende aus der Rolle der Begleiter:innen in die Rolle der Gutachter:innen. Die Beurteilung des Textes bezieht sich auf wissenschaftliche Normen und Qualitätskriterien des Faches. Fachhochschulen als kollektive Akteurinnen geben ihren Lehrenden häufig Bewertungsschemata und Fristen für die Bewertung und letztendliche Beurteilung vor, während Universitäten dies häufig den individuellen Akteurinnen und Akteuren offenlassen. So nimmt die Hochschule mehr oder weniger Einfluss auf den Bewertungsprozess. Die Beurteilung ist schlussendlich die zusammenfassende Bewertung einer Leistung in Form einer Ziffernote. Je klarer die bewertenden Prinzipien in der Beurteilung fassbar gemacht werden, desto verständlicher die Benotung (Becker-Mrotzek & Böttcher, 2009).

Die beiden hochschuldidaktischen Aufgaben des Begleitens und Beurteilens werden in der Literatur als problembehaftet und konfliktreich dargestellt: „Ein Problem ist die Doppelrolle des Betreuers als Ratgeber und als späterer Gutachter. Letztlich muss auch dem Studierenden klar sein, dass der Hochschullehrer (...) viele Tipps geben kann, aber am Ende zu einer begründeten Beurteilung bzw. Note kommen muss, die die Abschlussarbeit in den Kontext der Qualifikationsarbeiten anderer Absolventen stellt.“ (Bohlinger et al., 2018, S. 156). Die besondere Herausforderung ergibt sich in der Begleitung des Entstehungsprozesses und zudem aus der Anforderung, dennoch das Ergebnis neutral und objektiv zu beurteilen.

4 Offene und geschlossene Gutachten

Die hohe Anzahl an externen und fachlich heterogenen Gutachterinnen und Gutachtern an Fachhochschulen macht es notwendig, stärker standardisierte Bewertungsschemata anzuwenden als an Universitäten, um so den Blick der Externen auf (interne) Bewertungskriterien zu lenken. Dies macht die Beurteilung an sich transparenter und die verschiedenen Kriterien werden als Vorbereitung und Feedback an Studierende kommunizierbar. Gerade wenn mehrere und auch fachlich heterogene Lehrende bei der Betreuung von Arbeiten im Einsatz sind, ist davon auszugehen, dass eine größere Varianz von Bewertungsqualitäten hervortritt. Selbst so neutral erscheinende Kategorien wie die Anzahl der wissenschaftlichen Quellen können zu verschiedenen Positionen der betreuenden Personen führen. Im Gegensatz dazu bleibt eine Professorin an einer Universität innerhalb ihres eigenen Bewertungssystems (Gutachtens) homogener. Bei den oftmals an Fachhochschulen angewendeten Mischsystemen von externen als auch internen Betreuerinnen und Betreuern stellt

sich diese Frage umso deutlicher, inwiefern Institute qualitätssichernd auf einzelne Betreuer:innen einwirken.

Prinzipiell lassen sich zwei Arten der Bewertungspraxis für wissenschaftliche Arbeiten unterscheiden (Tab. 1.).

Tabelle 1: Bewertungspraktiken

Typ	Beschreibung	Hochschultypus
Typ 1: Gutachten, die sich basierend auf Kategorien quantifizierend gestalten	geschlossene Gutachten mit gewichteten Kriterien	häufig an Fachhochschulen vorzufinden
Typ 2: Gutachten, die sich als offener Bewertungsrahmen verstehen	offene Gutachten ohne Kriterien	häufig an Universitäten vorzufinden

Bei kriterienbasierten Vorgangsweisen (Typ 1), die tendenziell an Fachhochschulen gebräuchlich sind, werden wissenschaftliche Arbeiten durch die Vergabe von Punkten oder Prozentsätzen bewertet. Eine Gesamtpunkteanzahl ergibt sich dann etwa aus einer Zusammensetzung aus Forschungsfrage, Methodik und Ergebnissen der Arbeit sowie Formalkriterien. Damit steigt allerdings das Risiko, dass kleine Fehler ins Gewicht fallen, wenn etwa ein Formalfehler oder nicht-gendergerechte Formulierung des wissenschaftlichen Textes den Schreibenden die Bestnote „kostet“. Dies kann durch gewichtete Kriterien abgewendet werden. In den offenen Bewertungsschemata hingegen (Typ 2), die in der Regel an Universitäten angewendet werden, werden Studierende ebenfalls nach Kriterien bewertet, jedoch nach nicht expliziten. Das offene Gutachten liegt in der wissenschaftlichen Verantwortung der betreuenden Person mit Doktorat oder Habilitation und bietet aufgrund dessen Offenheit keinerlei Vergleichsmöglichkeit. Womöglich gehen Lehrende an Universitäten nach bestimmten Kriterien bei der Bewertung der Arbeiten vor, wie logischer Aufbau oder methodische Vorgangsweise, allerdings sind sie nicht „gezwungen“, diese offenzulegen, wie dies an Fachhochschulen vorgegeben und gängige Praxis ist.

Durch die Zugehörigkeit der Studierenden an Fachhochschulen zu Jahrgängen gleicht die finale Punktevergabe durch Gutachter:innen bei einer wissenschaftlichen Abschlussarbeit in letzter Konsequenz einem Ranking von wissenschaftlichen Texten und Personen. Von beispielsweise 0 bis 100 erreichbaren Punkten wissen die Studierenden, wo sie im Vergleich zu anderen ihres Jahrganges stehen. Diese Punkte oder Zahlen verändern, wie Studierende sich selbst und ihre wissenschaftliche Leistung wahrnehmen (Espeland & Sauder, 2007). Die Aufmerksamkeit der Lehrenden sowie der Studierenden an Fachhochschulen richtet sich demnach auf die zu erfüllenden Kriterien wissenschaftlichen Schreibens und diese wiederum verändern die Art und Weise, wie die handelnden Akteurinnen und Akteure Entscheidungen treffen (z. B. Themenwahl, Literaturauswahl, Forschungsfrage). Schließlich kommt es so zu einer Standardisierung und Rationalisierung der Schreibpraxis.

Eine standardisierte Bewertungspraxis hat allerdings auch Vorteile: Betreuende können den Aufwand der Betreuung besser einschätzen und Texte effizienter bewerten, da sie den Formalkriterien folgend relativ standardisierte Arbeiten bekommen, etwa Bachelorarbeiten im Umfang von 40 Seiten. Zudem können Studierende, die über die Bewertungskriterien informiert sind, gezielt darauf achten, so viele Punkte wie möglich zu erlangen, indem sie die expliziten Bewertungskriterien erfüllen (*criteria compliance*) (Mitchell, 2010).

Bewertungsschemata kreieren damit „self-fulfilling prophecies“ (Espeland & Sauder, 2007, S. 15), d. h. es werden wissenschaftliche Abschlussarbeiten abgegeben, die den definierten Kriterien entsprechen, und das Risiko steigt, dass abweichende Arbeiten mit kreativem oder innovativem Zugang so verunmöglicht werden bzw. mit einem hohen Risiko für die Studierenden einhergehen, nicht zu bestehen. Dieser Effekt kann mit der Foucaultschen Perspektive auf Sicherheitsdispositive als akademische, gouvernementale Praxis verstanden werden. Die wissenschaftlichen Abschlussarbei-

ten erfüllen als eine spezifische Form des Sicherheitsdispositivs „die wesentliche Funktion auf eine Realität zu antworten, so daß diese Antwort jene Realität aufhebt, auf die sie antwortet – sie aufhebt oder einschränkt oder bremst oder regelt.“ (Foucault 2019, S. 76). Die Bewertungsschemata bestimmen demnach die Praxis des Schreibens wissenschaftlicher Arbeiten mit, denn es werden infolgedessen Arbeiten produziert, die in den Bewertungsschemata erfassbar sind und so ein strategisches Durchkommen sichern. Dies wiederum verstärkt die Durchdringung des quantifizierbaren Bewertungsschemas: „What is counted usually counts.“ (Miller, 2001, S. 386). Die Beurteilung der Arbeit entscheidet aber möglicherweise über nachfolgende Bildungswege oder Karrierechancen der Studierenden. Konsequenzen haben die Beurteilungen auch dann, wenn sie auf Studierendenseite an Finanzen gekoppelt sind, z. B. zur Ablehnung eines Stipendiums führen. Dann erzeugen Beurteilungen Muster der Exklusion (Simmel, 1971, S. 23 ff.).

5 Formale und inhaltliche Bewertungskriterien

Die Lehrenden wenden in der Bewertung der wissenschaftlichen Texte entweder explizit oder implizit eine Mischung aus *inhaltlichen und formalen Kriterien* an, um zu einer nachvollziehbaren Bewertung der wissenschaftlichen Arbeiten zu gelangen. Inhaltliche Kriterien sind etwa die kritische Auseinandersetzung mit der Literatur oder die Methodik der Arbeit, während die formalen Kriterien sich u. a. auf Formatierung, Rechtschreibung, Mindestseitenanzahl oder die Plagiatsprüfung beziehen können (Bohlinger et al., 2018). Bewertungsschemata, die an Fachhochschulen angewendet werden, zeigen allerdings erhebliche Unterschiede in den Bewertungskriterien selbst sowie in der Gewichtung dieser Kriterien. Die geschlossenen und kriterienbasierten Bewertungsschemata (Typ 1) folgen der Praxiserfahrung der Autorinnen zufolge verschiedenen inhaltlichen und formalen Kriterien. Bei einigen sind Punktevergaben pro Kategorie möglich (z. B. Methodik der wissenschaftlichen Arbeit), manche sind eine Kombination aus Punktevergabe und freier Kommentierung bzw. Begründung. Die untenstehende Tabelle 2 zeigt das Bewertungsschema einer Fachhochschule mit insgesamt sechs Kriterien für wissenschaftliche Arbeiten auf Masterniveau, wovon sich fünf auf inhaltliche und eins auf formale Vorgaben bezieht. Insgesamt können 100 Prozent erreicht werden.

Tabelle 2: Beispiel für ein Bewertungsschema (Typ 1)

Beurteilungskriterien	Gewichtung
1. Aufgabenstellung Klare Formulierung, Abgrenzung von Zielen, wissenschaftlicher Anspruch	10 %
2. Aufbau und Struktur Formale Korrektheit (z. B. durch konsequente Gliederung) inhaltliche Verständlichkeit und Konsistenz	15 %
3. Methodik und Vorgehensweise wissenschaftliche Literaturrecherche Breite und Tiefe der verwendeten Literatur a) Einsatz empirischer Methoden Darlegung des Forschungsdesigns Durchführung und Interpretation der empirischen Untersuchungen b) Literaturarbeit Verwendung einschlägiger Literaturstellen Qualität und Abstützung der Argumentation auf Basis der berücksichtigten Literatur Zusammenhängende Darstellung der wissenschaftlichen Erkenntnisse	25 %
4. Durchdringung des Themas Konzeptionelle Eigenständigkeit bei der Problembearbeitung Niveau und Absicherung eigener Ansätze und Ideen Kritische Auseinandersetzung mit Literaturlücken und -widersprüchen	25 %

(Fortsetzung Tabelle 2)

Beurteilungskriterien	Gewichtung
5. Ergebnisse/praktische Verwertbarkeit Klare Formulierung, eindeutiger Zusammenhang mit der Fragestellung Widerspruchsfreie, konsistente und ausreichende Ergebnisdarstellung	10%
6. Ausdruck und Beachtung formaler Regeln Einhaltung der Form- und Zitiervorschriften Sorgfältige formale Darstellung der Arbeit (Layout, Schriftbild, Formatierung) Stil und Sprache der Arbeit (Verständlichkeit, Satzbau, Rechtschreibung)	15%

Golding u. a. (2014) zeigen auf, dass Gutachter:innen grundsätzlich eine positive Bilanz einer wissenschaftlichen Abschlussarbeit erwarten, d. h. sie haben eine positive Lesehaltung (Ulmi, Bürki, Verhein-Jarren & Marti, 2017). Sie haben sich in der Regel bereits im ersten oder zweiten Kapitel eine Meinung darüber gebildet, ob die Arbeit positiv zu bewerten ist oder nicht. Textkohärenz, die wissenschaftliche Bearbeitung der Literatur und publizierbare Ergebnisse werden in Bewertungen von Lehrenden positiv beurteilt, wohingegen Präsentationsfehler zu einem negativen Eindruck führen (Bohlinger et al., 2018). Nicolae (2018) argumentiert allerdings, wie kleinteilig und situativ – beispielsweise in Notenkonferenzen oder Kommissionen – Wertigkeit hergestellt wird. Die Kleinteiligkeit der Bewertung zeigt sich an den Kategorien, die in standardisierten Bewertungsschemata Anwendung finden:

Die (1) *inhaltlichen Qualitätskriterien* wissenschaftlicher Arbeiten stellen sich wie folgt dar: Die Kategorien „Aufarbeitung der Literatur und Methodik“ und „Aufbau, Struktur und Konzeption der Arbeit“ spielen in den meisten Bewertungsschemata an Fachhochschulen eine Rolle. Dies lässt sich auch aus der Literatur bestätigen: Bourke und Holbrook (2013) untersuchen die Bewertungskriterien für Doktorarbeiten und verdeutlichen, dass die Aufbereitung der Literatur, Konzeption und Methodik sowie die Analyse der Ergebnisse ausschlaggebende inhaltliche Kriterien für Betreuende sind. Die Forschungsfrage kann eine eigene zusätzliche Kategorie darstellen.

- Das Kriterium „Aufarbeitung der Literatur und Methodik“ beinhaltet meistens sowohl die Literaturbearbeitung als auch die in der Arbeit angewandten Methoden der empirischen Sozialforschung. Bei der Literaturbearbeitung wird Wert auf folgende Fragen gelegt: Wie wird wissenschaftliche Literatur in den Text eingearbeitet? Wie erfolgt die Auswahl, Aktualität und Passung der Literatur zur Forschungsfrage? Erwähnt wird außerdem in der Regel das kritische Hinterfragen der Literatur (Quellenkritik) als eigene Subkategorie. Es soll gelingen, den aktuellen Forschungsstand mit der vorhandenen Literatur darzustellen, das Thema des Textes relevant einzugrenzen und eine aktuelle Problemstellung zu schreiben.
- Das Kriterium „Aufbau, Struktur und Konzeption“ fällt bei den Arbeiten in der Regel mit einem erheblichen Anteil an der Gesamtnote ins Gewicht. Gutachter:innen bewerten dabei etwa, ob der Text logisch und schlüssig aufgebaut ist oder die einzelnen Abschnitte des Textes aufeinander Bezug nehmen. Kohärenz in Form eines roten Fadens anhand der Forschungsfrage wird ebenfalls in dieser Kategorie verlangt. In einem Bewertungsschema für Masterarbeiten an einer Fachhochschule wird die Kategorie „Aufbau und Struktur“ beispielsweise mit „formaler Korrektheit, konsequenter Gliederung, inhaltlicher Verständlichkeit und Konsistenz“ beschrieben und mit 15 von 100 Prozent der Note gewichtet. In einem weiteren Bewertungsschema auf Bachelor-niveau einer Fachhochschule (Tab. 3). wird die Kategorie „Aufbau und Struktur der Arbeit“ mit sieben Subfragen operationalisiert. Eine Bewertung kann dann auf einer Skala von 0 bis 2 erfolgen. Anhand der Subfrage 1.4. „Ist das Layout einheitlich und durchgängig, unterstützt es die Lesbarkeit?“ wird deutlich, dass hier ein Formalkriterium inmitten inhaltlicher Kriterien abgefragt wird.

Tabelle 3: Bewertungsschema in der Kategorie „Aufbau und Struktur“

	Aufbau und Struktur der Arbeit	0	1	2
1.1.	Werden die Ausgangssituation sowie die Problemstellung klar dargestellt?			
1.2.	Werden die Fragestellung sowie die Zielsetzung klar dargestellt?			
1.3.	Gibt es eine strukturierte Vorgangsweise zur Beantwortung der Forschungsfrage?			
1.4.	Ist das Layout einheitlich und durchgängig, unterstützt es die Lesbarkeit?			
1.5.	Sind alle Aussagen und Erkenntnisse aus der Arbeit abgeleitet?			
1.6.	Sind alle Teile der Arbeit in den Argumentationsgang logisch und nachvollziehbar eingebunden?			
1.7.	Ist das Verhältnis der einzelnen Abschnitte zueinander sinnvoll?			

- In manchen Fällen kann die Formulierung der Forschungsfrage eine eigene Kategorie in der Bewertung darstellen, die dann deren Relevanz und explizite Beantwortung erfordert: Ein Beispiel verdeutlicht das anhand der Bachelorarbeit (Tab. 4) mit den Subkriterien der relevanten Eingrenzung der Frage, der klaren Formulierung und der theoretischen Diskutierbarkeit der Forschungsfrage. Gewichtet wird die Forschungsfrage mit max. 12 von 100 Punkten.

Tabelle 4: Bewertungsschema in der Kategorie „Forschungsfrage“

Forschungsfrage (FF)		12 P.
1	Relevante Eingrenzung der FF im Vergleich zum Ausgangsthema	
2	Klare Formulierung der FF (und eventuell abgeleiteter Sub-FFs)	
3	Rein theoretische Diskutierbarkeit der FF und eventueller Sub-FFs	

Ein anderes Beispiel für ein Bewertungsschema in der Kategorie der Forschungsfrage verdeutlicht, dass wissenschaftliche Standards im Bewertungsschema vermittelt und expliziert werden können, indem einerseits Fragen zur Kategorie gestellt und andererseits die Idealkriterien in der jeweiligen Kategorie gegenübergestellt werden (Tab. 5).

Tabelle 5: Bewertungsschema in der Kategorie „Forschungsfrage“

Forschungsfrage FF	
Offen formulierte FF, die theoretischen und empirisch-explorativen Erkenntnisanspruch abdeckt	<i>Fragen, die an die Gutachter:innen gerichtet sind</i>
Klar ausgewiesene theoretische Sub- FF(s) mit analytischem Anspruch	
Klar ausgewiesene empirische Sub-FF(s) mit angewandten Methoden beantwortbar	
FF wird logisch aus Problemstellung und Zielsetzung abgeleitet, FF wird nicht geschlossen (ja/nein), kausal eindeutig (wenn/dann, je/desto) oder gerichtet (Ergebnis absehbar) formuliert. FF umfasst Sub-FFs inhaltlich. Theoretische Sub-FFs werden explizit als solche ausgewiesen, ordnen sich der FF inhaltlich unter, schlüsseln die wesentlichen Theoriereferenzen auf und formulieren Beziehungsannahmen. Empirische Sub- FFs werden explizit als solche ausgewiesen, ordnen sich der FF inhaltlich unter, schlüsseln Untersuchungsobjekt und -population auf und lassen sich mit einem qualitativ-explorativen Zugang beantworten (keine Messung, Hypothesenprüfung).	<i>Idealkriterien Beschreibung der Kategorie „Forschungsfrage“</i>

Diese inhaltlichen Qualitätskriterien wissenschaftlicher Arbeiten lassen sich auch in Studien in anderen Ländern wiederfinden: Zheng et al. (2020) geben Originalität bzw. Innovationsgrad (mit 15 % der Note), Aufarbeitung der Literatur (15 %), Relevanz der Arbeit (40 %), praktischen Nutzen (20 %) und Präsentation der Literatur (10 %) als gängige Bewertungskriterien für Masterarbeiten an. Hier zeigt sich im Unterschied zu den gängigen inhaltlichen Bewertungskriterien in Österreich, dass Originalität bzw. Innovationsgrad und der praktische Nutzen von Masterarbeiten in anderen Ländern höher bewertet werden. Die Relevanz der Arbeit wird in der Kategorie der Forschungsfrage mit beantwortet und bewertet. Über das Kriterium der Originalität bzw. Innovationsgrad auf Masterniveau wird in der Literatur allerdings ambivalent diskutiert – dies kann auch erst ab der Dissertation eine Rolle spielen (Zheng et al., 2020).

In ergebnisorientierten Bewertungspraktiken fallen neben inhaltlichen Kriterien die (2) *formalen Kriterien* ins Gewicht. Die Formalkriterien beziehen sich in ergebnisorientierten Bewertungsschemata vorwiegend auf das Einhalten von Zitier- und Formvorschriften (die auch Textgliederung und Anzahl der wissenschaftlichen Quellen beinhalten können), Rechtschreibung und wissenschaftlicher Schreibstil sowie Umfang der Arbeiten (Seitenzahl) (Bohlinger et al., 2018). Die Bachelorarbeiten sowohl an Fachhochschulen und Universitäten umfassen durchschnittlich ca. 30 bis 50 Seiten, je nach Hochschule und Fach, während Masterarbeiten umfangreicher sind und zwischen ca. 80 und 120 Seiten umfassen. Unzutreffende Seitenzahl und ein negativer Plagiatscheck werden sowohl an Fachhochschulen als auch an Universitäten als Knock-out-Kriterien gewertet. Ist eine der beiden Kategorien unzureichend erfüllt, ist die gesamte Arbeit als „nicht genügend“ zu beurteilen. Gender-sensible Schreibweise und die Seitenzahl sind hingegen nicht in allen Bewertungsschemata vorzufinden.

Die Anzahl der wissenschaftlichen Quellen unterliegt ebenso einer formalen Bewertung. Manche Fachhochschulen sehen hier Mindeststandards, z. B. mindestens 20 wissenschaftliche Quellen bei der Bachelorarbeit vor, wovon die Hälfte davon Artikel aus Fachzeitschriften sein müssen. Die verwendeten Publikationen sollen nicht älter als fünf Jahre alt sein. Diese und andere formale Kriterien werden in manchen Gutachten als binäre Kategorien „erfüllt“ bzw. „nicht erfüllt“ verstanden, da die Seitenanzahl oder Anzahl an Quellen quantifiziert werden kann. Hat ein/e Studierende:r nur 18 Quellen verwendet, sind die Formalkriterien nicht erfüllt. Dies birgt allerdings das Risiko, dass eine weitgehend „gute Arbeit“ aufgrund solcher Formalkriterien zurückgewiesen wird oder sinnwidrig überarbeitet werden muss. Inwieweit diese Formalkriterien als hilfreich oder einschränkend erlebt werden, müssten künftige empirische Arbeiten der Hochschuldidaktik erfassen (Bohlinger et al., 2018). Ein beispielhaftes Beurteilungsschema definiert *alle* formalen Kriterien (Zitierweise, Durchnummerierung, Abbildungen, Umfang der Arbeit nach Zeichenanzahl, Plagiatscheck, gendergerechte Sprachwahl) als Muss-Kriterien, ohne deren Einhaltung die Arbeit negativ beurteilt und wieder an die Studierenden zur Überarbeitung zurückgespielt wird. Das Beispiel einer Fachhochschule zeigt die Formalkriterien auf Bachelorniveau auf, die fünf von 21 der gewichteten Kriterien für das Gutachten ausmachen (Tab. 6).

Tabelle 6: Beispiel für Formalkriterien

Formalia
17. Gliederung von Kapiteln und Unterkapiteln
18. Sprachliche Korrektheit (gendersensible Schreibweise, Rechtschreibung, Grammatik, Satzbau)
19. Zitation, Literaturverzeichnis und Formatierung (entsprechend dem Manual)
20. Plagiatscheck
21. Einhaltung der Längen- und Einreichvorgaben sowie formaler Mindeststandards

Die einzelnen formalen Kriterien sind kleinteilig gelistet: So wird unter „sprachlicher Korrektheit“ die Einhaltung von gendersensibler Schreibweise, Rechtschreibung, Grammatik und Satzbau angeführt.

6 Fazit

Der vorliegende Beitrag hat gängige Bewertungspraktiken von Hochschullehrenden an österreichischen Fachhochschulen und Universitäten anhand von vier Differenzlinien in den Blick genommen: (1) *interne und externe Betreuung der Arbeit*, (2) *prozessorientierte und ergebnisorientierte Bewertung der Arbeit*, (3) *offene und geschlossene Gutachten* und (4) *inhaltliche und formale Bewertungskriterien*. Dabei ergab sich die Schwierigkeit, diese Praktiken mit bestehenden Forschungsergebnissen rückzukoppeln, da die Black Box des Bewertens lediglich am Rande wissenschaftlich reflexiv aufgegriffen wird und sich so eine inhärente Lücke in der Wissenschafts- und Hochschulforschung ergibt.

Wie sich gezeigt hat, hängt die letztendliche Beurteilung einer wissenschaftlichen Arbeit nicht nur von den Leistungen der Studierenden, sondern auch von den internen oder externen Lehrenden, die die Arbeiten betreuen, sowie den institutionellen Vorgaben in der Prüfungsordnung, dem Rollenbild, Selbstverständnis und den institutionellen Rahmenbedingungen ab. Studierende wissen oftmals bei der Wahl der Betreuung nicht, ob diese ein prozessorientiertes oder ergebnisorientiertes Verständnis von Betreuung verfolgen und wie viel Betreuungszeit, Rückmeldung und Unterstützung sie erhalten werden. Auch über die Bewertungskriterien und daraus resultierende Gutachten ist den Studierenden meist wenig bekannt. Während Lehrende an Fachhochschulen oftmals geschlossene und kriterienbasierte Bewertungsschemata durchlaufen, die formale und inhaltliche Dimensionen beinhalten, steht Lehrenden an Universitäten oftmals ein leeres Gutachten ohne Kriterien zur Verfügung. Gerade durch den hohen Einsatz von externen Lehrenden an Fachhochschulen ist es zentral, einheitliche und dadurch auch nachvollziehbare Beurteilungsschemata für wissenschaftliche Abschlussarbeiten zu verwenden. Allerdings: Weder eine reine Quantifizierung von Kriterien noch ein loser, offener Beurteilungstext sind schreibdidaktisch wünschenswert und ausreichend aussagekräftig oder vergleichbar. Es braucht beides: Kriterien zur Bewertung und eine textlich argumentierte Begründung. Nur so kann Lernen ermöglicht werden (James, 2014). Während Bewertungsprozesse mit einer Vielzahl an gewichteten Kriterien recht komplex sind, stellt die letztendliche Benotung eine Simplifizierung des vorher stattgefundenen Bewertungsprozesses dar. In der Praxis zeigt sich, dass sich die Qualität einer wissenschaftlichen Arbeit oftmals schwierig in numerischen Werten darstellen lässt, daher ist eine textliche Begründung der Note unbedingt nötig.

Der vorliegende Beitrag bietet einen Einblick in vier Spannungsfelder gängiger Bewertungspraktiken wissenschaftlicher Arbeiten und als Ergänzung einer disziplinen- und hochschulübergreifenden Diskussion über die gängige Praxis des Betreuens und Bewertens. Die systematische Reflexion solcher Bewertungsprozesse im Zuge des hochschulischen Lehrens kann im besten Fall zu einem Organisationsentwicklungsprozess führen (Espeland & Sauder, 2007). Denn gerade mit Blick auf organisationale Veränderungen stellen sich noch viele offene Fragen: Wer wirft einen kritischen Blick auf jene Hochschullehrende, die offene Gutachten schreiben oder Punkte, Prozente und Noten vergeben? Welche Möglichkeiten der Noten-Kalibrierung gibt es bzw. welche Erfahrungen machen Hochschullehrende mit Formen von Zweitbegutachtungen? Wie gehen diese mit kritisch zu bewertenden wissenschaftlichen Texten um? Künftige Forschung sollte sich ebenso auf das wissenschaftliche Selbstverständnis der externen, nicht-promovierten oder habilitierten Lehrenden beziehen: Wie gut eignen sich Praktiker:innen als Betreuer:innen von wissenschaftlichen Arbeiten? Wie wichtig ist die Integration in eine wissenschaftliche community und die Vermittlung eines entsprechenden wissenschaftlichen Habitus für Studierende? Es stellt sich zudem die Frage, ob vorherrschende Titelpkämpfe in der Wissensgesellschaft (Deleuze, 1993) eine prozessorientierte Vorgangsweise bei der Beurteilung wissenschaftlicher Arbeiten nicht ohnehin obsolet machen. Letztendlich hängen viele dieser Fragen davon ab, welche bildungspolitischen Steuerungsmechanismen Hochschulen

vorliegen, um hochschuldidaktische Veränderungen in Betreuungs- und Beurteilungsprozessen initiieren zu können. Eine Schlüsselrolle kommt hierbei jenen Lehrenden zu, die wissenschaftliches Schreiben lehren, denn durch sie kann ein Impuls zur Veränderung der Lehr- und Lernkultur an Hochschulen initiiert werden (Haacke & Frank, 2006).

Für künftige Forschung stellen sich zudem Fragen rund um das Thema der Innovationsverhinderung im Sinne der Entstehung neuen Wissens oder neuer Methodik: Inwiefern blockieren definierte Kriterien und Bewertungsstandards innovative Abschlussarbeiten mit inhaltlichem Wagnis? Inwiefern leisten Zeitdruck, Effizienzregeln und Formalismus auf Lehrenden- und Studierenden-seite ihren Beitrag zur Verhinderung der Generierung neuen Wissens?

Hochschullehrende sollten sich ihrer Doppelrolle zwischen Begleitung der Studierenden im Prozess und der Begutachtung des Textes nach den Kriterien des Faches stets bewusst sein, um in jeder Situation handlungsfähig zu bleiben und die passenden Anregungen zur Weiterentwicklung des Prozesses oder des Textes geben zu können. Auch Nachgespräche mit Studierenden über den Betreuungsprozess als Teil der hochschulischen Beratung könnten als unmittelbare Feedbackmöglichkeit an Lehrende und als Drehscheibe für weitere Entwicklungsschritte genutzt werden. Um die Black Box Bewertung zu öffnen, könnten inneruniversitäre (Teil)Öffentlichkeiten im Sinne von Austauschmöglichkeiten zwischen Lehrenden eines Faches geschaffen werden.

Der Beitrag macht Differenzbeobachtungen in den Bewertungspraktiken von Hochschullehrenden deutlich und regt dazu an, dass Hochschullehrende ihr Selbstbild und ihre Deutungshoheit als Betreuer:innen und Gutachter:innen von wissenschaftlichen Arbeiten reflektieren, um Studierende im wissenschaftlichen Schreiben bestmöglich zu unterstützen und gleichzeitig höchstmögliche Transparenz und Fairness im Beurteilungsprozess zu ermöglichen.

Literatur

- Becker-Mrotzek, M. & Böttchen, I. (2009). *Schreibkompetenz entwickeln und beurteilen*. Berlin: Cornelsen.
- Bohlinger, Chamoni, P., Fröhlich, M., Gerhardt, C., May, J. F., Möller, S., Neuneck, G., Rausch, M., Schneider, P., Stock, S. & Weißhaar, A. (2018). Erfahrungsberichte von Betreuern. S. 137–156. In S. Stock, P. Schneider, E. Peper & E. Molitor (Hrsg.), *Erfolgreich wissenschaftlich arbeiten*. Berlin, Heidelberg: Springer.
- Boltanski, L. & Thévenot, L. (2007). *Über die Rechtfertigung. Eine Soziologie der kritischen Urteilkraft*. Hamburg: Hamburger Edition.
- Bourke, S. & Holbrook, A. P. (2013). Examining PhD and Research Masters Theses. *Assessment and Evaluation in Higher Education*, 38(4), 407–416.
- Buff Keller, E. & Jorissen, S. (2015). *Abschlussarbeiten im Studium anleiten, betreuen und bewerten*. Opladen: Budrich.
- Deleuze, G. (1993). Postskriptum über Kontrollgesellschaften. In G. Deleuze (Hrsg.), *Unterhandlungen 1972–1990*. (S. 254–260). Frankfurt am Main: Suhrkamp.
- Drognitz, M., Mendzheritskaya, J. & Hansen, M. (2021). Beraten und Betreuen in der Hochschullehre. S. 103–116. In R. Kordts-Freudinger, N. Schaper, A. Scholkmann, B. Szczyrba, R. Krempkow, P. Salden, I. Ulrich, I. van den Berk & M. Wiemer (Hrsg.), *Handbuch Hochschuldidaktik*. Bielefeld: wbv Publikation.
- Espeland, W. N. & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113, 1–40.
- Foucault, M. (2019). *Sicherheit, Territorium, Bevölkerung. Geschichte der Gouvernementalität I*. Frankfurt am Main: Suhrkamp.
- Golding, C. et al. (2014). What Examiners Do: What Thesis Students Should Know. *Assessment and Evaluation in Higher Education*, 39(5), 563–576.
- Haacke, S. & Frank A. (2006). Typisch deutsch? Vom Schweigen über das Schreiben. In W. Kissling & G. Perko (Hrsg.), *Wissenschaftliches Schreiben in der Hochschule*. (S. 35–44). Innsbruck: Studienverlag.
- Hamann, J., Kaldewey, D., Bielecki, N., Bloch, R. & Flink, T. (2018). Aktuelle Herausforderungen der Wissenschafts- und Hochschulforschung. *Soziologie*, 47(2), 187–203.

- Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. *Zeitschrift für Soziologie*, 39, 162–181.
- Hu, Y., van der Rijst, R. M., van Veen, K. & Verloop, N. (2016). The Purposes and Processes of Master's Thesis Supervision: a Comparison of Chinese and Dutch Supervisors. *Higher Education Research and Development*, 35(5), 910–924.
- James, D. (2014). Investigating the Curriculum through Assessment Practice in Higher Education. *Higher Education*, 67(2), 155–169.
- Kauffeld, S. & Othmer, J. (2019). *Handbuch Innovative Lehre*. Wiesbaden: Springer.
- Kordts-Freudinger, R., Schaper, N., Scholkmann, A., Szczyrba, B., Krempkow, R., Salden, P., Ulrich, I., Berk van den, I., Wiemer, M. (2021). *Handbuch Hochschuldidaktik*. Bielefeld: wbv Publikation.
- Kropf, J. & Laser, S. (2018). Eine Bewertungssoziologie des Digitalen. In J. Kropf & S. Laser (Hrsg.), *Digitale Bewertungspraktiken* (S. 1–16). Wiesbaden: Springer Fachmedien.
- Kruse, O. & Ruhrmann, G. (2014). Prozessorientierte Schreibdidaktik: Grundlagen, Arbeitsformen, Perspektiven. In S. Dreyfürst & N. Sennewald (Hrsg.), *Schreiben. Grundlagentexte zur Theorie, Didaktik und Beratung* (S. 15–34). Opladen und Toronto: Verlag Barbara Budrich.
- Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology*, 38, 201–221.
- Macfadyen, A., English, C., Kelleher, M., Coates, M., Cameron, C. & Gibson, V. (2019). 'Am I Doing It Right?' Conceptualising the Practice of Supervising Master's Dissertation Students. *Higher Education Research and Development*, 38(5), 985–1000.
- Mennicken, A. & Kornberger, A. (2021). Von Performativität zu Generativität: Bewertung und ihre Folgen im Kontext der Digitalisierung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73(1), 451–478.
- Miller, P. (2001). Governing with Numbers. Why Calculative Practices Matter. *Social Research*, 28, 379–390.
- Mitchell, S. (2010). Now you don't see it; now you do. Writing made visible in the university. *Arts & Humanities in Higher Education*, 9(2), 133–148.
- Nicolae, S. (2018). Soziologie des Wertens und Bewertens. Zur Einführung in die Reihe „Sociology of Valuation and Evaluation“. In S. Nicolae, M. Endreß, O. Berli & D. Bischur (Hrsg.), *(Be)Werten. Beiträge Zur Sozialen Konstruktion Von Wertigkeit* (S. 3–20). Wiesbaden: Springer Fachmedien.
- Peters, J. H. & Dörfler, T. (2019). *Planen, Durchführen und Auswerten von Abschlussarbeiten in der Psychologie und den Sozialwissenschaften*. Hallbergmoos: Pearson.
- Renske A.M, Meijer, P. C., Brekelmans, M. & Pilot, A. (2015). Adaptive Research Supervision: Exploring Expert Thesis Supervisors' Practical Knowledge. *Higher Education Research and Development*, 34(1), 117–130.
- Schlichte, K. & Sievers, J. (2015). Abschlussarbeiten und Prüfungen. In K. Schlichte & J. Sievers (Hrsg.), *Einführung in die Arbeitstechniken der Politikwissenschaft* (S. 133–146). Wiesbaden: Springer Fachmedien.
- Schwartz, L. & Mehnert, G. (2008). *Der amerikanische Traum. Mit Green Card oder Visum in die USA*. Berlin: The American Dream.
- Simmel, G. (1971). *On Individuality and Social Forms*. Edited by D. N. Levine. Chicago: University of Chicago Press.
- Ulmi, M., Bürki, G., Verhein-Jarren, A. & Marti, M. (2017). *Textdiagnose und Schreibberatung: Fach- und Qualifizierungsarbeiten begleiten*. Opladen: Verlag Barbara Budrich.
- Zheng, X. Yanwu, L. & Zhemin, L. (2020). Assessing and Predicting the Quality of Research Master's Theses: an Application of Scientometrics. *Scientometrics*, 124(2), 953–972.
- Zembylas, T. (2018). Zur Kontextualisierung von Bewertungsprozessen. In S. Nicolae, M. Endreß, O. Berli & D. Bischur (Hrsg.), *(Be)Werten. Beiträge Zur Sozialen Konstruktion Von Wertigkeit* (S. 171–194). Wiesbaden: Springer Fachmedien.
- Zumbach, J. & Astleitner, H. (2016). *Effektives Lehren an der Hochschule: ein Handbuch zur Hochschuldidaktik*. Stuttgart: Verlag W. Kohlhammer.

Autorinnen

Dr. Katharina Resch, Universität Wien, Institut für Lehrer*innenbildung, Wien, Österreich; E-Mail: katharina.resch@univie.ac.at

Mag. Anna Schopf, Pädagogische Hochschule Wien, Institut für bildungswissenschaftliche Grundlagen und reflektierte Praxis, Wien, Österreich; E-Mail: anna.schopf@phwien.ac.at



Zitiervorschlag: Resch, K. & Schopf, A. (2022). Black Box Bewertung. Gängige Bewertungspraktiken von Hochschullehrenden bei der Beurteilung wissenschaftlicher Arbeiten. *die hochschullehre*, Jahrgang 8/2022. DOI: 10.3278/HSL2237W. Online unter: wbv.de/die-hochschullehre



die hochschullehre

Interdisziplinäre Zeitschrift für Studium und Lehre

Die Open-Access-Zeitschrift **die hochschullehre** ist ein wissenschaftliches Forum für Lehren und Lernen an Hochschulen.

Zielgruppe sind Forscherinnen und Forscher sowie Praktikerinnen und Praktiker in Hochschuldidaktik, Hochschulentwicklung und in angrenzenden Feldern, wie auch Lehrende, die an Forschung zu ihrer eigenen Lehre interessiert sind.

Themenschwerpunkte

- Lehr- und Lernumwelt für die Lernprozesse Studierender
- Lehren und Lernen
- Studienstrukturen
- Hochschulentwicklung und Hochschuldidaktik
- Verhältnis von Hochschullehre und ihrer gesellschaftlichen Funktion
- Fragen der Hochschule als Institution
- Fachkulturen
- Mediendidaktische Themen

wbv.de/die-hochschullehre



Alle Beiträge von **die hochschullehre** erscheinen im Open Access!