



Auf dem Weg zu einer neuen Prüfungskultur

Innovative Leistungsbewertung an der Privaten Pädagogischen Hochschule der Diözese Linz

Ursula Svoboda^a, Ute Vogl^a

^aPädagogische Hochschule der Diözese Linz,
ute.vogl@ph-linz.at

EINGEREICHT 21 MAR 2021

ÜBERARBEITET 19 MAI 2021

ANGENOMMEN 25 MAI 2021

In diesem Beitrag werden Forschungsberichte zu Innovative Grading dargestellt, diskutiert und mit aktuellen Forschungen von Masterthesis-Studierenden an der PHDL im Sinne eines Literature Reviews ergänzt. Damit werden die Initiativen der Projektgruppe Innovative Grading aufgegriffen, weitergeführt und daraus die Legitimation von Innovative Grading als gleichwertige Beurteilungsform neben konventioneller Ziffernnotenbeurteilung abgeleitet.

SCHLÜSSELWÖRTER: Innovative Grading, Leistungsbeurteilung, Kritik der Ziffernzensur, hochschulisches Qualitätsmanagement

1. Einleitung

Eine erfolgreiche Implementierung innovativer Ansätze der Leistungsfeststellung und Leistungsbeurteilung an Schulen kann nur im Zusammenwirken mit den tertiären Institutionen erfolgen (Schmidinger & Vierlinger, 2012, S.107). Die Private Pädagogische Hochschule der Diözese Linz (PHDL) trägt diesem Appell Rechnung, indem sie 2018 die Projektinitiative *Interessensgemeinschaft Innovative Grading IG²* gründete, die sowohl innovative Formen der Leistungsfeststellung in der Lehrer*innenausbildung praktiziert, beforscht und weiterentwickelt als auch Innovative Grading als zweistufige Beurteilungsform Pass/Fail implementiert. In der Folge werden die Begriffe Innovative Grading und Pass/Fail synonym verwendet.

Innovative Grading als einer Form alternativer Leistungsbeurteilung liegt das zweistufige Pass-/Fail-Prinzip zugrunde. Mit der Beschränkung auf die Formulierungen „Mit Erfolg teilgenommen“ und „Ohne Erfolg teilgenommen“ fällt die Klassifizierung der Qualität von Leistungen in Abstufungen weg, die in der Kritik eines Mangels an Objektivität und wissenschaftlicher Fundierung stehen (Reitinger et al., 2018, S.109). Wird Innovative Grading als Leistungsbeurteilung für eine Lehrveranstaltung an der PHDL gewählt, können Lehrende entweder Anforderungskataloge oder Kompetenzraster als Instrumente der Leistungsfeststel-

lung verwenden. Der Zuwachs an Kompetenzen wäre dadurch schon während des Zeitraums, in dem die Lehrveranstaltungen stattfinden, ersichtlich und würde nicht bloß durch eine Klausur erhoben werden (Reitinger et al., 2018, S. 110) – die Leistungsfeststellung erfolgt demnach lehrveranstaltungsimmanent.

Hervorzuheben ist, dass eine wissenschaftliche Auseinandersetzung mit Formen der Leistungsbeurteilung für Lehrende *und* Studierende an Pädagogischen Hochschulen relevant ist: Einerseits müssen Studierendenleistungen auf Basis lernpsychologischer und gesetzlicher Grundlagen beurteilt werden und andererseits müssen sich Lehramtsstudierende während ihrer Ausbildung nicht nur Kenntnisse über Formen der Leistungsfeststellung und -beurteilung aneignen, sondern basierend auf diesem Wissen Entscheidungen darüber treffen, welche Methoden der Leistungsfeststellung und -beurteilung sie für ihren eigenen Unterricht als geeignet erachten. An der PHDL forschen daher neben Lehrenden ebenso Studierende zum Thema Innovative Grading. Das Pass-/Fail-Format stellt derzeit ein noch junges Forschungsgebiet dar. Die bisher bekannten Forschungsergebnisse versprechen jedoch eine Steigerung des Wohlbefindens (White & Fantone, 2009), weniger Druck- und Stressempfinden (Rohe et al., 2006) sowie weniger Konkurrenzdenken unter den Lernenden (White & Fantone, 2009). Ein partizipativer Ansatz von Leistungsfeststellung und Leistungsbewertung involviert auch die Studierenden in die Entwicklung und Weiterentwicklung von Kompetenzrastern (Kohn, 2001, S. 31). Die Darlegung erworbener Kompetenzen geschieht nicht mehr punktuell durch Prüfungen oder – wie im herkömmlichen Modus – summativ, sondern formativ im Zuge der gesamten Lehrveranstaltung.

2. Innovative Grading - Aktuelle Forschungslage

Evidenzbasierte Lehre ist für das tertiäre Bildungsfeld ein Qualitätsmerkmal. Insofern ist es offensichtlich, dass Studien zum Vergleich unterschiedlicher Formen der Leistungsbeurteilung und deren Wirkungen für eine Pädagogische Hochschule von hoher – auch curricularer – Bedeutung sind. Im Folgenden werden relevante Studien und deren Ergebnisse zum Thema in aller Kürze dargestellt.

2.1. Studie von Valtin & Wagner (2002) zu den Wirkungen von Ziffernbeurteilung und verbaler Beurteilung

Im Jahr 2002 veröffentlichten Valtin & Wagner eine im Rahmen des Projektes „NOVARA (Noten oder Verbalbeurteilung: Akzeptanz, Realisierung, Auswirkungen)“ durchgeführte Studie, in welcher sie die Effekte von Ziffernbeurteilung und verbaler Leistungsbeurteilung auf die Entfaltung der kindlichen Persönlichkeit im Hinblick auf Leistung erforschten. Sie konnten feststellen, dass sich die verbale Beurteilung für Schüler*innen, welche schlechtere Leistungen erbringen und sich

mehr ängstigen, von größerem Vorteil erweist als für jene, die in der Schule besser abschneiden und mehr Selbstsicherheit besitzen. Die minimale Effizienz der verbalen Leistungsbeurteilung beruhe in dieser Angelegenheit auf der geringen Realisierung ihres zugrundeliegenden Konzepts in den Schulen beziehungsweise Klassen. Deshalb könne das Design der Forschung keine gültige Auskunft über die Konsequenzen einer verbalen Beurteilung geben. Ihre Untersuchung ermöglichte dennoch die Erkenntnis, dass Kinder, welche häufig mit negativer Benotung konfrontiert waren, dazu neigten, ihr Augenmerk auf eigene Fehlschläge zu lenken und sich verhältnismäßig mehr vor der Schule beziehungsweise vor Prüfungen zu fürchten. Leistungstärkere Schüler*innen zeigten einen gesteigerten extrinsischen Antrieb. Kinder, deren Leistungsbeurteilung mittels Noten erfolgte, nahmen die Ansprüche der Schule als anstrengender wahr (Valtin & Wagner, 2002).

2.2. Expertise von Brügelmann (2006) zur Ziffernbenotung und alternativen Leistungsbeurteilung

Für diese Expertise wurden verschiedene empirische Studien aus deutschsprachigen Ländern, internationale Forschungen und Reviews über die Verhältnisse in Großbritannien und den Vereinigten Staaten von Amerika zum Thema *Leistungsbeurteilung* einer nochmaligen kritischen Betrachtung unterzogen. Hinsichtlich der Effekte von Benotung und verbaler Beurteilung auf das emotionale Befinden der Schüler*innen zeigte sich, dass Kinder ein höheres Wohlbefinden in der Schule verspürten, wenn keine Ziffernbenotung erfolgte. Die Empfindungen der Kinder hingen jedoch stark davon ab, wie die jeweilige Lehrperson in ihrer Klasse bezüglich der Ziffernbeurteilung verfuhr. Der Grad des Selektionsdrucks stellte hier eine wesentlichere Komponente dar als die Art der Leistungsbeurteilung (Note oder verbale Deskription). Alles in allem hoben sich alle Studien hindurch kaum in puncto Wirkungen ab. Prinzipiell eröffnet zwar eine Variation der Leistungsbeurteilung die Potenzialität günstiger Effekte, in der Schule können sich diese allerdings bloß begrenzt entfalten. Der Grund dafür liegt zum einen darin, dass unterschiedliche Leistungsbeurteilungsformen, wie beispielsweise Schulnoten verbunden mit schriftlicher Ergänzung, miteinander verknüpft werden. Zum anderen erfährt das Konzept der verbalen Leistungsbeurteilung eine unzulängliche oder gar keine Verwirklichung. Außerdem verhindert der Selektionsdruck eine Leistungsbeurteilung, die auf Förderung ausgerichtet ist (Brügelmann, 2006).

2.3. Studie von Rohe, Barrier, Clark, Cook, Vickers & Decker (2006) zu den Vorteilen der Pass-/Fail-Beurteilung hinsichtlich Stress, Stimmung, Gruppenzusammenhalt und Prüfungsangst Medizinstudierender

Beginnend mit der Studierendenkohorte 2006 substituierte die medizinische Fakultät Mayo in Rochester (Minnesota) das fünfstufige Beurteilungssystem im ersten Studienjahr durch ein Pass-/Fail-Beurteilungssystem. Es zeigte sich, dass die Studierenden, deren Leistungsbeurteilung über das Pass-/Fail-System verlief, am Ende ihres zweiten Jahres weniger wahrgenommenen Stress und einen stärkeren Gruppenzusammenhalt als jene Kolleg*innen, welche via fünfstufigem System beurteilt wurden, hatten. Ferner befand sich die Pass-/Fail-Gruppe in besserer Stimmung als die benotete Gruppe, wobei die Differenz aber keine statistische Signifikanz aufwies. In puncto Prüfungsangst und bei den Resultaten der medizinischen Zulassungsprüfung der Vereinigten Staaten Schritt traten keine signifikanten Unterschiede auf (Rohe et al., 2006).

2.4. Studie von Dahlgren, Fejes, Abrandt-Dahlgren & Trowald (2009) zu Beurteilungssystemen, Beurteilungsmerkmalen und den Herangehensweisen Studierender ans Lernen

Durch die Studie konnte festgestellt werden, dass eine klare Beziehung zwischen Beurteilungssystemen, Beurteilungsaufgaben und Lernen existiert. Bei Programmen mit mehrstufigen Beurteilungssystemen konzentrierten sich die Beurteilungsaufgaben mehr auf unwichtige Inhalte des Kurses, Faktenwissen und die Reproduktion von Wissen. Des Weiteren leitete ein mehrstufiges Beurteilungssystem in größerem Ausmaß als ein Pass-/Fail-Beurteilungssystem die Aufmerksamkeit der Studierenden bei der Planung ihres Studiums stärker auf die Beurteilungsaufgabe an sich als auf das Lernen, förderte aber die Kooperation zwischen Studierenden in geringerem Umfang. Studierende eines Studiengangs mit mehrstufigem Beurteilungssystem erlebten die Beurteilung zudem weniger als eine zusätzliche Gelegenheit zum Lernen als Studierende eines Studiengangs mit Pass-/Fail-Beurteilung. Beurteilungen selbst, insbesondere vorhergehende, spielten für Studierende in benoteten Systemen eine größere Rolle als für jene in Pass-/Fail-Systemen (Dahlgren et al., 2009).

2.5. Studie von Bloodgood, Short, Jackson & Martindale (2009) zum Wechsel zur Pass-/Fail-Beurteilung an einer medizinischen Fakultät

Im Jahr 2002 wechselte die University of Virginia School of Medicine für jeden einzelnen Kurs in den ersten zwei Jahren der medizinischen Fakultät von einem fünfstufigen (A, B, C, D, F) zu einem zweistufigen Beurteilungssystem. Im selben Jahr

startete eine Langzeitstudie, die die Wirkung jener Veränderung auf die akademische Leistung der Medizinstudent*innen, deren Anwesenheit, die Bewerbung für einen Praxisplatz, Zufriedenheit und psychisches Wohlbefinden erforschte. Die Untersuchung endete im Jahr 2007 und zeigte, dass durch den Wechsel der Beurteilung von Buchstabenbenotung zu Pass/Fail in den ersten zwei Studienjahren an der medizinischen Fakultät eine deutliche Verbesserung des psychischen Wohlbefindens und der Zufriedenheit der Studierenden eintrat. Eine Leistungsminderung in Kursen, Referendariaten, bei den Testergebnissen der medizinischen Zulassungsprüfung der Vereinigten Staaten, des Erfolges bei der Praktikumsvermittlung oder des Anwesenheitsgrads blieb aus (Bloodgood et al., 2009).

2.6. Studie von McDuff, McDuff, Farace, Kelly, Savoia & Mandel (2014) zum Wechsel zur Pass-/Fail-Beurteilung an einer medizinischen Fakultät

Ursprünglich sah der präklinische Lehrplan der medizinischen Fakultät an der University of California San Diego ein Beurteilungssystem mittels Honors (Auszeichnung)/Pass/Fail vor. Beginnend mit dem akademischen Jahr 2010/11 wechselte die medizinische Fakultät zur Verwendung eines Pass-/Fail-Beurteilungssystems. Nun sollten die Auswirkungen dieser Veränderung auf die akademische Leistung der Medizinstudent*innen erforscht werden. Bei der medizinischen Zulassungsprüfung Schritt 1 kristallisierte sich jedoch keine Verringerung der Leistung heraus. Diese Resultate wurden im breiteren Kontext einer Vielzahl an Faktoren, welche bei der Einschätzung der Vorzüge verschiedener Beurteilungssysteme berücksichtigt werden sollten, diskutiert. Die Autor*innen empfahlen schließlich die Fortsetzung der Pass-/Fail-Beurteilung an der medizinischen Fakultät der University of California San Diego (McDuff et al., 2014).

2.7. Studie von Lüttert et al. (1997) zu Ansichten von Lehrpersonen, Eltern sowie Schüler*innen bezüglich Ziffernnotenbeurteilung und verbaler Beurteilung

Das Projekt „Leistungsbeurteilung und Leistungsrückmeldung an Hamburger Schulen“ (LeiHS- Projekt) von Lüttert et al. (1997) untersuchte die Ansichten von Lehrpersonen, Eltern sowie Schüler*innen bezüglich Ziffernbeurteilung und verbaler Beurteilung und konnte bestätigen, dass sich die Beurteilung mit Noten auf die Anstrengungsbereitschaft auswirken kann. Mehr als die Hälfte der teilnehmenden Schüler*innen stimmten überwiegend oder voll zu, sich bei Noten besonders anzustrengen. Jedoch sprachen sich knapp 60 % gegen die Beurteilung von Noten aus. Daraus lässt sich schließen, dass Noten Kinder sehr wohl zu erhöhter Anstrengungsbereitschaft treiben können, obwohl sie mit diesem Beurteilungsmodus nicht zufrieden sind.

2.8. Studien von Jacobs (2009), Alm & Colnerud (2015) und Guskey (2015) zu Leistungssteigerung durch Notendruck

Jacobs (2009) konnte zeigen, dass Noten eine Erhöhung der Leistung hervorrufen können. Dafür mussten Studierende in einem seiner Seminare etliche Quiz lösen, die zum Teil benotet, zum Teil unbenotet wurden. An einem Seminarplan war genau ersichtlich, an welchen Terminen benotet bzw. nicht benotet wurde. Es konnte gezeigt werden, dass die benoteten Quiz bessere Testleistungen hervorbrachten als die unbenoteten Quiz mit ausführlicher Rückmeldung. Aktuellere Studien (Alm & Colnerud, 2015; Guskey, 2015) zeigten, dass sich die Leistungsbeurteilung durch Noten auf leistungsschwache Lernende besonders negativ auswirkt. Müller (2013, S. 57) erläutert dazu, dass die Motivation von Schüler*innen, ihre Ausdauer und die Aufgabenauswahl direkt mit den Erfolgserwartungen gekoppelt sind. Können jedoch keine Erfolgserlebnisse aufgrund schlechter Noten erlebt werden, wird auch die Anstrengungsbereitschaft sowie die Motivation sinken.

2.9. Studie von Möller et al. 2014 zu Kompetenzrasterzeugnissen

Im Rahmen des Hamburger Schulversuchs *alles>>können*, dessen Ziel es war, systematisch kompetenzorientierten Unterricht zur individuellen Förderung aller Schüler*innen zu entwickeln, wurden Kompetenzrasterzeugnisse zur Leistungsbeurteilung und -rückmeldung erprobt. Im Laufe des fünfjährigen Schulversuchs nahm die Verwendung von Kompetenzrasterzeugnissen deutlich zu. Eine Befragung der teilnehmenden Lehrer*innen ergab eine deutliche Präferenz von Kompetenzrasterzeugnissen gegenüber Notenzeugnissen. Lehrer*innen begründeten ihre Präferenz mit dem in ihren Augen hohen Potenzial zur Förderung der zukünftigen Leistungsentwicklung der Schüler*innen. (Möller et al., zitiert nach Beutel, S.-I & Pant, H. A. 2020, S. 75).

2.10. Studie von Reitinger et al. 2020 zu Pass-/Fail-Beurteilung, Ziffernzensur und Lernerleben in der Lehramtsbildung

Eine erste Evaluierung zu Pass/Fail wurde bereits 2020 an der PHDL durchgeführt. Vordergründig zeigt sich (ähnlich wie in manchen englischsprachigen Studien) kein wesentlicher Unterschied zwischen den Beurteilungsformen Ziffernnoten versus Pass/Fail. Bei genauerer Betrachtung muss kritisch angemerkt werden, dass in der untersuchten Lehrveranstaltung „Anwendung wissenschaftlicher Forschungsmethoden“ kognitives Wissen und forschungspraktische Kompetenzen aufgebaut und bewertet werden. Es liegt die Vermutung nahe, dass Innovative Grading eher jenen Lehrveranstaltungen entspricht, in denen personale, soziale und künstlerische Kompetenzen aufgebaut und erweitert werden sollen. Auf die-

sen Aspekt wird hier nicht weiter eingegangen. Angesichts der komplexen Inhalte einer Lehrer*innenausbildung und deren Beurteilung wäre genau diese Themenstellung von großer Bedeutung für weitere Forschungen.

3. Aktuelle Forschungsarbeiten zu IG an der PHDL – Zusammenfassung und Ergebnisse

Im Rahmen einer von Svoboda und Vogl gegründeten Masterthesis-Arbeitsgruppe beschäftigten sich vier Studierende mit unterschiedlichen Aspekten zum Thema. Die dazu von den Studierenden gestellten Forschungsfragen lauten a) Welchen Einfluss hat Innovative Grading auf die Entscheidung künftiger Lehrer*innen für oder gegen alternative Leistungsbeurteilung? (Katharina Baumgartner), b) Erhöht der Einsatz von Kompetenzrastern die Objektivität der Leistungsbeurteilung? (Johanna Strasser), c) Welchen Einfluss haben unterschiedliche Beurteilungsformen auf die Anstrengungsbereitschaft von Studierenden der PPHDL? (Maria Plöchl) und d) Wie wirken sich alternative Bewertungsmethoden auf das Wohlbefinden von Studierenden an der PHDL aus? (Anja Büringer; Anm.: Die Arbeit ist zum Zeitpunkt der Veröffentlichung dieses Artikels noch nicht fertig gestellt). Ähnlich wie in einem Literature Review werden die Arbeiten und deren Ergebnisse dargestellt. Masterthesen unterliegen per se diversen Limitationen, sind aber in der Zusammenschau eine wichtige Ergänzung der vorgestellten Studien und ein Abbild der vielfältigen bildungswissenschaftlichen Diskussionen am Haus.

3.1. Der Einfluss unterschiedlicher Beurteilungsformen auf die Anstrengungsbereitschaft (Plöchl, 2021)

In einer Online-Gruppendiskussion mit Studierenden wurden der schulische Werdegang und wichtige Ereignisse der eigenen Bildungsbiografie kommuniziert. Außerdem wurde reflektiert, wie sich die Einstellungen zur Notengebung über die Jahre verändert, welche Auswirkungen diese auf das persönliche Selbstbild und den jeweiligen Lernprozess haben und welche Haltung sich zu den Beurteilungsmodi im Studium gebildet haben. Dafür wurde die Gruppendiskussion als Datenhebungsmethode gewählt, die von Lamnek (2005, S. 408) als spezielle Form des Gruppeninterviews bestimmt wird. Im Anschluss wurde das Datenmaterial durch die inhaltlich strukturierte, qualitative Inhaltsanalyse nach Mayring (2015) ausgewertet.

Die sechs Studierenden des berufsbegleitenden Studiums Elementarpädagogik teilten in der Untersuchung eigene Erfahrungen mit Leistungsbeurteilung aus der Volksschule, Mittelschule/ Gymnasium und der berufsbildenden „Bildungsanstalt für Elementarpädagogik“ (BAfEP). Ebenso wurden Meinungen und Ansichten bezüglich der Beurteilungsmodi an der Privaten Pädagogischen Hochschule ausge-

tauscht, woraus sich ein Dreischritt formulieren lässt. Denn aus den biografischen Erfahrungen und den aktuellen Erkenntnissen aus dem Studium ergeben sich allgemein gültige Feststellungen für die Pädagogik.

Die zentralen Resultate der vorliegenden Studie zeigen, dass unterschiedliche Formen der Leistungsbeurteilung tatsächlich unterschiedliche Auswirkungen auf die Anstrengungsbereitschaft von Lernenden haben können, diese aber ebenso von zahlreichen anderen Faktoren beeinflusst wird. Diese sind die bereits genannte Bildungsbiografie, der sozioökonomische Hintergrund sowie der Einfluss von Familie und Freunden. Aus all diesen Aspekten ergibt sich ein komplexes Zusammenspiel mit erheblichem Einfluss auf Lernverhalten, Leistungsmotivation und die damit verbundene Anstrengungsbereitschaft.

In Tabelle 1 wurden die wichtigsten Erkenntnisse zur Beantwortung der Forschungsfrage nach Kategorien aufgelistet, welche durch die Gruppendiskussion entwickelt wurden.

TAB. 1. *Zentrale Erkenntnisse zum Einfluss von Anstrengungsbereitschaft in Bezug auf unterschiedliche Beurteilungsformen*

Kategorie	Ergebnisse
Ziffernbeurteilung	<ul style="list-style-type: none"> • kann sich lernhemmend/lernförderlich auswirken • kann Druck, Stress und Angst auslösen • kann Selbstwert positiv/negativ beeinflussen
Innovative Grading	<ul style="list-style-type: none"> • Anstrengungslevel kann variiert werden • ermöglicht unterschiedliche Zielvorstellungen • erhöht das Wohlbefinden • löst weniger Stress aus • größere Leistungsspanne?

Anm.: Darstellung Maria Plöchl

Es konnte festgestellt werden, dass die teilnehmenden Kindergarten- und Krabbelstufenpädagoginnen die Beurteilungsform selbst nicht als einflussreichsten Faktor für die Anstrengungsbereitschaft ansehen, diese aber sehr wohl Einfluss nehmen kann. So wurde von einer Untersuchungsbeteiligten geäußert, dass die Ziffernbeurteilung mit Anstrengung assoziiert wird, während bei der alternativen Beurteilungsform mittels Pass-/Fail-Format ein entspannteres Lernverhalten erlebt wird. Weiters sind mehrere Interviewteilnehmerinnen der Ansicht, dass bei Innovative Grading das Anstrengungslevel und die persönlichen Zielvorstellungen von den Lernenden selbst beeinflusst werden können. So können sich die einzelnen Studierenden, je nach Interesse und verfügbarer Zeit, in unterschiedlichem Ausmaß engagieren und einbringen. Dieser Aspekt warf die Frage auf, ob die Leistungsschere bei einem Pass-/Fail-Format demnach auseinander gehen würde. Die Teilnehmerinnen diskutierten im gemeinsamen Gespräch über psychische Auswirkungen der unterschiedlichen Beurteilungsmodi. Diesbezüglich wurde festgestellt, dass Noten bei einigen Probandinnen in der Vergangenheit Druck, Stress und Angst auslösten, während durch das Pass-/Fail-Format an der PHDL eine entspanntere Arbeitshaltung generiert werden

konnte. Aus der vorliegenden Untersuchung geht außerdem hervor, dass vorhandenes Interesse in Bezug auf die Anstrengungsbereitschaft eine wichtige Dimension darstellt. Laut Brandstätter et al. (2018, S. 120–121) ist Interesse ein entscheidender Motivationsfaktor, durch den ausdauerndes und engagiertes Lernen möglich gemacht wird. Wird ehrliches Interesse im Lernprozess geweckt, handelt man als Lernende*r selbstbestimmt und ist zudem gewillt, erhöhte Anstrengungsbereitschaft aufzubringen. Zudem werden mehr positive Erfahrungen im Lernprozess erlebt und generelle Freude an der Tätigkeit selbst aufgebaut. Rheinberg (1996, S. 14) ergänzt dazu: Wenn die Lernenden durch äußere Reize motiviert werden, werden im Lernprozess weniger nachhaltige Lern- und Lösungsstrategien angewendet. Sind Lernende aber am Lernstoff interessiert und zeigen sich intrinsisch motiviert, kommt tiefgreifenderes und differenzierteres Lernverhalten zum Tragen.

Es konnte gezeigt werden, dass Noten oder Zeugnisse extrinsisch motivieren können, diese Motivation jedoch nur bis zur Erreichung des jeweiligen Ziels anhält. Intrinsische Motivationsfaktoren wie Interesse, Sinnempfinden oder der Wunsch nach Kompetenz- und Wissenserweiterung sind für die Leistungsmotivation ausschlaggebend und stellen für die Anstrengungsbereitschaft eine wichtigere Dimension als die Art der Leistungsbeurteilung dar. Durch reduzierten Stress und erhöhtes Wohlbefinden kann interpretiert werden, dass Beurteilungsprozesse mittels Innovative Grading einen geringeren extrinsischen Motivationsfaktor darstellen und man als Lernende eher intrinsisch motiviert handelt.

Nachdem mit den bereits beschriebenen Forschungsergebnissen die Forschungsfrage beantwortet werden konnte, wird im Folgenden der Fokus auf andere Einflussfaktoren gelegt, die für die Bereitschaft zur Anstrengung entscheidend sein können. Neben den bereits behandelten intrinsischen Motivationsfaktoren, die wohl den größten Einfluss für die Anstrengungsbereitschaft darstellen, wurde eine weitere wichtige Komponente von den Teilnehmerinnen der Gruppendiskussion mehrfach betont: die Lehrperson. So teilten die Pädagoginnen die Ansicht, dass eine vertrauensvolle Beziehung zur Lehrperson sowohl in der Volksschule als auch im Studium die Basis eines erfolgreichen Lernprozesses bildet. Dies wird auch von zahlreichen Studien belegt. Die 2009 vom neuseeländischen Bildungsforscher John Hattie veröffentlichte Hattie-Studie trägt den Titel „Visible Learning“. In der Metastudie wird die Frage thematisiert, welche Aspekte für den schulischen Lernerfolg relevant sind. Lotz und Lipowsky (2015, S. 105–106) halten ein zentrales Ergebnis dieser Meta-Metaanalyse fest: „Teachers make the difference“. „Visible Learning“ unterstreicht laut Lotz und Lipowsky (2015, S. 127–128) nicht nur die Einfühlungsvermögen und Beziehungsfähigkeit der Lehrperson, sondern auch deren Einstellungen und Wertvorstellungen sowie das Interesse, das den Lernenden entgegengebracht wird. Die Studie arbeitet heraus, dass für 30 % der Leistungsunterschiede die Lehrperson verantwortlich ist. Dies stellt einen beeindruckenden Einflussfaktor dar, dessen sich Lehrpersonen bewusst sein sollten.

Neben der Beziehungsgestaltung wurden lernförderliche Charakteristika von Lehrkräften thematisiert. Dabei wurden Begeisterungs- und Motivationsfähigkeit sowie empfundene Sympathie genannt, welche die Vertrauensentwicklung und den daraus resultierenden Motivationsaufbau fördern (Schweer, 2014, S. 253). Wird von Lehrenden Begeisterung für einen Themeninhalt wahrgenommen, wirkt dies für die Diskussionsteilnehmerinnen sogar ansteckend. Dies lässt sich neurobiologisch mit Spiegel-Nervenzellen beschreiben: In unserem Gehirn werden Emotionen oder Gefühle, die sich in unserem Gegenüber abspielen, durch Spiegelneuronen gespiegelt. Diese Spiegel-Nervenzellen arbeiten unbewusst in uns und lassen uns fühlen, was unser Gegenüber fühlt. Einfach gesagt: Spiegelneuronen geben uns nicht nur die Information über die Gefühle anderer Menschen, man wird mit diesen Gefühlen sogar angesteckt (Bauer, 2017, S. 4–10).

Die letzte der von den Probandinnen genannten Einflusskomponenten für anstrengungsbereites Verhalten in Schule oder Studium ist eine angenehme soziale Lernatmosphäre. Dabei wird eine kleinere Gruppengröße bevorzugt, da diese einen familiäreren Rahmen für die Interviewbeteiligten ermöglicht und bessere Austauschmöglichkeiten bietet. Hier muss jedoch angemerkt werden, dass das Empfinden von Wohlgefühlen in der Schule natürlich nicht nur von den sozialen Beziehungen abhängt. Leistungszufriedenheit, Selbstwirksamkeit, das Klima in der Schulklasse oder die schulischen Rahmenbedingungen können als weitere Aspekte genannt werden (vgl. Hascher, 2004).

Zusammengefasst: Durch die vorliegende Studie wird erstmals gezeigt, dass die Art der Beurteilung Einfluss auf die Anstrengungsbereitschaft haben kann, jedoch intrinsische Motivationsfaktoren wie Interesse zum Gegenstand, das Empfinden von Sinnhaftigkeit im Lernprozess oder Wissensdrang einen noch größeren Einflussfaktor darstellen.

3.2. Objektivität durch Kompetenzraster (Strasser, 2020)

Noten werden vielerorts als subjektiv und demnach als nicht gerechtfertigt empfunden, obwohl die Rechtslage der LBVO (2020, § 11 Abs. 2) festlegt, dass die Lehrperson größtmögliche Objektivität im Prozess der Leistungsbeurteilung anzustreben hat. Diese Diskrepanz wurde zum Anstoß dieser Masterthesis herangezogen, in der der Frage nachgegangen wird, ob Kompetenzraster die Objektivität der Beurteilung erhöhen. Dazu sei zuerst der Kompetenzbegriff nach Weinert (2001) zitiert. Kompetenzen sind „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll zu nutzen“ (Weinert, 2001).

Grundsätzlich ist das Kompetenzraster eine Möglichkeit zur formativen Leistungsfeststellung, das Auskunft über das Kompetenzprofil der Lernenden gibt. Dieses Profil hält fest, was jemand in einem bestimmten Themenbereich kann, wie gut das Wissen verankert ist (Neuweg, 2019, S. 58) und welche Ziele ein Lernender bzw. eine Lernende noch erreichen kann. Ziener und Kessler (2012, S. 210) sprechen von „Kompetenzentwicklungsplan“, da das Raster während des gesamten Bildungswegs einer Person geführt werden kann und somit einen Plan zur Entwicklung von Kompetenzen darstellt. Ein Kompetenzraster „muss Lernprozesse sequenzierbar und graduierbar machen“ bzw. so aufgeschlüsselt sein, dass die etappenweise Entwicklung einer Kompetenz verständlich aufgelistet ist. (Ziener, 2016, S. 75). Neuweg (2019, S. 66–68) beschreibt acht essentielle Funktionen des Kompetenzrasters:

1. **LEHRZIELDEFINITION.** Die fachlich basierten Unterrichtsziele werden bei der Erstellung eines Kompetenzrasters hinterfragt. Es wird formuliert, welche Fähigkeiten erworben werden sollen, um Anforderungen im jeweiligen Kompetenzbereich zu erfüllen. Dabei ist eine intensive Auseinandersetzung mit Mindestanforderungen nötig.
2. **LEHRZIELKOMMUNIKATION.** Die festgelegten Ziele werden den Lernenden am Anfang eines Lernprozesses mittels Kompetenzraster dargelegt. Durch diese Transparenz wissen die Lernenden, was auf sie zukommt und können sich darauf einstellen.
3. **UNTERRICHTSPLANUNG.** Ein Kompetenzraster lenkt den Fokus auf die zu erlernenden Kompetenzen und die Unterrichtsplanung wird den Zielen angepasst. Diese Outputorientierung führt zum nachhaltigen Aufbau von Können.
4. **PRÜFUNGSGESTALTUNG.** Die Anwendung von Kompetenzrastern bringt eine vielfältige(re) Prüfungskultur hervor und lässt die vorherrschende starke Gewichtung von beispielsweise Schularbeiten verblassen. Die Entscheidung der Lehrpersonen, welche Form der Leistungsfeststellung passend ist, fällt damit bewusster aus. Damit wirken Kompetenzraster an der Entwicklung valider Leistungsfeststellung mit.
5. **AUFMERKSAMKEITSTEUERUNG.** Durch ein Kompetenzraster können Lehrpersonen, Erziehungsberechtigte sowie Lernende ihren Fokus auf den Erwerb von Kompetenzen legen anstatt bloß auf Noten, Punkte oder Schul- und Hausübungen.
6. **FORMATIVE LEISTUNGSFESTSTELLUNG UND FEEDBACK.** Mithilfe des Kompetenzrasters wird die formative Leistungsfeststellung detaillierter, da ein aussagekräftiges Stärken-Schwächen-Profil anstatt Noten, Punkten oder Fehlerzahlen entsteht. Ferner bietet das Kompetenzprofil eine differenzierte Feedback-Grundlage.
7. **SELBSTBEURTEILUNG.** Lernende können durch das Kompetenzraster in den Prozess der Leistungsfeststellung und – beurteilung integriert werden. Sie erhalten

die Möglichkeit, an der Formulierung von Kriterien zur Leistungsfeststellung mitzuarbeiten, ihre Leistungen selbst einzuschätzen und zu beurteilen.

8. **TRANSPARENTE NOTENGEbung.** Mit dem Kompetenzraster kann analysiert werden, welche Fähigkeiten in welcher Ausprägung sich in Schüler*innenleistungen zeigen. So entsteht am Ende eines Beurteilungszeitraums ein individuelles Kompetenzprofil, das eine sachlich-orientierte (kriteriale) Bezugsnorm ermöglicht.

Es lässt sich festhalten, dass es nicht „das“ Kompetenzraster gibt, sondern eine Vielzahl an Möglichkeiten zur Gestaltung desselben. Diese Vielfalt ist auf die internen Bedingungen, an welchen sich die Generierung des Kompetenzrasters richtet, zurückzuführen. Dazu gehören beispielsweise die Individualität von Lehrpersonen oder auch Schulkonzepte (Kayser, 2008, S. 14). Neuweg (2019, S. 177) weist bei der Entwicklung eines Kompetenzrasters darauf hin, dass Pädagogen*innen stets „das realistisch Erreichbare“ aus dem Lehrplan herausarbeiten und in das Kompetenzraster einarbeiten sollten. Ferner ist es essentiell, die Inhalte fachspezifisch anzupassen, da in jedem Fach unterschiedliche Kompetenzen im Vordergrund stehen.

Um die Qualität dieses Instrumentes zu gewährleisten, können folgende Gütekriterien berücksichtigt werden (Stiggins, 2008, S. 173, zit. n. Keller, 2011, S. 147–148):

Gute Raster sollten:

- einen Expertenkonsens abbilden, was das Kernkonzept zu einer guten Leistung in einem bestimmten Lerngebiet ausmacht;
- auf weiterführende oder übergeordnete Bildungsziele bezogen sein, z. B. Lehrpläne, Bildungsstandards usw.
- genügend Kriterien in unterschiedlichen Kategorien enthalten, die genau beschreiben, was für eine „sehr gute“, „gute“, „genügende“, bzw. ungenügende“ Leistung zu erbringen ist;
- in einer klaren und prägnanten Sprache verfasst sein, um Fehlinterpretationen und Missverständnisse vorzubeugen;
- positive Formulierungen enthalten und beschreiben, was eine Person in einem bestimmten Gebiet kann;
- den Lernenden bereits während der Lernphase, auf jeden Fall vor Bekanntgabe der Beurteilung, zur Verfügung stehen.

In der vorgenommenen Forschungsarbeit wurde ein Leitfadenterview mit fünf Lehrpersonen der PHDL durchgeführt, die bereits mit Hilfe von Kompetenzrastern im Pass-/Fail-Modus beurteilen. Das wesentliche Ziel dieser Forschung war, herauszufinden, ob die Arbeit mit Kompetenzrastern eine erhöhte Objektivität im Prozess der Leistungsfeststellung garantiert. Als Ergebnis der Befragung fanden sich fünf Faktoren, die zu einer objektiven Leistungsfeststellung mit Kompetenzrastern führen:

1. **VEREINHEITLICHE LERNZIELE.** Die Befragten gaben an, dass aufgrund von festgelegten Kriterien gleiche Bedingungen für alle Studierenden herrschen und somit ein objektiver Rahmen gegeben ist, in dem eine Beurteilung stattfinden kann.
2. **FUNDIERTE UND MESSBARE INHALTE.** Das Ziel dabei ist, Lernangebote entsprechend den zu erreichenden Kompetenzen zu erstellen, sodass die Konzentration auf jenen Inhalten liegt, die zu einem Zuwachs von Wissen bzw. von Kompetenzen führen. Dadurch werden überflüssige Inhalte wie beispielsweise reine Beschäftigungsarbeiten vermieden und es kann eine Lernkultur, in der Qualität vor Quantität steht, entwickelt werden.
3. **EINE PRÄZISE FORMULIERUNG DER KOMPETENZEN.** Die Suche nach präzisen Begrifflichkeiten stellt sich als herausfordernd dar, weil oftmals komplexe Kompetenzen und Inhalte auf einfache und prägnante Sprache „übersetzt“ werden müssen. Hierbei geben die Interviewten an, dass die Zusammenstellung und Benennung von Kompetenzen erfahrenen Lehrpersonen leichter fällt als Junglehrer*innen.
4. **INTERSUBJEKTIVITÄT.** Die Befragung der Lehrenden machte deutlich, dass der intersubjektive Austausch über Leistungsanforderungen nicht nur die Zusammenarbeit im Team des Fachbereichs erhöht, sondern auch die Objektivität der Beurteilung steigert. Ein weiterer Aspekt der Intersubjektivität wird durch Selbstbeurteilung der Studierenden erreicht. Dazu wird das Kompetenzraster zur Selbsteinschätzung genutzt, über die es im Nachhinein einen intersubjektiven Austausch gibt.
5. **EINE REFLEXIVE HALTUNG** im Hinblick auf das eigene, subjektive Handeln. Die Untersuchung zeigte bei allen Teilnehmenden eine reflexive Wahrnehmung eigener Ansichten bei der Arbeit mit Kompetenzrastern, was eine Voraussetzung darstellt, um objektiv arbeiten zu können.

3.3. Motive für die Präferenz eines Formats der Leistungsbeurteilung (Baumgartner, 2021)

Für die Klärung der Forschungsfrage „Aus welchen Gründen präferieren angehende Lehrpersonen Ziffernbenotung oder alternative Leistungsbeurteilung beziehungsweise eine Mischform daraus für ihre künftige Praxis?“ wurde mit leitfadengestützten Interviews, welche sich an der Theorie des überlegten Handelns der Sozialpsychologen Ajzen und Fishbein orientierten, gearbeitet. Es wurde untersucht, inwieweit bereits in der Vergangenheit liegende Ereignisse (Hintergrundfaktoren) Einfluss auf die eigene Biographie respektive den persönlichen Lebenslauf nehmen, und es wurden individuelle Sichtweisen (Einstellungen, Verhaltensintention, subjektive Norm), welche zum Zeitpunkt der Studie vorherrschten, ermittelt. An der Befragung nahmen drei weibliche und drei männliche Studierende des Mas-

terstudiums Primarstufe der PHDL teil. Die Ergebnisse zeigten, dass a) keine*r der interviewten Student*innen die Anwendung der Ziffernbenotung in seiner*ihrer zukünftigen Lehrer*innenlaufbahn anstrebte, b) zwei Studentinnen und zwei Studenten die alternative Leistungsbeurteilung bevorzugten und c) eine Studentin und ein Student sich für eine Mischform von Ziffernbenotung und alternativer Leistungsbeurteilung aussprachen.

Die vier Studierenden („b“), deren Wahl auf die alternative Leistungsbeurteilung fiel, äußerten hauptsächlich unangenehme Erfahrungen mit Ziffernbenotung in der eigenen Schulzeit und daher auch damit verbundene negative Gefühle. Im Hinblick auf Innovative Grading hingegen zeigte sich eine positive Tendenz, was die emotionale Lage anbelangte. Die persönlichen Einstellungen zu Innovative Grading und alternativer Leistungsbeurteilung waren durchwegs positiv. Der Ziffernbenotung standen diese Interviewpartner*innen allesamt negativ gegenüber. Unter ihren Mitmenschen vermuteten die befragten Personen sowohl Unterstützer*innen als auch Gegner*innen ihrer Entscheidung. Deren Ansichten spielten für die beiden Studentinnen und ihre Abgabe einer Empfehlung sehr wohl eine Rolle. Im Gegensatz dazu würden die zwei Studenten die fremden Anschauungen – ihrer Einschätzung nach – nicht tangieren. Die Interviewten sahen viele Vorzüge im Einsatz der alternativen Leistungsbeurteilung. Sie trüge mehr Vorteile in sich, da sie eine deskriptive, differenziertere Darlegung der Leistung ermögliche, womit eine höhere Aussagekraft der Dokumentation einhergehe. Der Fokus liege hierbei auf den vorhandenen Kompetenzen, Potenzialen und Schwächen der Schüler*innen sowie geeigneten Fördermaßnahmen, deren Detektion auf akkurater Beobachtung fuße. Darüber hinaus führe alternative Leistungsbeurteilung zu einer Minimierung von Druck und einer positiven Prägung der Haltung der Kinder gegenüber der Institution Schule.

Jene Studierenden („c“), welche für die Mischform von Ziffernbenotung und alternativer Leistungsbeurteilung plädierten, schilderten, dass sie mit Innovative Grading positive Empfindungen assoziierten. Ebenso vertraten sie eine positive Meinung zu Innovative Grading und gleichfalls alternativer Leistungsbeurteilung. Ihre Position zur Ziffernbenotung stuften sie einesteils neutral, andernteils interessanterweise negativ ein. Bei der imaginierten Verwirklichung ihrer vorläufigen Empfehlung erwarteten sie sich von ihren Mitmenschen in erster Linie Unterstützung. Welche Überzeugungen die Außenstehenden in dieser Angelegenheit hegten, wäre für die weibliche Befragte und ihren gefassten Entschluss von Relevanz, während dies, laut Erzählung des männlichen Interviewpartners, keinen Effekt auf dessen weitere Vorgehensweise hätte. In der inhaltlichen Argumentation der beiden spiegelte sich im Besonderen die Auffassung wider, nicht nur den Bedürfnissen der Schüler*innen, sondern auch jenen der Erziehungsberechtigten und sekundären Bildungseinrichtungen gerecht werden zu müssen. Grundsätzlich mangle es der alleinigen Ziffernbenotung an genügend Aussagekraft. Dieser Umstand

könne jedoch durch die Hinzuziehung der alternativen Leistungsbeurteilung ausgeglichen und sohin adäquate Differenzierung erreicht werden. Überdies erhielten die Kinder dadurch transparente, verständlichere Einsicht in ihre Fertigkeiten und geringeren Begabungen, in das Zustandekommen ihrer Ziffernnoten sowie Feedback für ihren fortführenden Lernprozess. Mittels alternativer Leistungsbeurteilung ließe sich außerdem besser auf die Ansprüche der Schüler*innen reagieren. Des Weiteren hatte der männliche Befragte in einem früheren Praktikum schon eine Form der alternativen Leistungsbeurteilung erprobt, wodurch er für sich ihren Erfolg und demnach Nutzen sehen konnte. Aus all diesen Begründungen kristallisierte sich die Perspektive der Studierenden, alternative Leistungsbeurteilung diene vorrangig den Kindern, heraus. An der Ziffernbenotung hielten sie eigentlich bloß aus Rücksichtnahme auf die Erwachsenen fest, weil vermutlich der Mehrheit jener ein alternatives und daher andersartiges, für sie ungewohntes Beurteilungsformat „Schwierigkeiten“ bereiten würde.

Positive Gefühle verbunden mit Innovative Grading oder negative Gemütsregungen, welche Ziffernbenotung anbelangten, drückten sich auch in entsprechender Färbung der diesbezüglichen Meinung aus. Jedoch offenbart sich hier durch sachliche Argumentationen der Befragten eine zuzügliche rationale Konstruktion der Einstellung (Gollwitzer & Schmitt, 2009). Diese stützte sich aber wiederum partiell auf persönliche Erfahrungen und deren innewohnende Empfindungen, was eventuell auf klassische Konditionierungen (Pavlov, 1927; zit. n. Kiesel & Koch, 2012) und/oder operantes Konditionieren (Winkel, Petermann & Petermann, 2006) als Quelle(n) hindeutet. Abgesehen davon ist anzunehmen, dass sich die Geisteshaltung der Studierenden desgleichen bis zu einem gewissen Grad via Modelllernen an der PHDL ausbildete.

Neben diesem Verhältnis zwischen Emotionen und Einstellungen tritt genauso eine potenziell existente Kausalität zwischen Einstellungen und Verhaltensintention zutage. Jene vier Interviewpartner*innen etwa, welche die Empfehlung von alternativer Leistungsbeurteilung anstrebten, teilten im Vorfeld geschlossen eine positive Gesinnung betreffend alternativer Leistungsbeurteilung. Im Falle der zwei Studierenden, die eine Mischform aus Ziffernbenotung und alternativer Leistungsbeurteilung am sinnvollsten erachteten, erweist sich die Relation hingegen nicht gänzlich. Beide äußerten zwar ein positives Verständnis von alternativer Leistungsbeurteilung; ihre neutrale beziehungsweise negative Auffassung von Ziffernbenotung widerspricht ihrer Intention aber. Hierbei muss ins Bewusstsein gerufen werden, dass eben nicht nur Einstellungen, sondern auch der wahrgenommene Erwartungsdruck von außen – die subjektive Norm – die Verhaltensintention beeinflusst. Jene Studentin und jener Student versuchten sich insofern nämlich an die möglichen Wünsche der Erziehungsberechtigten und die Gepflogenheiten weiterführender Schulen im Hinblick auf Ziffernbenotung anzupassen, obwohl sie vor allen Dingen Befürworter*innen der alternativen Leistungsbeurteilung waren.

Ansonsten scheint es, als verifiziert sich die Wirkkraft der subjektiven Norm auf die Verhaltensintention bloß zu 50 %. Denn während die drei weiblichen Studierenden ihr Verhalten an die vorhandene oder nicht gegebene Unterstützungsbereitschaft ihrer Mitmenschen, was die Empfehlung einer Art der Leistungsbeurteilung angeht, adaptieren würden, blieben die männlichen Studierenden, laut eigener Auskünfte, auf ihrem gewählten Kurs.

Grundsätzlich zeigte diese Masterthesis, dass Innovative Grading beinahe ausnahmslos positive Resonanz bei den Studierenden auslöste. Infolgedessen liegt die Vermutung nahe, ihre positiven Erfahrungen mit Innovative Grading prägten ihre generelle Einstellung zu alternativer Leistungsbeurteilung und bestärkten somit ihren Willen zu deren Einsatz. Drei limitierende Aspekte der Arbeit sollen erwähnt werden: a) die Größe der Stichprobe von sechs Personen, b) die Verhaltensüberzeugungen, auf welchen die Einstellungen basieren, und die normativen Überzeugungen, die als Fundament der subjektiven Normen gelten, waren weder im Leitfaden noch in den Interviews integriert und c) die Tatsache, dass die Interviews aktuelle Einstellungen der Studierenden darstellten, diese aber durch weitere Erfahrungen möglicherweise ihre Verhaltensintention modifizieren könnten.

Mit Blick auf den praktischen Alltag erlauben die Resultate die Schlussfolgerung, nicht die Quantität, sondern die Qualität der jeweiligen Art der Leistungsbeurteilung sei für die damit einhergehenden Erfahrungen/Emotionen und die wiederum daraus erwachsenden Einstellungen verantwortlich. Denn obwohl alle Interviewpartner*innen während ihrer Schul- und Studienzzeit hauptsächlich Ziffernoten als Beurteilung ihrer Leistung erhalten hatten und mit alternativen Beurteilungsformen in geringem Maße in Berührung gekommen waren, schälte sich diesbezüglich kein Hinweis auf eine Einflussnahme dieser oder jener Frequenz auf das damalige Erleben und/oder aktuelle Anschauungen heraus. Vielmehr hingen diese zwei Aspekte davon ab, auf welche Weise die dafür zuständigen Lehrpersonen die Leistungsbeurteilung umsetzten.

Zumal das österreichische Schulsystem seinen Schüler*innen augenscheinlich kaum Konfrontation mit alternativer Leistungsbeurteilung bietet, eröffnet sich für pädagogische Hochschulen die Chance, durch theoretische Vermittlung und praktische Anwendung unterschiedlicher alternativer Beurteilungsversionen eine Kompensation dieses Defizits zumindest für Lehramtsstudierende zu schaffen. Gleichzeitig ermöglichten sie mithilfe solcher Initiativen eine Erweiterung des Horizonts der jungen Pädagog*innen in diesem Bereich. Mithilfe zuzüglicher Informationen darüber, wie sich später Bündnisse mit Erziehungsberechtigten, dem Lehrer*innenkollegium und der Direktion schließen lassen, würde ihnen essenzielles Werkzeug an die Hand gegeben, um in Zukunft selbst neue Pfade der Leistungsbeurteilung zu beschreiten. Fügten die Professor*innen und Mentor*innen der Hochschulen noch eine Prise Ermutigung hinzu, so sollte der Weg für eine günstige Fortentwicklung des Bildungswesens Österreichs geebnet oder, genauer

gesagt, gesichert sein. Die wichtigste Voraussetzung dafür liegt, wie sich zweifellos zeigte, aber ohnehin bereits vor: der elementare, stark ausgeprägte Wille, alternative Leistungsbeurteilung in der Schulpraxis, zum Wohl des Kindes und seiner Entfaltung, zum Einsatz zu bringen.

4. Rückblick und Ausblick

Alternativen zur Ziffernnotenbeurteilung gibt es an der PHDL bereits seit dem WS 2007/08 durch eine Initiative von Praxisberater*innen. Diese stellten fest, dass die Notenbeurteilung in den Pädagogisch Praktischen Studien unpassend bzw. kontraproduktiv ist. Denn gerade eine pädagogische Ausbildung an hochschulischen Einrichtungen basiert auf einem elaborierteren Theorie-Praxis-Verständnis, mit dem Ziel, der tatsächlichen Komplexität pädagogischer Handlungszusammenhänge Rechnung zu tragen. Und pädagogische Handlungssituationen sind komplex, ...

- ... weil sie in ihrem Ausgang immer offen sind. Ob und wie Kinder und Erwachsene pädagogische Impulse aufgreifen und weiterspinnen, bleibt unbestimmbar.
- ... weil sie einzigartig sind. Es treffen Menschen mit einer nie dagewesenen Kombination von Lebenserfahrungen, Lebensweltbezügen und Handlungsqualitäten aufeinander und bringen noch nie Dagewesenes hervor.
- ... weil pädagogisches Handeln immer unter bestimmten (räumlichen) Bedingungen unmittelbar, d. h. unter zeitlichem Druck erfolgt. Pädagoginnen und Pädagogen sind gefordert, in immer neuen (sozialen) Konstellationen sofort Handlungsentscheidungen zu treffen (Thiersch, 2002).

Die Einführung der zweistufigen Beurteilung ergänzt durch eine ausführliche schriftliche Rückmeldung führte zu einer erheblichen Verbesserung der Praxisqualität. Endlich stand nun nicht mehr die Note, sondern die Kompetenzentwicklung im Vordergrund. Wenn man heute Studierenden oder Praxisberater*innen über die frühere Beurteilung der Pädagogisch Praktischen Studien erzählt, reagieren diese mit erstauntem Kopfschütteln.

Durch diesen Paradigmenwechsel äußerten zunehmend auch Lehrende aus den Fachbereichen Kunst, Musik und dem Querschnittsthema Persönlichkeitsbildung Zweifel an der Stimmigkeit von Ziffernbeurteilung für die Beurteilung von Kompetenzentwicklung in ihren Lehrveranstaltungen. Im WS 2017/18 startete das Forschungsprojekt Innovative Grading. Erstmals wurde erhoben, in wie vielen und welchen Lehrveranstaltungen eine andere Beurteilungsform als die Ziffernbeurteilung angestrebt wurde. Erste Veröffentlichungen, hausinterne Fortbildungen und fachbereichsinterner Austausch führten zu einer vermehrten Anwendung der zweistufigen Leistungsbeurteilung. Inzwischen hat sich Innovative Grading

an der PHDL etabliert. Im ersten Durchgang des Masterstudiums der Primarstufe (2019/20) wurden von den Autorinnen Themen ausgelobt und damit auch die Forschung zum Themenbereich Innovative Grading befeuert. Die Ergebnisse liegen in dieser Arbeit vor.

Welche Veränderungen sind nun im Studien- und Prüfungsalltag der PHDL ersichtlich? Rückblickend auf das Studienjahr 2019/20 ergibt sich folgendes Bild: Im Bachelorstudium Primarstufe wurden 12% aller Lehrveranstaltungen mit Pass/Fail beurteilt, im Bachelorstudium Elementarpädagogik (im Erhebungszeitraum waren erst vier Semester des sechssemestrigen Studiums ausgerollt) wurden 27% der Lehrveranstaltungen mit Pass/Fail beurteilt. Die beteiligten Fachbereiche waren neben den Pädagogisch Praktischen Studien auch Musik, Kunst, Bewegung und Sport sowie Lehrveranstaltungen aus der Querschnittthematik Persönlichkeitsbildung. Etwas weniger stark, aber ebenso beteiligt waren Lehrveranstaltungen aus der Bildungswissenschaft, dem wissenschaftlichen Arbeiten sowie der Fachdidaktiken (Sprachliche Bildung, Mathematik, Nawi).

Die Anzahl der insgesamt mitdenkenden und mitgestaltenden Personen war und ist erfreulich groß. Die rege Beteiligung am Innovative Grading quer durch alle Fachbereiche zeigt, dass mit einer neuen Beurteilungsform an der PHDL fast unbemerkt eine neue Prüfungskultur Einzug gehalten hat, die (vermutlich) mit einer neuen Lehr- und Lernkultur einhergeht. Es ist davon auszugehen, dass Innovative Grading als Beurteilungsform nur dann gewählt wird, wenn Bildung als dialogischer Prozess mit einer lebendigen Feedback-Kultur verstanden wird, was dem Leitbild der PHDL entspricht (https://www.phdl.at/ueber_uns/organisation/profil/leitbild/).

Mit dieser Veröffentlichung wird die langjährige und spannende Arbeit am Forschungsprojekt Innovative Grading abgeschlossen. Die PHDL zeigt damit beispielhaft vor, wie Leistungsfeststellung und Leistungsbeurteilung als ein wichtiges Themenfeld diskutiert, sukzessive optimiert und studierendenfreundlicher gestaltet werden kann. Die Sinnhaftigkeit und Legitimität der Leistungsbeurteilung mit einem zweistufigen Beurteilungsformat wurde dargestellt und bildet die Grundlage für ein Plädoyer dafür, Innovative Grading gleichwertig mit der Ziffernbeurteilung in der Prüfungsordnung der Studienpläne Primarstufe und Elementarpädagogik an der PHDL zu verankern.

Literatur

- Alm, R., & Colnerud, G. (2015). Teachers' experiences of unfair grading. *Educational Assessment, 20*, 132–150. <https://doi.org/10.1080/10627197.2015.1028620>
- Bastian, J., Combe, A., & Langer, R. (2016). *Feedback-Methoden. Erprobte Konzepte, evaluierte Erfahrungen* (4. erweiterte und überarbeitete Aufl.). Beltz.

- Bauer, J. (2017). Die pädagogische Beziehung. Neurowissenschaften und Pädagogik im Dialog. Ein Überblick unter besonderer Berücksichtigung der Vorschulzeit. *Lehren & Lernen* 1/2017.
- Baumgartner, K. (2021). *Motive für die Präferenz eines Formats der Leistungsbeurteilung*. Unveröffentlichte Masterarbeit Primarpädagogik an der Pädagogischen Hochschule der Diözese Linz.
- Beutel, S.-I., & Pant, H. A. (2019). *Lernen ohne Noten: Alternative Konzepte der Leistungsbeurteilung*. Kohlhammer.
- Bloodgood, R. A., Short, J. G., Jackson, J. M., & Martindale, J. R. (2009). A change to pass/fail grading in the first two years at one medical school results in improvement of psychological well-being. *Academic Medicine*, 84(5), 655–662. <https://doi.org/10.1097/acm.0bo13e31819f6d78>
- Brandstätter, V., Schüler, J., Puca, R. M., & Lozo, L. (2018). *Motivation und Emotion*. Springer.
- Brügelmann, H. (2006). Sind Noten nützlich – und nötig? Ziffernzensuren und ihre Alternativen im empirischen Vergleich. Eine Expertise der Arbeitsgruppe Primarstufe an der Universität Siegen im Auftrag des Grundschulverbands e. V., Frankfurt (Kurzfassung). In H. Bartnitzky, H. Brügelmann, U. Hecker, & G. Schönknecht (Hrsg.), *Pädagogische Leistungskultur* (S. 17–46). Grundschulverband – Arbeitskreis Grundschule e.V.
- Dahlgren, L. O., Fejes, A., Abrandt-Dahlgren, M., & Trowald, N. (2009). *Grading systems, features of assessment and students' approaches to learning*. *Teaching in Higher Education*, 14(2), 185–194. <https://doi.org/10.1080/13562510902757260>
- Fishbein, M., & Ajzen, I. (2015). *Predicting and changing behavior. The reasoned action approach*. Routledge.
- Gollwitzer, M. & Schmitt, M. (2009). *Sozialpsychologie kompakt*. Weinheim: Beltz.
- Guskey, T. R. (2015). *On your mark: Challenging the conventions of grading and reporting*. solution tree press.
- Hascher, T. (2004). *Wohlbefinden in der Schule*. Waxmann.
- Jacobs, B. (2009). *Leistungssteigerung durch Notendruck? – Die Wirkung der Benotung auf die Studierleistungen in einem Seminar*. <http://psydok.psycharchives.de/jspui/bitstream/20.500.11780/3347/1/notenanreiz.pdf>
- Kayser, J. (2008). Der Umgang mit dem Kompetenzbegriff. *Betrifft: Lehrerausbildung und Schule* (2), 10–17.
- Keller, S. (2011). Beurteilungsraster und Kompetenzmodelle. In W. Sacher & F. Winter (Hrsg.), *Diagnose und Beurteilung von Schülerleistungen – Grundlagen und Reformansätze* (S. 143–160). Schneider.
- Kiesel, A., & Koch, I. (2012). *Basiswissen Psychologie. Lernen. Grundlagen der Lernpsychologie*. VS.
- Kohn, A. (2011). The case against grades. *Educational Leadership*, 69(3), 28–33.

- Lamnek, S. (2005). *Qualitative Sozialforschung*. Beltz.
- Lotz, M. & Lipowsky, F. (2015). Die Hattie-Studie und ihre Bedeutung für den Unterricht – Ein Blick auf ausgewählte Aspekte der Lehrer-Schüler-Interaktion. In G. Mehlhorn, F. Schulz, & K. Schöppe (Hrsg.), *Begabungen entwickeln & Kreativität fördern* (S. 97–136). kopaed
- Lütgert, W., Tillmann, K.-J., Beutel, I., Jachmann, M., & Vollstädt, W. (2000). *Leistungsbeurteilung und Leistungsrückmeldung an Hamburger Schulen. Bericht über ein Forschungsprojekt*. reset.Grafische Medien GmbH.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (12. Aufl.). Beltz.
- McDuff, S. G. R., McDuff, D., Farace, J. A., Kelly, C. J., Savoia, M. C., & Mandel, J. (2014). Evaluating a grading change at UCSD school of medicine: pass/fail grading is associated with decreased performance on preclinical exams but unchanged performance on USMLE step 1 scores. *BMC Medical Education*, 14, 1–9. <https://doi.org/10.1186/1472-6920-14-127>
- Müller, A. (2013). *Die Schule schwänzt das Lernen. Und niemand sitzt nach*. hep.
- Neuweg, G. H. (2019). *Kompetenzorientierte Leistungsbeurteilung. Pädagogische und rechtliche Hilfestellungen für die Schulpraxis*. Trauner.
- Plöchl, M. (2021). *Der Einfluss unterschiedlicher Beurteilungsformen auf die Anstrengungsbereitschaft*. Unveröffentlichte Masterarbeit Primarpädagogik an der Pädagogischen Hochschule der Diözese Linz.
- Reitinger, J., Weinberger, A., Svoboda, U., Nausner, E., Vogl, U., Bachl, W., Bauer, M., Boxhofer, E., Grabner, V., Keplinger, G., Kittinger, C., Kitzmüller, B., Neuling, A., Seyfried, C., & Wöß, E. (2018). Innovative Leistungsfeststellung und Leistungsbeurteilung an der Privaten Pädagogischen Hochschule der Diözese Linz – Vision Statement der Interessensgemeinschaft Innovative Grading IG2. *Pädagogische Horizonte*, 2(1), 107–112.
- Reitinger, J., Nausner, E., & Seyfried, C. (2020). Pass/Fail-Beurteilung. Zifferzensur und Lernerleben in der Lehramtsausbildung. *Lehrerbildung auf dem Prüfstand*, (2), 199–217.
- Rheinberg, F. (1996). Von der Lernmotivation zur Lernleistung: Was liegt dazwischen? In J. Möller & O. Köller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 23 – 52). PVU.
- Rohe, D.E., Barrier, P.A., Clark, M.M., Cook, D.A., Vickers, K.S., & Decker, P.A. (2006). The Benefits of Pass-Fail Grading on Stress, Mood and Group Cohesion in Medical Students. *Mayo Clinical Proceedings*, 81(11), 1443–1448. <https://doi.org/10.4065/81.11.1443>
- Schmidinger, E., & Vierlinger, R. (2012). *Zeitgemäße Leistungsbeurteilung*. Jugend & Volk.

- Schweer, M. (2014). Lehrer-Schüler-Verhältnis. In W. Einsiedler, M. Götz, A. Hartinger, F. Heinzl, J. Kahlert, & U. Sandfuchs, (Hrsg.). *Handbuch Grundschulpädagogik und Grundschuldidaktik* (S. 251–255). Julius Klinkhardt.
- Strasser, J. (2020). *Erhöhen Kompetenzraster die Objektivität der Leistungsfeststellung?* Unveröffentlichte Masterarbeit Primarpädagogik an der Pädagogische Hochschule der Diözese Linz.
- Thiersch, Hans (2002): Sozialpädagogik – Handeln in Widersprüchen? In H.-U. Otto, T. Rauschenbach, & P. Vogel (Hrsg.), *Erziehungswissenschaft: Professionalität und Kompetenz* (S. 209–222). Springer.
- Valtin, R. (2002). Informationen zum Projekt NOVARA. In R. Valtin, C. Schmude, H. Rosenfeld, K. Darge, G. Ostrup, O. Thiel, M. Wagener & C. Wagner, *Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand* (S. 149–153). Juventa.
- Valtin, R., & Wagner, C. (2002). Wie wirken sich Notengebung und verbale Beurteilung auf die leistungsbezogene Persönlichkeitsentwicklung von Schülerinnen und Schülern aus? In R. Valtin, C. Schmude, H. Rosenfeld, K. Darge, G. Ostrup, O. Thiel, M. Wagener, & C. Wagner, *Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand* (S. 113–139). Juventa.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17– 31). Beltz.
- White, C. B., & Fantone, J. C. (2010). Pass-fail grading: laying the foundation for self-regulated learning. *Advances in Health Sciences Education*, 15, 469–477. <https://doi.org/10.1007/s10459-009-9211-1>
- Winkel, S., Petermann, F., & Petermann, U. (2006). *Lernpsychologie*. Ferdinand Schöningh.
- Ziener, G., & Kessler, M. (2012). *Kompetenzorientiert unterrichten? mit Methode*. Kallmeyer.
- Ziener, G. (2016). *Herausforderung Vielfalt. Kompetenzorientiert unterrichten zwischen Standardisierung und Individualisierung*. Friedrich.

