



WAS WIR LEISTEN, WENN WIR BEURTEILEN.

Altes und Neues zum (leidigen) Thema
Leistungsbeurteilung



ANTRITTSVORLESUNGEN AN DER KPH WIEN/KREMS: BAND 4

REIHE ANTRITTSVORLESUNGEN AN DER KPH, BAND 4

Eva Sattlberger

Was wir leisten, wenn wir beurteilen.

Altes und Neues zum (leidigen) Thema
Leistungsbeurteilung

Reihe: Antrittsvorlesungen an der KPH Wien/Krems, Band 4

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind unter <http://dnb.d-nb.de> abrufbar.

ISBN: 978-3-904046-03-9

IMPRESSUM

Kirchliche Pädagogische Hochschule Wien/Krems
Herausgeber: Rektorat der KPH Wien/Krems
Redaktion: Thomas Handschuh, Thomas Krobath, Doris Lindner

Layout/Satz/Grafik: Karin Gratiana Wurm
Druck: DRUCKEREI FUCHS, 2100 Korneuburg

Wien/Krems 2020 (1. Auflage Dezember 2020)

Vorwort

Das an der Kirchlichen Pädagogischen Hochschule Wien/Krems schon etablierte Format der Antrittsvorlesung zu neu eingerichteten Professuren wird in diesem Band fortgesetzt mit einer Nachschau zur eingerichteten Professur für Schulpädagogik und Allgemeine Didaktik.

Eva Sattlberger wählt zum Antritt ihrer Professur einen Vortrag zum Thema Leistungsbeurteilung und Leistungsfeststellung. Bereits im Titel des Vortrags „Was wir leisten, wenn wir beurteilen – Altes und Neues zum (leidigen) Thema Leistungsbeurteilung“ kommt die komplexe Gemengelage zum Ausdruck. Das Thema Leistungsbeurteilung stellt den geradezu paradigmatischen Schnittpunkt von lerntheoretischen Erkenntnissen, rechtlichen Vorgaben, schulpraktischen Sachzwängen und bildungspolitischen Forderungen dar. Kein anderes Thema mit Bildungsbezug wurde in den vergangenen Jahren intensiver und durchaus kontroversiell diskutiert als das Thema Leistungsbeurteilung, allen voran hinsichtlich der Reform zur Zentralmatura.

Leistungsbeurteilung zum Thema einer Antrittsvorlesung zu wählen zeugt also nicht nur von der Courage der Vortragenden, sondern ist auch Beleg für die institutionelle Bereitschaft, auf „heiße Eisen“ offen zuzugehen. Seit jeher versteht sich die Kirchliche Pädagogische Hochschule Wien/Krems als ein Ort, an dem wissenschaftlicher und schulpraktischer Diskurs aufeinandertreffen, offen verhandelt und im bestem Fall auch zu einem gemeinsamen Diskurs werden.

In diesem Sinne wünscht das Rektorat Eva Sattlberger viel Erfolg bei der Ausgestaltung der Professur.

Christoph Berger
Rektor der KPH Wien/Krems

Andreas Weissenböck
Vize-Rektor für Ausbildung

Vorwort	3
Inhalt	5
<i>Was wir leisten, wenn wir beurteilen – Altes und Neues zum (leidigen) Thema Leistungsbeurteilung</i>	7
1. Einleitung	7
2. Problemaufriss I	8
3. Problemaufriss II	21
4. Zukünftige Forschungsvorhaben	26
5. Ausblick	28
6. Literatur	29
<i>Akademischer Werdegang von HS-Prof.ⁱⁿ Mag.^a Dr.ⁱⁿ Eva Sattlberger</i>	31
<i>Ausgewählte Publikationen von HS-Prof.ⁱⁿ Mag.^a Dr.ⁱⁿ Eva Sattlberger</i>	32

Was wir leisten, wenn wir beurteilen

Altes und Neues zum (leidigen) Thema Leistungsbeurteilung

1. Einleitung

Ich begrüße Sie alle ganz herzlich zu meiner Antrittsvorlesung als Hochschulprofessorin für Schulpädagogik und Allgemeine Didaktik an der KPH Wien/Krems und freue mich, dass Sie heute gekommen sind, dass ihr heute gekommen seid.

Ich habe für diesen Vortrag das Thema Leistungsbeurteilung gewählt, da es mich in den vergangenen Jahren massiv beschäftigt hat und auch mit Sicherheit in Zukunft noch beschäftigen wird. Die Einführung der standardisierten Reife- und Diplomprüfung (SRDP) vor allem im Fach Mathematik an AHS hat in Österreich eine Diskussion um Leistungsbeurteilung wieder neu entfacht. Themen wie Notengerechtigkeit, Was bedeutet Allgemeinbildung (und gehört die Mathematik wirklich dazu)?, punktuelle Überprüfungen als Zeugnis erworbener Kompetenzen einerseits und andererseits Fragen zum Einfluss der Politik auf Reformen im Bildungsbereich sowie zur Umsetzungsproblematik von Idealvorstellungen, was die Beurteilung von Leistung betrifft, wurden vielseitig bespielt. Es soll ein Bogen gespannt werden von der Primarstufe über die Reifeprüfung und darüber hinaus zur Ausbildung von angehenden Primarstufenlehrer*innen.

2. Problemaufriss I

Leistungsbeurteilung ist sowohl im wissenschaftlichen als auch im privaten Bereich ein vielbesprochenes, sehr umfangreiches und in seinen Facetten insgesamt sehr komplexes Thema.

Allgemein wird zwischen *Leistungsfeststellung* und *Leistungsbeurteilung* unterschieden. Leistungsfeststellung meint den Prozess des Messens und Leistungsbeurteilung ist der Vorgang des Bewertens und der Gewichtung der gemessenen Leistung. Aus Sicht der Allgemeinen Didaktik (als Wissenschaft des schulunterrichtlichen Lernens und Lehrens) dienen Leistungsüberprüfungen und -bewertungen dem Erfassen von Lernwirkungen des geplanten Unterrichts. Es wird unter Berücksichtigung der Lernvoraussetzungen der Schüler*innen das Ausmaß der intendierten Lernstände erfasst. Dieser Ansatz findet sich auch in der gesetzlichen Grundlage, der sogenannten Leistungsbeurteilungsverordnung, wieder, wo Leistungsfeststellungen „zur Information darüber dienen, auf welchen Teilgebieten die Schüler die Lehrziele erreicht haben und auf welchen Teilgebieten noch ergänzender Unterricht nötig ist“ [LBVO 2020].

Eder, Neuweg und Thonhauser geben als Funktionen der Leistungsfeststellung und -beurteilung folgende Formen an [vgl. Eder, Neuweg & Thonhauser 2009: 248]:

Pädagogische Funktionen beziehen sich auf die Steuerung des Lehr-Lern-Geschehens. Wenn ein Unterrichtsinhalt Gegenstand einer Leistungsfeststellung wird, dann ist er bedeutsamer als jene Inhalte, die nicht geprüft werden (Signalfunktion). Leistungsfeststellungen ermöglichen den Lernenden auch Feedback über ihre Lernprozesse (Feedbackfunktion). Ebenfalls in den Bereich des pädagogisch Wirksamen fallen die Motivierungs- und Disziplinierungsfunktion durch Noten, die aber höchst umstritten sind. Im pädagogischen Sinn sollen Leistungsfeststellungen und -beurteilungen die Lernprozesse der Schüler*innen bestmöglich fördern.

Gesellschaftliche Funktionen betonen vor allem die Unterschiede zwischen den Schüler*innen. „Die Zuordnung zu unterschiedlichen Bewertungsklassen (Klassifizierungsfunktion) ist Voraussetzung für die Zuweisung zu verschiedenen Laufbahnen innerhalb und außerhalb der Schule (Allokationsfunktion) und zur Vergabe von Berechtigungen (Selektionsfunktion)“. [Eder, Neuweg & Thonhauser 2009: 248] Vor allem im Selektionsgeschehen wird hier die Note zum zentralen Instrument (Berichtsfunktion) und bietet eine scheinbar leicht überschaubare Information über den Erfolg bei der Bewältigung schulischer Anforderungen.

Personenbezogene Funktionen der Leistungsbeurteilung betreffen die Beeinflussung der Einstellung zur Schule. Sie wirken auf das Selbstkonzept und die Vorstellungen über die eigenen Leistungsmöglichkeiten der Schüler*innen. Oft ist Leistungsbeurteilung freilich auch mit Schul- bzw. Notenangst verbunden, unidirektionale Beurteilungsverfahren führen zu Ohnmachtserfahrungen, die Bildungskarrieren beeinflussen.

Leistungsmessungen müssen bestimmten Gütekriterien genügen, wenn ihre Ergebnisse sinnvoll verwertbar sein sollen:

- *Objektivität* bezeichnet den Grad, in welchem die Ergebnisse der Messung unabhängig von der Person des Messenden sind. Unterschiedliche Personen kommen zu denselben Ergebnissen.

- *Reliabilität* bedeutet, dass bestimmte Merkmale genau und zuverlässig gemessen werden. Mehrfache Messungen führen zu denselben Ergebnissen.
- *Validität* gewährleistet, dass das Messinstrument misst, was es zu messen vorgibt. [vgl. Sacher 2014: 35f]

Mit der Umsetzung der im Schuljahr 2014/15 an AHS bzw. 2015/16 an BHS flächendeckend eingeführten neuen Reife- und Diplomprüfung wurde im Wesentlichen versucht diesen grundlegenden Qualitätsmerkmalen von Leistungsbeurteilung näher zu kommen. So wurden durch die Standardisierung der Aufgabenstellungen in einzelnen Unterrichtsfächern „Klausurprüfungen unabhängig von einzelnen Lehrenden auf ein gemeinsames österreichisches Qualitätsniveau gebracht und damit objektiver. Die Grundidee der neuen Reife- und Diplomprüfung umfasst demnach die Standardisierung (der schriftlichen Klausuren) und die Kompetenzorientierung, durch die Kandidatinnen und Kandidaten ihre Teilprüfungen hinsichtlich klar definierter Anforderungen ablegen, die auch empirisch gut ausgewiesen werden können.“ [BIFIE 2013: 2]

Die Schlagworte sind also *Objektivität* – um in der Vergangenheit doch relativ unterschiedlichen Prüfungsanforderungen im Rahmen der Matura entgegenzuwirken –, *Standardisierung* und *Kompetenzorientierung*.

Weiters heißt es im sogenannten Grundsatzpapier zur SRDP: Die standardisierte Reife- und Diplomprüfung „garantiert Schülerinnen und Schülern durch verbindliche Beurteilungsrichtlinien mehr Fairness und Objektivität in der Beurteilung ihrer Leistungen an einer zentralen Schnittstelle in ihrer Bildungslaufbahn. Tertiären Institutionen und künftigen Arbeitgebern bietet sie zuverlässigere Aussagen über die tatsächlich erworbenen Kompetenzen der Schulabgänger/innen.“ [BIFIE 2013: 4]

In den meisten Fächern hat dieser Paradigmenwechsel auch zu keiner allzu lang dauernden bzw. allzu großen Aufregung geführt, also eigentlich in allen Fächern außer Mathematik und hier vor allem in der AHS.

Hier einige Zitate aus den Medien:

„So eine Matura gibt man nicht“ [Hämmerle 2016]

„Mathematik Matura: Wer das nicht versteht, sollte nicht studieren.“ [Beba 2016]

„Schüler frustriert. Mathe-Matura bekommt ein ‚Nicht Genügend‘. Eine verärgerte AHS-Lehrerin zur ‚Krone‘: ‚Warum muss man Kinder dermaßen quälen?‘“ [Kronen Zeitung 2019]

Was hat hier vermeintlich so schief gehen können, dass die Aufregung über die Mathematikmatura an der AHS so groß war? Meine These dazu ist, dass zwar unumstritten ist, dass Maturant*innen bestimmte Kompetenzen beherrschen sollen – für meine Ausführungen heute stelle ich den Kompetenzbegriff an sich nicht in Frage, Kompetenzen sind vereinfacht gesprochen formulierte Lernziele, welche die Inhalte der einzelnen Lehrpläne präzisieren. Die Einführung der jeweiligen Lehrpläne in den Jahren 2000 bzw. 2004 haben in den einzelnen Fächern inhaltlich auch keine großen Aufregungen hervorgerufen. Schüler*innen wird durch diese Präzisierung einerseits garantiert, dass sie nach zwölf (oder mehr) Schuljahren ein gewisses Grundgerüst an Inhalten garantiert mitnehmen. Für weiterführende Institutionen dienen die Kompetenzkataloge bzw. die Reifeprüfung als Anhaltspunkte für das Ausbildungs-

niveau. Der Streit oder die Aufregung hat sich letztendlich daran entzündet, dass unterschiedliche Personen ein recht unterschiedliches Verständnis von den *klaren Anforderungen*, d.h. vom *Anspruchsniveau* der Klausurprüfungen hatten, und damit auch vom Leistungsbegriff.

Im Zuge meiner Tätigkeit als Verantwortliche für die Standardisierte Reifeprüfung in Mathematik an AHS habe ich praktisch keinen Prüfungstermin erlebt, dessen Anspruchsniveau nicht beanstandet wurde. In Hoch-Zeiten gab es Tage, an denen ich zuerst Mails/Anfragen/Briefe beantworten musste, die das viel zu hohe Anspruchsniveau der Aufgaben kritisiert haben, um als Nächstes Schreiben zu beantworten, welche das viel zu niedrige Niveau als Zeichen für den Verfall der gymnasialen Mathematikausbildung gesehen haben („viel zu einfach“, „einer Matura nicht würdig“, „nur Unterstufenrechnungen“, usw.). Das Anspruchsniveau dieser Prüfung wurde also öffentlich diskutiert und beanstandet – wohlgemerkt in beide Richtungen. Daraufhin gab es – meiner Meinung nach oft jeder logischen Grundlage entbehrende – vor allem politisch motivierte Änderungen/Entscheidungen, um den angeblich zu hohen Selektionsprozess der SRDP Mathematik aufzuweichen ¹.

Die oben beschriebenen Diskussionen und medialen Aufregungen kamen für die Verantwortlichen der Bildungspolitik großteils überraschend, stellt doch die standardisierte Reifeprüfung eine logische Konsequenz zu den Bildungsstandardtestungen auf den Schulstufen 4 und 8 dar. Diese wurden aber in der Öffentlichkeit weitaus weniger diskutiert als die Reifeprüfung.

Bei beiden Testungen (Bildungsstandards und SRDP) wird die Idee der Outputorientierung im Bildungssystem umgesetzt. 2004 schon konkretisierte die sogenannte Zukunftskommission [vgl. Zukunftskommission 2004: 1, zit. nach Schreiner & Wiesner 2017: 18] in einem Grundsatzpapier die Bedeutung der Bildungsstandards: Lehrpläne würden „hauptsächlich auf den Inhalt des Unterrichts“ (also den Input) fokussieren, was zu einer stark kontextbezogenen Leistungsbeurteilung führt, die über die tatsächliche Erreichung von Grundkompetenzen (als „Output“) wenig aussagen würde. Will man also etwas über den Output eines Bildungssystems wissen, dann ist der Nachweis von klar ausgewiesenen Kompetenzen zu einem bestimmten Zeitpunkt erforderlich. Bei der standardisierten Reifeprüfung geschieht dies im Rahmen einer Berechtigungsvergabe, bei den Bildungsstandards im Sinne eines allgemeinen Systemmonitorings.

Im Konzept der Bildungsstandards findet sich folgende Formulierung: „Bildungsstandards legen aus Sicht der Schüler/innen konkret formulierte Lernergebnisse in Form von Könnensbeschreibungen (...) fest, die aus den Lehrplänen der Schularten und Schulstufen abgeleitet wurden.“ [BIFIE 2020] In diesem Sinne ist also die Grundlage bzw. der Zugang ähnlich wie bei der SRDP: Aus den durch die Lehrpläne festgelegten Inhalten werden Kompetenzen abgeleitet. Sowohl mit den Bildungsstandardüberprüfungen als auch mit der SRDP möchte man wissen, ob die Schüler*innen diese Kompetenzen auch beherrschen. Beide Zugänge verwenden für diese Leistungsmessung aus Aufgaben zusammengesetzte Tests, die von externen (im Gegen-

¹ Ein Tiefpunkt in jener Zeit, als ich für die SRP in Mathematik (AHS) zuständig war, war z. B. die Entscheidung auf höherer Ebene, dass künftig für Typ 1-Aufgaben halbe Punkte vergeben werden – eine meines Erachtens völlig absurde Zugangsweise bei der Bewertung von Leistung, die im Anschluss an die Schule praktisch nirgendwo mehr auftritt. Ich ziehe hier den Vergleich mit dem Abbiegen bei einer Führerscheinprüfung, wo der Kandidat*die Kandidatin zwar geblinkt, aber vergessen hat zu lenken und daher gegen ein Hindernis gefahren ist. Niemand würde in dieser Situation dem Kandidaten*der Kandidatin einen halben Punkt gewähren.

satz zu schulinternen Lehrer*innen) Personen bzw. Gruppen von Expert*innen zusammengestellt und qualitätsgesichert wurden. Im Unterschied zur SRDP, die durch die Vergabe einer Berechtigung auch eine Selektionsfunktion erfüllt, hat das Erreichen oder Nichterreichen der Bildungsstandards für die Schüler*innen keinerlei Konsequenzen. Daher fehlt bislang auch die Diskussion zum Anspruchsniveau der Testungen bei den Bildungsstandards. Diese wurde bei der SRDP ja auch erst dann geführt, als die ersten Noten vergeben wurden und die damit einhergehenden Konsequenzen getragen werden mussten.

Mein erstes Fazit lautet also: Standardisierte Testungen basieren auf aus Lehrplänen formulierten Grundkompetenzen und erfüllen im Sinne einer größeren Transparenz der Leistungsanforderungen eine Orientierungsfunktion sowohl für Schüler*innen als auch für Lehrer*innen – und das ist gut so. Das Anspruchsniveau der Messinstrumente muss aber (gesellschaftlich?) vorab offenbar ausdiskutiert werden, sofern die Testungen (offiziell oder inoffiziell) mit einer Berechtigungsfunktion einhergehen.²

Standardisierte Testungen ermöglichen viele Auswertungen, die neben der Messung von Leistungen grundlegende Informationen über unser Bildungssystem ermöglichen. Exemplarisch möchte ich hier auf einzelne Ergebnisse eingehen.

Bestimmte Gruppen werden in unserem Bildungssystem benachteiligt.

Die Ergebnisse der ersten Durchgänge der SRDP haben z. B. gezeigt, dass Mädchen und Burschen durchaus unterschiedliche Ergebnisse erzielen. Exemplarisch seien hier die Ergebnisse des Haupttermins 2018 aus dem Portal der Statistik Austria angeführt: Sowohl in Mathematik als auch in Englisch haben die Mädchen an AHS und BHS bei der schriftlichen Prüfung (also vor der mündlichen Kompensationsprüfung) schlechter abgeschnitten als die Burschen. So erhielten beim Sommertermin 2017/2018 der Abschlussjahrgänge an AHS 18,0% der Burschen und 25,5% der Mädchen ein Nicht genügend in Mathematik (gesamt 22,4%), in Englisch 7,2% der Burschen und 8,8% der Mädchen (gesamt 8,1%). An BHS erhielten 13,1% der Burschen und 25,5% der Mädchen eine negative Note in Angewandter Mathematik (gesamt 19,7) und in Englisch 9,1% der Burschen und 16,1% der Mädchen (gesamt 12,9%)³. Ähnliche Ergebnisse konnten auch in den Prüfungsterminen der Vorjahre beobachtet werden. Die unterschiedlichen Leistungen der Mädchen und Burschen sollen hier nur als ein Beispiel angeführt werden.

Ähnliche Ergebnisse für Mathematikleistungen können auch in anderen Überprüfungen beobachtet werden. Die Ursachen dafür sind höchst unterschiedlich und sehr komplex, darauf möchte ich aber heute nicht eingehen. Als Stichwort sollen hier nur die unterschiedlichen Zugänge zu eigenen Leistungszuschreibungen von Mädchen und Burschen genannt werden, diese spiegeln sich im sogenannten Selbstkonzept wider, welches einen großen Einfluss auf die tatsächlichen Leistungen hat.

² Wobei ich der Meinung bin, dass das Anspruchsniveau prinzipiell gar nicht ausdiskutiert werden kann.

³ Prüfungsergebnisse vor der Kompensationsprüfung, Quelle: STATISTIK AUSTRIA, SRDP Abschlussjahrgang 2017/18 Sommertermin

Unser Bildungssystem ist sozial ungerecht.

Untersuchungen im Rahmen der Bildungsstandardüberprüfungen erheben Daten der Schüler*innen, die auf ihren sozialen Hintergrund schließen lassen (dies ist bei der SRDP [noch] nicht der Fall). Diese werden mit den Leistungsdaten verknüpft und ergeben ein Bild, das im Dezember 2019 auch bei den PISA-Ergebnissen (erneut) bestätigt wurde.

Hier exemplarisch ein Ergebnis aus der Bildungsstandardtestung 2018 der 4. Schulstufe (also am Ende der Volksschule): „Zahlreiche internationale Studien, wie beispielsweise PISA 2015 oder PIRLS und TIMSS 2011 (...) sowie die bisherigen Standardüberprüfungen (...) belegen, dass Schülerleistungen mit demografischen Merkmalen, wie Geschlecht, Migration oder familiärer Herkunft zusammenhängen. Vor allem die Bildung der Eltern erscheint als starker Indikator für den Kompetenzerwerb bei den Schülerinnen und Schülern.“ [BIFIE 2019: 40]

Bruneforth, Weber und Bacher führen die Chancengleichheit unter einer normativ-legistischen Begründung als zentrales bildungspolitisches Ziel an. „Zusammenfassend zeigt sich folgendes Bild zur Chancengleichheit. Die soziale Herkunft ist die zentrale Ungleichheitsdimension. Sie wirkt sich durchgehend in der Bildungslaufbahn auf den Kompetenzerwerb und den Schulbesuch aus. Ihre Wirkung bleibt auch bestehen, wenn der Einfluss anderer Merkmale statistisch kontrolliert wird. Die anderen Merkmale folgen mit deutlichem Abstand in der Bedeutung: Kinder von Eltern mit geringer Bildung zeigen in der 4. Schulstufe deutlich schwächere Leistungen. Ihr Rückstand im Lernfortschritt liegt bei mehr als zwei Schuljahren (21,7 Schulmonate) hinter den Kindern von Eltern mit Maturaabschluss. Für die 8. Schulstufe ergeben sich ähnliche Befunde.“ [Bruneforth, Weber & Bacher 2012: 200]

Es gäbe hier noch sehr viel darüber zu sprechen, ich möchte nur zusammenfassend anmerken, dass bestimmte Schüler*innengruppen in unserem Bildungssystem zu wenig Förderung erhalten, dies wirkt sich auch auf die Bildungslaufbahnen aus. Es entstehen zudem sogenannte Kompositionseffekte: Die schlechteren Startbedingungen dieser Kinder werden in der Schule nicht kompensiert und sie besuchen oft Klassen mit Rahmenbedingungen, die zusätzlich zu einem schlechteren Lernerfolg führen.

Weiters zeigen Analysen zu den PISA-Erhebungen, dass sich das Erstselektionsalter als empirisch besonders bedeutsam für die Chancengleichheit erwiesen hat [vgl. Bruneforth, Weber & Bacher 2012: 202]. Abbildung 1 soll dies verdeutlichen. Sie zeigt die Abhängigkeit der Leseleistung vom sozioökonomischen Status der Eltern in verschiedenen Ländern und einen Zusammenhang mit dem Erstselektionsalter.

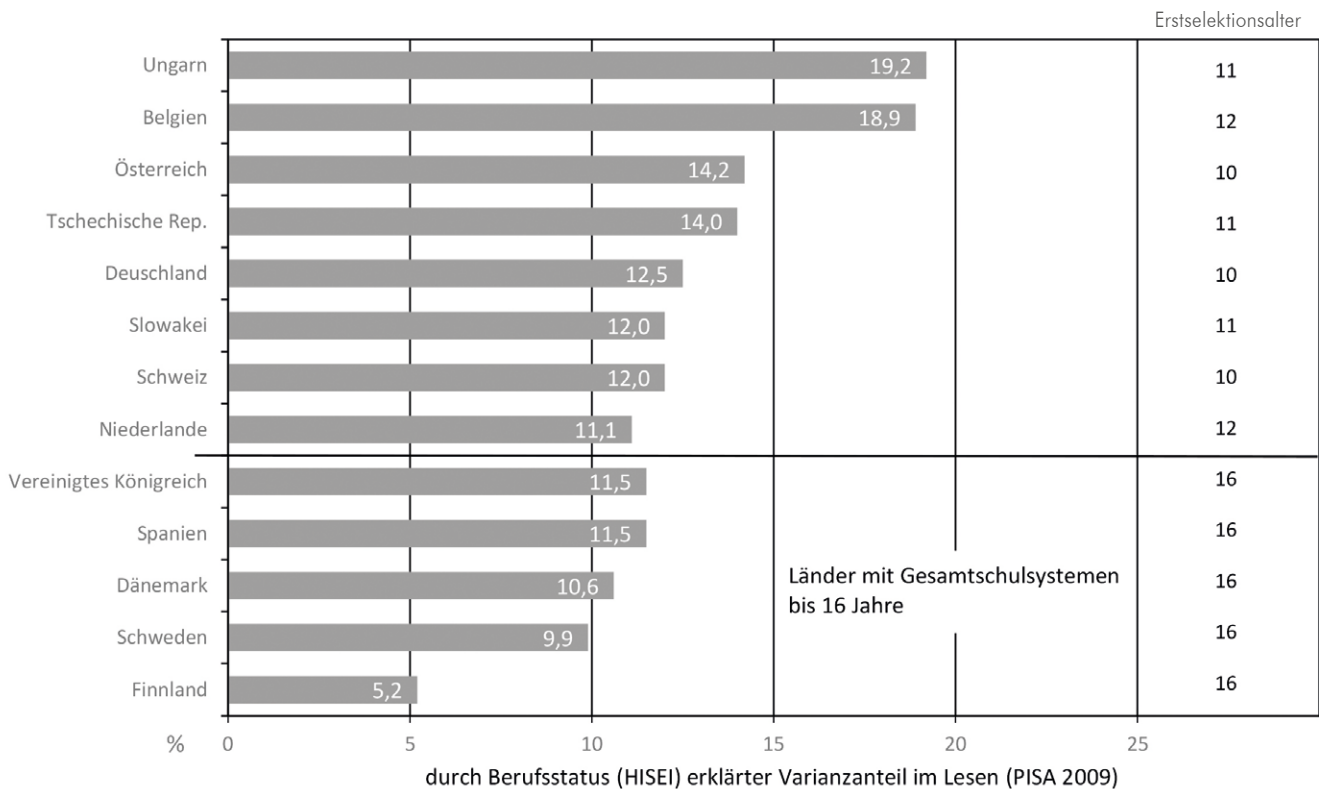


Abbildung 1: Abhängigkeit der Leseleistung vom sozioökonomischen Status der Eltern in unterschiedlichen Ländern [Bruneforth, Weber & Bacher 2012: 205, Abb. 5.2, Quelle PISA 2009, eigene Bearbeitung]

*Noten spiegeln nicht immer die von außen gemessenen Testleistungen der Schüler*innen wider.*

De facto wird in Österreich nicht mit zehn Jahren selektiert, sondern ausschlaggebend ist *bis dato* das Semesterzeugnis der 4. Schulstufe als Entscheidungsgrundlage für Gymnasialreife oder nicht. Zu dem Zeitpunkt wo über ihre Zukunft entschieden wird, sind die Schüler*innen meist noch neun Jahre alt. Künftig soll dieser Umstand aber noch verschärft werden. So liest sich zumindest diesbezüglich das aktuelle Regierungsabkommen, in dem es heißt: „Die Entscheidung über die weiterführende Bildungslaufbahn soll nicht mehr nur von einer Leistungsfeststellung (Schulnachricht der vierten Schulstufe) abhängig gemacht werden, sondern auf der Basis der Ergebnisse einer individualisierten Kompetenzfeststellung in der dritten Schulstufe, des Jahreszeugnisses der dritten Klasse und der Schulnachricht der vierten Klasse getroffen werden.“⁴ [Bundeskanzleramt 2020: 209]

Schmid, Paasch & Katstaller [2016] beschäftigen sich mit dem Thema der leistungsangemessenen Benotung durch Lehrkräfte in der Volksschule. Sie gehen dabei davon aus, dass aufgrund rechtlicher Bestimmungen Chancengleichheit für Schüler*innen dahingehend gewährleistet werden soll, dass gleiche Noten ein vergleichbares Leistungsniveau zum Ausdruck bringen. In ihrem Artikel erwähnen sie, dass die Notenverteilung am Ende der Primarstufe – wenn es

⁴ Davon abgesehen, dass der Begriff der Leistungsfeststellung hier falsch verwendet wird, wird das Erstselektionsalter de facto noch weiter nach unten verschoben. Bereits in der 3. Klasse, also bei Kindern mit ca. acht Jahren, wird über ihre weitere Zukunft entschieden.

um den Übergang in eine weiterführende Schule geht – eine rechtsschiefe Verteilung aufweist, d.h. dass Schüler*innen im Schnitt besser benotet werden als ihre Leistungen bei den Bildungsstandardtestungen dies erwarten lassen würden. „Im Unterschied zu den Testleistungen, die (qua Testkonstruktion) normalverteilt sind, sind die Noten wie erwartet rechtsschief verteilt: 43% erhielten im Halbjahreszeugnis die Note 1, 33% die Note 2, 17 % die Note 3, 6% die Note 4 und 1 % die Note 5. (...) Schüler/innen aus Elternhäusern mit einem höheren sozioökonomischen Status (SES) erhalten bei gleichen Testleistungen signifikant bessere Noten als jene aus Elternhäusern mit einem niedrigeren SES.“ [Schmid, Paasch & Katstaller 2016: 276] Ein weiterer Befund aus dieser Studie zeigt, dass innerhalb von Klassen Kinder mit Migrationshintergrund bei gleichen Testleistungen, gleichem sozioökonomischen Status und gleichem Geschlecht etwas schlechtere Noten erhalten als Kinder ohne Migrationshintergrund. Demnach spiegeln gleiche Noten nicht gleiche Leistungen wider, was insofern sehr kritisch zu betrachten ist, als eben die Mathematiknote des Halbjahreszeugnisses für den Übergang in die Sekundarstufe 1 ein wichtiges Kriterium darstellt. [vgl. Schmid, Paasch & Katstaller 2016: 278]

Diese Ergebnisse sind relativ beunruhigend, wobei die vielfältige Problematik der Notengebung an sich nicht neu ist. Sacher [2014: 42] schreibt z. B. „Die Objektivität der schulischen Leistungsüberprüfungen und -bewertungen müsste sich darin zeigen, dass verschiedene Lehrkräfte dieselbe Schülerleistung gleich beurteilen. In vielen entsprechenden Studien fand man aber beträchtliche Beurteilungsdivergenzen.“ Die Beurteilungsunterschiede gehen z. T. auf unterschiedliche Persönlichkeitsmerkmale zurück: Alter, Geschlecht, Ausbildung und berufliche Motivation, Auffassung der eigenen Rolle, Vorstellungen über die zu Prüfenden und dgl. beeinflussen die Notengebung maßgeblich. Ähnliches gilt auch für die Reliabilität. Aus Untersuchungen ist längst bekannt, dass Lehrkräfte zu unterschiedlichen Zeiten dieselbe Leistung unterschiedlich beurteilen und dass die Reihenfolge der Korrektur Einfluss auf die Beurteilung hat [vgl. dazu Sacher 2014, Neuweg 2019, ...]. Zur Validität der Leistungsbeurteilung schreibt Neuweg [2019: 5], dass „keineswegs eindeutig festgelegt ist, was die Lehrerin bzw. der Lehrer überhaupt zu messen hat. Das „Messobjekt“ ist ihr bzw. ihm durch die recht allgemein gehaltenen Lehrpläne nur sehr unbestimmt vorgegeben, sodass verschiedene Lehrerinnen und Lehrer verschiedene Vorstellungen darüber entwickeln, welche Schülerleistungen überhaupt gemessen werden sollen.“ Die Lehrer*innen bestimmen, was zu messen ist, fertigen dazu das Messinstrument an, führen die Messung selbst durch und nehmen auch die Bewertung vor, wobei das zu Messende auf den von den Lehrer*innen durchgeführten Unterricht zurückzuführen ist [vgl. Kiper 2001: 22, zit. nach Benischek 2006b: 296].

In ihrem Beitrag zum Nationalen Bildungsbericht 2009 schreiben Eder, Neuweg & Thonhauser: „Die Heranziehung der lehrzielbezogenen Norm ist – v.a. wenn diese für individuelle Rückmeldungen um die individuelle Norm ergänzt wird – in der erziehungswissenschaftlichen Literatur unumstritten. Sie ist die vorzugswürdige Norm

- im Interesse der Lehr-/Lern-Steuerung,
- unter dem Aspekt der Beurteilungsgerechtigkeit,
- mit Blick auf die Berichtsfunktion von Noten und Zeugnissen und
- unter dem Gesichtspunkt der Hinführung der Schüler/innen zur sachlich begründeten Selbsteinschätzung.“ [Eder, Neuweg & Thonhauser 2009: 249]

Breidenstein & Zaborowski unterstreichen, „[dass] Zensuren wenig geeignet sind für Zwecke einer pädagogisch-didaktischen Leistungsrückmeldung (...) stellt einen relativ weitreichenden Konsens in der schulpädagogischen Diskussion dar. Dafür sind Zensuren zu undifferenziert, zu informationsarm und zu sehr dem sozialen Vergleich verhaftet.“ [zit. nach Friedrich & Heim 2019: 885]. Schreiner und Kraler [2019: 866] fordern darüber hinaus, dass sich pädagogisch-diagnostisches Handeln weniger über die Betonung einer unidirektionalen mehrstelligen Relation darstellen sollte, sondern als Interaktion zwischen diagnostizierender und diagnostizierter Person. Dies würde auch Lehrer*innen die Möglichkeit eröffnen, Rückschlüsse auf den eigenen Unterricht zu ziehen.

Hierzu ein kleiner Exkurs: Ich habe mich im Bildungsministerium erkundigt wie viele Schulversuche es in den vergangenen Jahren im Bereich „Alternative Leistungsbeurteilung“ – also Bewertungen von Schüler*innenleistungen ohne Notenvergabe mit Pensensbüchern oder dgl. – gegeben hat. Zurzeit gibt es lediglich einen diesbezüglichen Schulversuch in der „Lernwerkstatt Brigittenu“. Das war aber nicht immer so. Schulversuche zur Alternativen Leistungsbeurteilung gibt es in Österreich schon seit 34 Jahren. Bis 2015/16 hatten von ca. 3100 Volksschulen 2165 Schulen (5292 Klassen bzw. 94706 Schüler*innen) einen derartigen Schulversuch. Nach 2016 konnte man die genauen Zahlen nicht mehr erheben, da auf Empfehlung des Rechnungshofes möglichst viele Schulversuche in die Autonomie der Schulen übergeführt wurden und daher nicht mehr gemeldet werden mussten. Die Entscheidung, ob Schüler*innen alternativ beurteilt werden sollen, lag in diesem Zeitraum im Klassen- bzw. Schulforum, war also eine gemeinsame Entscheidung von Schule, Eltern und Lehrer*innen. Sie konnte bis zum Ende der 3. Schulstufe getroffen werden. Die große Anzahl der Volksschulen mit alternativer Leistungsbeurteilung zeigt, dass diese Forderung vor allem aus der Praxis kam, also von jenen, die tagtäglich mit Leistungsbeurteilung umgehen mussten.⁵

Wie sehen die aktuellen Entwicklungen in diesem Bereich in Österreich aus? Zurzeit wird der Lehrplan der Volksschule kompetenzorientiert formuliert und parallel dazu werden Kompetenzraster erprobt, die eine Übersetzung der erreichten Lernziele in Noten gewährleisten können. Nach also über 30 Jahren Beobachtung der Praxis und nach zehn Jahren an Überlegungen zu den Ergebnissen aus den Nationalen Bildungsberichten hat das Bildungsministerium auf die Schulpraxis reagiert. Das ist prinzipiell gutzuheißen, wäre da nicht die Gesetzesänderung durch die zuständige Bundesregierung gewesen, die Noten ab der 2. Schulstufe wieder verbindlich eingeführt hat und die auch im aktuellen Regierungsprogramm zu finden ist. Bildungsminister Faßmann meinte in einer hoffentlich sehr verkürzten Darstellung, dass ihn viele Eltern angerufen haben, weil sie wieder Noten wollten und dass deshalb auch wieder Noten gegeben werden müssen. Ich wünschte es hätten ihn Lehrer*innen angerufen oder Studienautor*innen⁶.

Als weiteres, neues Reformanliegen im Rahmen des Pädagogik-Pakets 2018 soll es in Zukunft statt den Bildungsstandardüberprüfungen (diese sind nach zwei Durchgängen laut Beschluss des Ministeriums bereits wieder Geschichte) eine sogenannte informelle Kompetenz- und Potentialmessung (IKPM) geben. Diese dient einerseits der Diagnose von Kompetenzen und

5 Eine weitere Forderung aus der Praxis war – laut Auskunft des Ministeriums – eine verbindliche Anleitung, wie man die verwendeten Kompetenzraster möglichst gut in Noten umsetzen kann, um hier für gleiche Leistungen gleiche Noten vergeben zu können.

6 Im Juni 2020 wurde eine Petition „Die Ermöglichung der alternativen Leistungsbeurteilung ohne Noten im Rahmen der Schulautonomie“ mit ca. 12 000 Unterschriften im Parlament zur weiteren Behandlung eingereicht.

der Förderung der Schüler*innen, andererseits aber auch dem Schulqualitätsmanagement und dem Bildungsministerium zum Monitoring der Leistung und zur Entwicklung des Bildungssystems. Sie soll eine Weiterentwicklung der beiden bestehenden Instrumente Bildungsstandardüberprüfungen und Informelle Kompetenzmessung (IKM) darstellen und wird diese künftig ersetzen bzw. ergänzen. [vgl. BMBWF 2019]⁷ Getestet werden soll jedes Jahr flächendeckend mit einem 45 Minuten dauernden Test pro Fach (D und M in der Volksschule und D, M, E in der Sekundarstufe 1). Zur Vorbereitung darauf soll ab Herbst 2020 eine verpflichtende Nutzung der bestehenden IKM vorgeschrieben werden.⁸

Genauere Informationen über diese IKP-Messungen zu bekommen war sehr schwierig. Es gibt praktisch keine offiziellen Dokumente. So etwas macht immer ein wenig stutzig ... Warum? Weil es z. B. eine Testung sein soll, die gleichzeitig als Systemmonitoring und Individualdiagnose dienen soll. Dies ist testtheoretisch besonders hinterfragenswert: Kompetenzmessungen auf Individualebene werden normalerweise über einen Zeitraum von 1,5 bis zwei Stunden angesetzt. Dabei werden meist verschiedene Aufgaben gestellt, wobei mehrere Aufgaben sich auf eine zu messende Kompetenz beziehen – da es sonst sehr stark davon abhängt, welche Einzelaufgaben für eine bestimmte Kompetenz gegeben werden. Die vorgesehene Testung soll aber nur 45 Minuten dauern. So können zwar auf Systemebene Rückmeldungen entstehen, für die Individualdiagnose mit dieser Tragweite wäre dieser Test aber zu kurz. Zudem ist dieses Testinstrument so angelegt, dass sowohl in der 3. und 4. Schulstufe als auch in der 7. und 8. Schulstufe eine Überprüfung der Schüler*innen erfolgen soll, um so auch den individuellen Lernfortschritt feststellen zu können. Es könnten also in Zukunft Kinder bereits im Alter von acht Jahren, aufgrund momentaner Aufnahmen, in ein Selbstbild kippen, das sie ihre ganze Schullaufbahn hindurch nicht mehr loswerden.

Prinzipiell ist es sehr begrüßenswert Lehrer*innen Instrumente zur Verfügung zu stellen, die ihnen bei der Unterrichtsplanung helfen und gleichzeitig ermöglichen, ihre Schüler*innen zielgenau zu fördern. Zusammen mit meinen obigen Ausführungen, dass die Lehrpläne kompetenzorientiert formuliert werden und Lehrer*innen Kompetenzraster zur Verfügung stehen, an denen sie sich orientieren können, könnte man sagen: Es geht in die richtige Richtung.

Dennoch müssen diesbezüglich ein paar grundlegende Fragen gestellt bzw. Überlegungen beachtet werden:

- *Das Anspruchsniveau der einzelnen IKPM ist nicht definiert.*

Genau wie bei der standardisierten Reifeprüfung kommt es im Endeffekt auf die einzelnen Testhefte an, mit denen die Kompetenzen der Schüler*innen zu einem bestimmten Zeitpunkt gemessen werden. Laut Auskunft des BIFIE⁹ durchlaufen die Aufgaben zwar ein sogenanntes Standardsetting, mit dem die Testheftschwierigkeit bestimmt werden kann, aber wie man bei der Reifeprüfung in Mathematik gesehen hat, wo die Aufgaben auch ein Standardsetting durchlaufen haben, war das

⁷ Der derzeitige Stand (August 2020) der Website des Ministeriums zur IKPM sieht keine Potentialmessung mehr vor, sondern eine sogenannte individuelle Kompetenzmessung (iKM^{plus}). Zudem ist vorgesehen, dass die Ergebnisse nicht in die Leistungsbeurteilung miteinfließen und nicht als Entscheidungsgrundlage zur Aufnahme an eine höhere Schule herangezogen werden [vgl. BMBWF 2020]. Anscheinend ist das Ministerium diesbezüglich lernfähig.

⁸ Nach derzeitigem Stand wird eine Nutzung nur mehr empfohlen.

⁹ Am 1. Juli 2020 wurde aus dem BIFIE gemäß IQS-Gesetz, BGBl. I Nr. 50/2019, das IQS – Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen.

der öffentlichen Meinung im Endeffekt egal. Hinzu kommt, dass es diesbezüglich ja keinerlei Erfahrungen gibt. Für die 3. Schulstufe hat es im Rahmen der Bildungsstandards zwar Baselinetestungen gegeben, diese beziehen sich aber auf eine ganz andere Schüler*innenpopulation. Laut BIFIE laufen hier gerade Pilotierungen. Es wird also versucht die bisherigen Daten mit den neuen zu verlinken, wohlgermerkt ohne dass der neue Lehrplan fertig ist. Demnach können die Lehrer*innen auch nicht nach dem neuen Lehrplan unterrichten und Kompetenzraster gibt es auch noch keine dazu. Es herrschen also völlig andere Voraussetzungen.

- *Die Bildungsstandards waren als Regelstandards formuliert, für die IKPM (bzw. iKM^{Plus}) braucht man meiner Meinung nach Mindeststandards, also die Vereinbarung, welche Inhalte mindestens für eine bestimmte Note gekonnt werden müssen. Sonst haben ja auch die gerade in Entwicklung stehenden Kompetenzraster keinen Sinn. Hier stellt sich die Frage wie die Setzung dieser Mindeststandards erfolgt und wie diese in eine Testung umgesetzt werden. Wer verhandelt dies?*
- *Laut Auskunft des BIFIE muss die Umsetzung der geplanten IKPM nach einem extrem straffen Zeitplan erfolgen. Da die Einführung der IKPM schon im Schuljahr 2021/2022 geplant ist, ist für die Vorbereitung weniger Zeit als für die Einführung der Bildungsstandards. Es muss daher die Frage erlaubt sein, ob der Zeitplan für die Umsetzung eines derart großen Projekts nicht etwas zu ambitioniert ist, da zusätzlich gleichzeitig die neuen Lehrpläne und die Kompetenzraster eingeführt werden. Es ist zu hoffen, dass sich das Feld noch stark bewegt und die derzeitige Entwicklungsarbeit noch gut in das Projekt einfließen kann.*
- *Lehrer*innen sind nicht geschult im Umgang mit Daten von Testungen bzw. Testergebnissen. Geplant ist zwar eine groß angelegte Lehrer*innenfortbildung, es bleibt dabei zu hoffen, dass alle betroffenen Lehrer*innen den Umgang mit diesen neuen Instrumenten auch rechtzeitig lernen. Zudem muss ja dann auch noch entsprechend unterrichtet werden, d.h. es ist zu hinterfragen, inwieweit a) die Lehrpläne und b) die didaktischen Zugänge der Lehrer*innen auf die Testungen vorbereiten.*
- *Die Frage nach dem P habe ich auch gestellt. Hier werden die Aussagen relativ diffus. Es sollen ja mit der IKPM nicht nur die Kompetenzen der Schüler*innen gemessen werden, sondern auch deren Potential. Darüber, wie das angestellt werden soll und welche Aussagen dann über acht- bis neunjährige Kinder getroffen werden, konnte mir bislang niemand wirklich beantworten. Noch dazu nehme ich an, dass Potentialmessungen regional und schulbezogen recht unterschiedlich ausfallen können. Zurzeit gehe ich davon aus, dass IKPM als Arbeitstitel zu sehen ist, vielleicht kann das P zumindest in der Volksschule ja noch verhindert werden.¹⁰*

¹⁰ Siehe dazu Fußnote 7.

Ich möchte an dieser Stelle eine Passage aus einem Zeitungsartikel von Prof. Karlheinz Gruber zitieren: „Wie fair und prognostisch wertvoll kann eine Beurteilung der Begabungen und Interessen von achteinhalb Jahre alten Kindern sein, bei denen Lehrerinnen gerade einmal zweieinhalb Jahre Zeit hatten, familiär bedingte Lerndefizite zu kompensieren und den Kindern über die Vermittlung von Kulturtechniken hinaus die vielfältige Welt des Lernens zu erschließen?“ [Gruber 2020]

- *Zu guter Letzt bin ich überzeugt davon, dass der Plan die Ergebnisse dieser Messungen nicht in die Leistungsbeurteilung der Schüler*innen einzubeziehen, nicht aufgehen wird.* Geplant ist, dass die Ergebnisse in die verpflichtend zu führenden sogenannten KEL-Gesprächen miteinfließen sollen. Die aus den Testungen erhaltene Leistungsinformation soll also als Grundlage für Elterngespräche dienen, wobei hier noch ein Leitfaden zur Objektivierung der Gesprächsführung entwickelt werden soll. Jetzt wird aber bei den offiziellen Testungen nur 45 Minuten geprüft, dabei können nicht alle Kompetenzen abgeprüft werden. Dazukommend stelle ich mir folgendes Szenario vor: Ein Kind hat bei der IK(P)M (iKM^{plus}) sehr gut abgeschnitten, der Lehrer*die Lehrerin ist aber der Meinung, dass dieses Kind nicht gut genug für ein Sehr gut, Gut oder dgl. ist. Diesen Umstand muss man dann bei einem Elternsprechtag einmal erklären. Sie können das Beispiel auch in umgekehrter Form durchdenken.

Anders herum gedacht: Wenn Sie sich nun an meine obigen Ausführungen zur fehlenden Objektivität von Noten erinnern, dann kommt bei dieser Vorgehensweise noch erschwerend hinzu, dass die Lehrer*innen ja genau wissen, wie ein Kind bei der offiziellen Testung abgeschnitten hat, und dass diese Erkenntnis nicht in die Notengebung einfließt, kann ich mir beim besten Willen nicht vorstellen.

Mein zweites Fazit lautet daher: Es wird gerade in viel zu kurzer Zeit ein Instrument entwickelt, das sowohl den Kompetenzstand von Schüler*innen als auch deren Entwicklungspotential messen soll. Begonnen wird mit der Messung bereits im Alter von ca. acht Jahren. Dabei sollen Rückmeldungen gleichzeitig sowohl auf Individual- als auch auf Systemebene generiert werden, was testtheoretisch hinterfragt werden muss. Die Ergebnisse sollen von den Lehrer*innen einerseits genutzt werden, damit sie ihre Schüler*innen sinnvoll fördern können und ihre Gymnasialreife konstatieren, andererseits müssen sie laut Verordnung die Ergebnisse ignorieren, wenn es um die Beurteilung von Leistungen geht. Das Ganze findet in einem Schulsystem statt, das von großer Chancenungleichheit geprägt ist und Kinder viel zu früh selektiert.

Was wäre also meiner Meinung nach ideal/besser?

- Die Entwicklung von Kompetenzrastern wird bereits vorangetrieben, das ist prinzipiell gut. Somit wird entweder für jede Schulstufe *oder viel besser am Ende einer neu zu definierenden Periode in der Laufbahn eines Schülers bzw. einer Schülerin* klar, was die Ziele sind, welche Kompetenzen mindestens erreicht werden müssen und welche individuellen Förderungsmaßnahmen die Schüler*innen benötigen, damit sie diese Ziele bestmöglich erreichen können. Festzuhalten ist, dass die neu entwickelten Kompetenzraster auch in der Praxis gut umsetzbar sein müssen – zurzeit ist dies noch nicht ganz der Fall, die Entwürfe dazu sind relativ umfangreich und natürlich an die geltende LBVO angepasst, also notenbasiert.

- Werden mit den Kompetenzrastern diese Forderungen eingehalten, dann würde es meines Erachtens letztendlich keine Noten mehr brauchen, dies bedingt aber eine Änderung der seit über 45 Jahren gültigen Leistungsbeurteilungsverordnung. Und wahrscheinlich vor allem ein Umdenken in den Köpfen aller, die mit Noten sozialisiert worden sind. Andere Formen der Leistungsfeststellung und -beurteilung wie z. B. direkte Leistungsvorlagen oder verbale Beurteilungen können an Stelle von Noten treten [vgl. Benischek 2006a: 177f].
- Zudem dürfen die neu entwickelten Instrumente nicht zu einer noch früheren Selektion von Schüler*innen in gymnasialreif und nicht-gymnasialreif führen. Wenn überhaupt, dann sollte Selektion nach Kriterien des Interesses und der Affinitäten möglichst spät stattfinden, frühestens das erste Mal mit 14, besser sogar noch mit 15 oder 16 Jahren, auf keinen Fall schon mit acht Jahren.
- Ein institutionalisierter, professioneller und entspannter Umgang mit Messinstrumenten und Unterstützungsmaterial könnte dazu führen, dass sowohl Schüler*innen als auch Lehrer*innen diese Instrumente im Sinne der Förderung von Kompetenzen ernst nehmen und die sich daraus ergebenden Möglichkeiten jeweils für sich nutzen. Vielleicht würde dann auch die ewige Ausrede, warum österreichische Schüler*innen bei BiSta-Testungen oder bei PISA schlecht abschneiden – weil sie den Test an sich aufgrund seiner fehlenden Notenrelevanz nicht ernst nehmen – wegfallen (es gäbe ja keine Noten mehr); die Ergebnisse könnten dann seriöser als bisher diskutiert werden.¹¹
- In der Praxis stehende Lehrer*innen müssen in einem engmaschigen Netz an Fortbildungen und Arbeitsmaterialien Unterstützung erhalten, um fachgerecht und professionell mit den Themen der Leistungsfeststellung und Leistungsbeurteilung umgehen zu können. Diesbezüglich soll vor allem die pädagogische Funktion von Leistungsbeurteilung hervorgehoben werden. Die Entwicklung einer positiven Feedbackkultur muss damit einhergehen. Dabei sollte auch gleich die Rolle der Lehrkraft als praktisch alleinige Selektionsinstanz hinterfragt werden.
- Durch die Arbeit mit Kompetenzrastern wird der Spielraum für Lehrer*innen, was die Notengebung betrifft, kleiner. Das ist in vielen Fällen kein Schaden, könnte es doch zu mehr Chancengerechtigkeit und Objektivität im Bildungsbereich führen. Die Konzentration auf die einzellehrer*innenbezogene Leistungsbeurteilung im Sinne des Selektionsgedankens sollte zurückgedrängt werden.

Soweit zu den Empfehlungen auf Schul- und Unterrichtsebene.

- Auf Systemebene bedarf es weiterhin verschiedener Monitoringinstrumente, die aber getrennt von der Individualebene ablaufen sollten. PISA, TIMSS, BiSta, VERA (so wie in Deutschland) – egal welche Bezeichnung dafür gewählt wird, ein Blick ins System, ein internationaler Vergleich kann – wenn auf Systemebene gut damit umgegangen wird – nur von Vorteil sein. Ziel dabei müsste sein, dass dies zu einer gelingenden Nutzung von datenbasierter Rückmeldung führt, im Gegensatz zu

¹¹ Siehe dazu auch den Artikel von Sabine Geist (Direktorin der Laborschule Bielefeld) zum Thema pädagogisches Diagnostizieren, das als partizipativer, ko-konstruktiver und dialogischer Prozess angelegt ist. [Geist 2019]

- unangemessenen Geltungsansprüchen evidenzbasierter Urteile.
- Getrennt von Lehrer*innen und Unterricht sollte auch im Rahmen einer vertiefenden Berufsorientierung Beratung der Schüler*innen stattfinden. Hier könnten mögliche Potentiale ermittelt werden. Dies aber zu gegebener Zeit in der Schullaufbahn.
 - Als letzten Punkt möchte ich den Blick auf die Ausbildung lenken. Lehramtsstudierende sollten sich früh genug mit der Problematik der Leistungsfeststellung und Leistungsbeurteilung auseinandersetzen, den Umgang mit Kompetenzrastern lernen und sich in diesem Bereich möglichst früh professionalisieren.

3. Problemaufriss II

Ein Teil meiner obigen Ausführungen sollte verdeutlichen, dass die von den Lehrer*innen vergebenen Leistungsbeurteilungen nicht immer mit den gemessenen Leistungen bei standardisierten Prüfungen übereinstimmen, dazu ist der Spielraum bei der Vergabe von (Jahres-) Noten einfach zu groß – nicht immer zum Nachteil der Schüler*innen, aber auch nicht immer zu deren Vorteil. Nicht zuletzt hat dies auch bei der standardisierten Reifeprüfung zu großen Unsicherheiten auf allen Seiten (Schüler*innen, Eltern und Lehrer*innen) geführt. Schüler*innen, die in Mathematik immer gute bzw. sehr gute Noten aufweisen konnten, wurden bei der SRDP plötzlich mit Nicht genügend beurteilt. Offiziell wurden die Daten der Jahresnoten leider noch nicht mit jenen der SRDP verknüpft.

Auch umgekehrt haben Schüler*innen mit befriedigenden oder genügenden Beurteilungen im Jahreszeugnis bei der SRDP sehr gut abgeschnitten. Abgesehen davon, dass es sich um unterschiedliche Formen der Leistungsfeststellung handelt (kontinuierliche vs. punktuelle), könnte es auch sein, dass Lehrer*innen zu wenig Expertise im Umgang mit den zur Verfügung gestellten Instrumenten haben. Ich möchte mich im Folgenden auf die Primarstufe beziehen, obwohl natürlich allgemein Untersuchungen zur Notengebungsfindung von Lehrer*innen auch im Sekundarstufenbereich von größtem Interesse wären.

Wo entsteht die Problematik vor allem in der Volksschule? Da Österreich als eines der wenigen EU- und OECD-Länder ein hochselektives Schulsystem am Ende der 4. Klasse Volksschule hat, ist es für viele Schüler*innen und auch für Eltern (vor allem in städtischen Ballungsräumen) von besonderer Bedeutung, dass die Noten am Ende der Volksschule einen Einstieg ins Gymnasium ermöglichen. Hier taucht vor allem das Gespenst der Hauptschule bzw. (neuen) Mittelschule als Restschule auf. Interessant sind hier Statistiken, die zeigen, dass neben regionalen Unterschieden (Stadt/Land) auch soziale Aspekte bei der Schulwahl (Hauptschule oder Gymnasium) eine große Rolle spielen. „So liegt der AHS-Schüleranteil in den Wiener Innenstadtbezirken (Innere Stadt, Josefstadt, Döbling) bei über 70%. In den Außenbezirken (Favoriten, Simmering, Meidling, Ottakring) aber unter 40 %“ [Schmid 2004: 15]. Die Richtung, die Kinder einschlagen, AHS-Matura und Studium oder eben nicht, wird schon am Ende der Volksschule durch verschiedenste Faktoren quasi vorherbestimmt (auch wenn das Schulsystem von offizieller Seite immer als sehr durchlässig beschrieben wird). Dies erzeugt einen enormen Druck auf die Lehrer*innen in der Volksschule. Umso genauer müssen sie darlegen, wie die Beurteilung von Leistungen der Schüler*innen zustande kommt.

Befragt man Berufseinsteiger*innen, dann werden als anfängliche Herausforderungen v.a. Classroom Management und das Problem der Bewertung von Leistungen genannt. Erste Ergebnisse einer Interviewstudie mit Masterstudierenden an der KPH Wien/Krems deuten darauf hin, dass

- die interviewten Berufseinsteiger*innen vor allem traditionelle Elemente der Leistungsfeststellung (Kriterienkataloge, Lernzielkontrollen, Schulübungen, Hausübungen, Mitarbeit, ...) einsetzen. Das Spektrum variiert dabei durchaus.
- Wenn es um Leistungsbeurteilung geht, dann übernehmen junge Lehrer*innen oftmals die bereits bestehenden Konzepte erfahrenerer Kolleg*innen. Sie verrechnen einzelne Leistungen summativ (meist ohne die wesentlichen Bereiche an die Schüler*innen zu

kommunizieren – und daher nicht LBVO-konform), messen Mitarbeit, Hausübungen und dgl. bestimmte Prozente zu und

- wenden eine Mischung aus den verschiedenen Ebenen der Bezugsnormen (kriterial, sachlich und individuell) an.¹²

Die Tradierung dieser Systeme ist aber oft – unter Beachtung meiner Ausführungen aus Teil 1 – relativ problematisch.

- Die Notendefinition wird nur teilweise und dabei recht unterschiedlich angewendet. Manche sehen Schüler*innen sehr gut, „wenn alles passt“, andere gehen vom Befriedigend aus und stellen diese Note als sozusagen „unteres Niveau“ dar.
- Was das Anspruchsniveau betrifft, so spielen eigene Erfahrungen eine Rolle und auch die Schule (z. B. Parallelklassen usw.) wird als Norm herangezogen. Wie vermutet, besteht keine einheitliche Herangehensweise an die Bestimmung des Anspruchsniveaus, es erscheint relativ willkürlich, wenn auch in den Einzelfällen nicht unüberlegt.

Was das Studium betrifft, so sind die Antworten doch sehr erstaunlich:

- Die Befragten geben an, dass sie vom Studium zur Thematik der Leistungsbeurteilung relativ wenig mitgenommen haben. Sie haben zwar gelernt Schularbeiten zusammenzustellen bzw. zu beurteilen, ein differenziertes Konzept der Beurteilung von Leistungen fehlt ihnen aber völlig. Der Wunsch nach einer intensiveren Beschäftigung mit dem Thema Leistungsbeurteilung während des Studiums wurde von allen Befragten formuliert.

Dabei haben diese jungen Kolleg*innen ein achtsemestriges Studium abgeschlossen und müssten daher bestens für den Berufseinstieg gerüstet sein. Geht man also der Frage „Was müssen Lehrer*innen können, um Leistungen von Schüler*innen fair zu bewerten?“ nach, so liegt der Grundstein sicherlich in der Ausbildung.

Daher habe ich einen näheren Blick in das für das BA-Studium des Lehramts Primarstufe gültige Curriculum geworfen, und es haben sich meines Erachtens ganz interessante Erkenntnisse ergeben.

Vorauszuschicken ist, dass der erfolgreiche Abschluss des BA- bzw. MA-Studiums „Grundlage der Befähigung zur Ausübung des Lehramts im Bereich der Primarstufe“ [KPH 2017: 4] ist. Und zur Ausübung des Berufs gehört natürlich auch das Feststellen und Bewerten von Leistungen.

Ich bin nun ganz pragmatisch-mathematisch an die Sache herangegangen und habe gezählt, wie oft das Wort Leistung im 672-seitigen Curriculum des BA-Studiums Lehramt Primarstufe vorkommt. Insgesamt 684 Mal allein oder in einer Wortkombination. Hier muss eine Trennung in die Kategorien inhaltlich und organisatorisch vorgenommen werden, also wo

¹² Diese Ergebnisse decken sich mit der von Friedrich & Heim durchgeführten Studie mit Junglehrer*innen [vgl. Friedrich & Heim 2019].

geht es um Studieninhalte für Studierende und welche Ausdrücke beschreiben Merkmale auf einer organisatorischen Ebene wie z. B. den Leistungsnachweis für eine bestimmte Lehrveranstaltung oder dgl. .

Zieht man die organisatorischen ab, dann bleiben immerhin noch 318 inhaltliche Beiträge mit dem Wortteil -leistung- übrig. Z. B.

Leistungsgesellschaft, Leistungskultur, Begabung und Leistung,
Leistungsbeurteilung, Leistungserwartungen, Leistungs- und Rückmeldekultur,
Leistungsverständnis, Leistungsfeststellung, Leistungsbegriff, Leistungsbewertung,
Lernleistung, Leistungserziehung, Leistungszuversicht, Leistungsfähigkeit,
Leistungsethos, Leistungsvoraussetzungen, Leistungsbereitschaft,
Schulleistungstests, Leistung, leistungsstark, leistungsschwach, Evaluation von
Lehrerleistung, Leistungsunterschiede.¹³

Laut Studienplan ist die Thematik der Leistungsfeststellung und -beurteilung auf verschiedenen Ebenen abgedeckt.

Ebene 1: Im Qualifikationsprofil des Curriculums sind im Abschnitt „Erwartete Lernergebnisse/Kompetenzen“ verschiedene Grundkompetenzen angeführt. Der Kompetenzbereich „Diagnostizieren, Beurteilen und Beraten“ formuliert sieben Kompetenzen auf unterschiedlichen Anforderungsniveaus (Stufen I-V). Jene drei für die Beurteilung von Leistung relevanten sind:

- D2: Lehrerinnen und Lehrer können ihr Grundwissen für pädagogische Diagnostik, Beurteilung und Beratung nutzen. (III)
- D5: Lehrerinnen und Lehrer fördern Schülerinnen und Schüler auf der Grundlage eines pädagogischen Lern- und Leistungsverständnisses und beraten Lernende und deren Eltern gezielt. (IV)
- D6: Lehrerinnen und Lehrer nutzen die Vielfalt an Formen und Möglichkeiten für die Leistungsfeststellung. Sie beurteilen Lernprozesse und Lernergebnisse unter Berücksichtigung der geltenden Rechtslage auf der Grundlage transparenter Beurteilungsmaßstäbe. (IV) [KPH 2017: 19]

Diese Grundkompetenzen finden sich in den Beschreibungen der einzelnen Lehrveranstaltungen und Module wieder. Sie werden dort als Querschnittsmaterie abgebildet. Interessant ist auch hier die Häufigkeit der Verteilungen:

Grob sind sie in zwei Kategorien einzuteilen: Diagnose (D1, D3, D4, D7) und Beurteilung (D2, D5, D6), insgesamt werden die inhaltlich/thematisch relevanten Ds 198 Mal im Curriculum angeführt und das bei insgesamt 546 unterschiedlichen Lehrveranstaltungsbeschreibungen.

¹³ In diesem Zusammenhang habe ich mir auch gleich die Häufigkeit der Begriffe Diagnostizieren (11 mal), Beurteilen (46 mal), Beraten, Bewerten (je 16 mal) und Fördern (44 mal) angesehen, ohne aber eine Trennung in organisatorisch und inhaltlich vorzunehmen.

Ich habe hier auch noch eine detailliertere Analyse durchgeführt, in welchen Lehrveranstaltungen sie vorkommen, das würde hier aber zu weit führen.

Ebene 2: Lehrveranstaltungsebene – Überschriften

Hier muss man zwischen den verpflichtenden Lehrveranstaltungen, die alle Studierenden durchlaufen, und den Wahlangeboten trennen. Verpflichtend haben alle Studierenden die Module Begabung und Leistung (4-02) und Pädagogisch-praktische Studien: Diagnostizieren und Fördern, Beurteilen und Beraten (5-02) – also im 4. und im 5. Semester – zu absolvieren. Als Wahlangebote stehen ihnen zusätzlich noch in vier Modulen Vorlesungen, Seminare oder Übungen zur Verfügung, die sich mit dem Thema „Leisten“ beschäftigen. Diese werden von den Studierenden aber eben nur zufällig belegt.

Es gibt sicherlich noch verschiedene andere Lehrveranstaltungen, in denen das Thema „Leistungsfeststellung und -beurteilung“ behandelt wird, wo dies aber aus dem Titel der Lehrveranstaltung nicht hervorgeht. Daher habe ich mir auch noch eine dritte Ebene angesehen.

Ebene 3: Modulbeschreibungen auf Ebene der Bildungsziele und Bildungsinhalte sowie der Lernergebnisse/Kompetenzen

Auch hier wurde zwischen Pflicht- und Wahlbereichen unterschieden. Verpflichtend müssen alle Studierenden die Module Fachdidaktisches und methodisches Handeln in Musik sowie in Bewegung und Sport 1, Instrumentalmusik (3-06), Begabung und Leistung (4-02) und Fachdidaktisches und methodisches Handeln in Kunst und Werken 1, Instrumentalmusik (4-06) absolvieren. Im Wahlbereich gibt es noch neun Module, die auf der Ebene der Modulbeschreibungen im Studienplan Ansätze im Bereich Leistung beinhalten. Spätestens auf dieser Stufe der Analyse wird aber sichtbar, dass oft Diagnose und Förderung inhaltliche Schwerpunkte zu sein scheinen, genauso wie Individualisierung und Differenzierung. Der Begriff der Kompetenzorientierung kommt relativ oft vor, genauso wie die kriteriengeleitete Planung. Großteils bleiben aber die inhaltlichen Formulierungen auf dem Gebiet der Diagnose stehen. Der Beurteilung als nächstem Schritt wird hierbei womöglich (noch) zu wenig Bedeutung zugemessen, was auch mit den Rückmeldungen der befragten BA-Absolventinnen übereinstimmt.

Was anfänglich beim reinen Abzählen der relevanten Begriffe den Eindruck erweckt hat, dass das Thema einen sehr großen Stellenwert im Studienplan hat, erscheint beim genaueren Hinschauen nun nicht mehr so¹⁴. Letztendlich scheint also das Curriculum relativ wenige Möglichkeiten anzubieten, bei denen sich die Studierenden intensiver mit dem Thema Leistungsfeststellung und -beurteilung auseinandersetzen können (je nachdem welche Wahlmodule bzw. Schwerpunkte sie gewählt haben).

¹⁴ Eine Ausnahme bilden die Lehrveranstaltungen, die sich mit der Erstellung und Beurteilung von Schularbeiten in D und M befassen, die auch von den Befragten in den Interviews als in Erinnerung erwähnt wurden.

Fazit von Problemaufriss II

- Oft existiert eine Diskrepanz zwischen von Lehrer*innen vergebenen Noten und Testungen von außen.
- Gründe dafür könnten sein, dass Lehrer*innen ihre Noten sozusagen freifliegend „nach Gefühl“ geben (es gibt ja keine Benchmarks für bestimmte Noten) und kein gemeinsames Verständnis über die Umsetzung der LBVO herrscht. Es werden Notenvergabepraktiken von Kolleg*innen übernommen oder die Ansprüche an den jeweiligen Schulstandort angepasst.
- Abhilfe sollen in Zukunft Kompetenzraster bringen, wo relativ genau definiert ist, welche Kompetenzen erreicht werden müssen, um eine bestimmte Note zu erlangen. Idealer wäre noch der Ansatz, der in Norwegen vorherrscht, dass Lehrer*innen mit Schüler*innen Kriterien der Leistungsbeurteilung vereinbaren und ein gemeinsames Förderkonzept entwickeln.
- Der Problematik von Leistungsfeststellung und Leistungsbeurteilung muss im Studium (ich vermute nicht nur hier an der KPH, sondern auch an anderen Pädagogischen Hochschulen und Universitäten) ein größerer Schwerpunkt eingeräumt werden. Studierende müssen lernen mit Kompetenzrastern umzugehen und sie müssen sich auch darüber bewusst sein, dass der Umgang mit Leistungsbeurteilung in Österreich ein hochselektiver und sozial extrem ungerechter ist. Nur mit diesem Bewusstsein kann es möglich werden, dass alle Schüler*innen ihren Bedürfnissen entsprechend gefördert werden, da diese Art von Selektion pädagogisch nicht legitimierbar ist.
- Auf sozusagen gesellschaftlicher Ebene muss das Anspruchsniveau der anstehenden IK(P)M/iKM^{Plus} ausverhandelt werden – sonst werden dieselben Probleme wie bei der SRDP auf die Beteiligten zukommen, nämlich dass unterschiedliche Ergebnisse in standardisierten Testungen verglichen mit den Bewertungen aus den Kompetenzrastern zu Unverständnis bei den Beurteilten und deren Eltern führt.

Hier wird es also in Zukunft wichtig sein, die Entwicklungen genau zu beobachten und auf die eben beschriebenen Problemfelder aufmerksam zu machen.

Die eben genannten Themen bilden nur einen kleinen Ausschnitt der Forschungsanliegen zum Thema Leistungsbeurteilung. Daher möchte ich am Schluss meines Vortrags noch skizzieren, mit welchen Forschungsschwerpunkten ich mich in nächster Zeit beschäftigen möchte.

4. Zukünftige Forschungsvorhaben

Analyse

- Analyse des MA-Studienplanes nach denselben Kriterien wie im vorigen Abschnitt
- Genauere Analyse der Lehrveranstaltungskonzepte zum Thema Leistungsbeurteilung im inhaltlichen Sinn
 - o Hier ist es z. B. wichtig mit den Lehrveranstaltungsleitenden des Bereichs Angewandte Förderdiagnostik in Kontakt zu treten, um herauszufinden, welche Inhalte zum Thema Leistungsbeurteilung und Leistungsfeststellung in dieser Lehrveranstaltung bearbeitet werden. Zudem gibt es an der KPH um die 400 Praxislehrer*innen in den Pädagogisch-praktischen Studien. Eine Analyse der vorhandenen Unterlagen (es gibt Kompetenzblätter zur Beurteilung der Studierenden für jedes Semester, die aber nicht mehr alle im Einsatz sind) ist hier sicher auch interessant.
 - o Darüber hinaus könnte man als größere Untersuchung die Praxislehrer*innen zu ihrer Beurteilungspraxis befragen. „Welche Formen der Leistungsfeststellung verwenden Sie? Wie kommen Sie von den einzelnen Leistungsfeststellungen zu einer bestimmten Note? Wie haben Sie das Anspruchsniveau für bestimmte Noten festgelegt? Usw.
Die Auswertungen dieser Untersuchung könnten dann in weiterer Folge auch für die Hochschule interessant sein.
- Damit einhergehend kann auch eine größere Lehrendenbefragung durchgeführt werden, um so zu einem Gesamtkonzept bzw. in einem ersten Schritt zu einer Gesamtschau – auch im Hinblick auf Problemaufriss I – zu gelangen.

Transfer Studium – Praxis, Professionalisierung der pädagogischen Diagnostik und Bewertung

- Ausgehend von den Ergebnissen meiner Interviews (Vorstudie) sollen weiterführende Untersuchungen folgen, etwa in Form einer flächendeckenden Studierendenbefragung mittels Fragebogen
 - o Was haben Sie zum Thema Leistungsbeurteilung aus Ihrem Studium mitgenommen?
 - o Wie können Sie diese Inhalte in Ihrer täglichen Praxis umsetzen?
 - o Nach welchen Kriterien konzipieren Sie Leistungsbeurteilung in Ihrer täglichen Praxis?
 - o Auf welche Aspekte der Vielfalt achten Lehrer*innen bei der Beurteilung von Leistungen (z. B. auch im Hinblick auf mehrsprachige Schüler*innen)?
 - o Wo würden Sie noch Unterstützung brauchen?

Durchaus vorstellbar ist auch eine Ausweitung auf andere Hochschulen, etwa eine großflächigere Erhebung von Daten in diesem Bereich. „Wie kommen Lehrer*innen in österreichischen Volksschulen generell zu ihren Noten?“ An der KPH Wien/Krems könnte eine Sensibilisierung der Lehrenden stattfinden, sofern dies gewünscht ist.

- o Diskussionen zum Leistungsbegriff an sich
 - o Entwicklung eines gemeinsamen Verständnisses der Problematik der Leistungsbeurteilung
 - o Es könnten auch weitere Schwerpunkte gesetzt werden wie z. B. durch Vorstellung von relevanten Studien zum Selbstkonzept und zur Selbstwirksamkeit von Schüler*innen und deren Auswirkungen auf ihre Leistungen.
 - o Sensibilisierung hinsichtlich der kulturellen Validität von Leistungsbeurteilung – Stichwort kultursensible Leistungsbeurteilung [vgl. Altrichter, Herzog-Punzenberger & Fellner 2019]
 - o Entwicklung eines gemeinsamen Konzepts zur Leistungsbeurteilung in den jeweiligen Lehrveranstaltungen. Welche Qualifizierungen benötigen Lehrer*innen dafür?
- Auf Lehrveranstaltungsebene wäre es durchaus interessant auch im Hochschulbereich sozusagen alternative Formen der Leistungsbeurteilung – sofern sie natürlich den gesetzlichen Rahmenbedingungen entsprechen – auszuprobieren. Ich würde gerne bei der Beurteilung meiner Studierenden mehr Fokus auf die Kompetenzorientierung legen, um in einem Aushandlungsprozess gemeinsam mit den Studierenden Kriterien der Leistungsbeurteilung festzulegen, die ihnen dann als Vorbild für ihre eigene Unterrichtstätigkeit dienen können.
 - Und natürlich spielt auch das Konzept des Geschlechts bei diesem Thema eine große Rolle. Hier wird es in Zukunft vor allem darum gehen, dass genaue Analysen bzgl. der Benachteiligung von Mädchen, genauso wie von Burschen, im Rahmen von durchgeführten Leistungsbeurteilungen gemacht werden. Stichwort: Zugangsweise zu Testungen; der Bogen zur SRDP könnte hier gespannt werden.

Zu den in Problemaufriss I dargestellten Überlegungen könnten auch noch weitere Untersuchungen angestellt werden bzw. sollte man die diesbezüglichen Entwicklungen genau beobachten.

5. Ausblick

Es gibt also noch einiges zu tun und ich freue mich darauf. Schließen möchte ich meine Ausführungen mit einer Aussage meines Sohnes Leo in folgender Situation:

Leo hat nach Beendigung seines ersten Schuljahres in der Volksschule neben der Leistungs- und Fortschrittsinformation ein Pensenbuch erhalten mit einer Liste von Kompetenzen, die er im ersten Schuljahr erreicht hat. Ich habe mir das durchgesehen und zu ihm gesagt, dass er ziemlich stolz auf sich sein kann wie viel er im ersten Jahr schon gelernt hat. Seine Antwort darauf war: Ja Mama, aber ganz ehrlich, völlig unabsichtlich!!!

Zuletzt bleibt mir noch die Gelegenheit mich bei all jenen zu bedanken, die mich auf meinem beruflichen Lebensweg unterstützt haben. Sowohl Kolleginnen und Kollegen als auch Freundinnen und Freunden sei hier gedankt – ich weiß, es war und ist nicht immer leicht mit mir. Vor allem aber, möchte ich auch meiner Familie und besonders meinen Eltern danken, die all meine Entscheidungen immer unterstützend hingenommen haben. Ich danke euch dafür, dass ihr trotz aller, nicht immer idealer Voraussetzungen, stets Wert auf eine gute Ausbildung eurer Kinder gelegt habt. Zu guter Letzt gilt mein Dank auch Christoph, für alles! Und das mit großer Geduld!

6. Literatur

- Altrichter, Herbert, Herzog-Punzenberger, Barbara & Fellner, Magdalena (2019). *Leistungsbeurteilung angesichts sprachlicher und kultureller Diversität*. Vortrag am ÖFEB-Kongress 2019. Vermessen? Zum Verhältnis von Bildungsforschung, Bildungspolitik und Bildungspraxis. Linz, PH OÖ: 17.-20.9.2019.
- Beba (2016). *Mathe-Matura: „Wer das nicht versteht, sollte nicht studieren“*. Die Presse am 23.5.2016. Online verfügbar unter: <https://www.diepresse.com/4994383/mathe-maturawer-das-nicht-versteht-sollte-nicht-studieren>, Zugriff am 20.8.2020.
- Benischek, Isabella (2006a). *Die Leistung von SchülerInnen – ein mehrdimensionaler Begriff*. In Beranek, Wilhelm & Weidinger, Walter (Hrsg.). *Erziehung und Unterricht*, Heft 1-2/2006. S. 171-180. Wien: öbvht.
- Benischek, Isabella (2006b). *Aspekte der Lehrerurteile an der Nahtstelle Primarstufe – Sekundarstufe 1*. In Beranek, Wilhelm & Weidinger, Walter (Hrsg.). *Erziehung und Unterricht*, Heft 3-4/2006. S. 295-301. Wien: öbvht.
- BIFIE (2013). *Standardisierte kompetenzorientierte Reifeprüfung/Reife- und Diplomprüfung. Grundlagen – Entwicklung – Implementierung*. Online verfügbar unter: <https://www.matura.gv.at/index.php?eID=dumpFile&t=f&f=1786&token=6d36c3b29ea588d46f540972351fed64cc783ae8>, Zugriff am 27.8.2020.
- BIFIE (2019). *Bundesergebnisbericht Standardüberprüfung 2018. Mathematik 4. Schulstufe*. Online verfügbar unter: https://www.bifie.at/wp-content/uploads/2019/02/BiSt_UE_M4_2018_Bundesergebnisbericht.pdf, Zugriff am 8.12.2019.
- BIFIE (2020). *Bildungsstandards und kompetenzorientierter Unterricht*. Online verfügbar unter: <https://www.bifie.at/bildungsstandards-und-kompetenzorientierter-unterricht/>, Zugriff am 24.10.2019.
- Bundeskanzleramt (2020). *Aus Verantwortung für Österreich, Regierungsprogramm 2020-2024*. Online verfügbar unter: <https://www.bundeskanzleramt.gv.at/bundeskanzleramt/die-bundesregierung/regierungsdokumente.html>, Zugriff am 14.1.2020.
- BMBWF (2019). *Bildungsstandards in der Allgemeinbildung*. Website des BMBWF, Zugriff am 18.11.2019. Nicht mehr verfügbar.
- BMBWF (2020). *Bildungsstandards in der Allgemeinbildung*. Online verfügbar unter: <https://www.bmbwf.gv.at/Themen/schule/bef/bist/bildungsstandards.html>, Zugriff am 28.8.2020.
- Bruneforth, Michael, Weber, Christoph & Bacher, Johann (2012). *Chancengleichheit und garantiertes Bildungsminimum in Österreich*. In Herzog-Punzenberger, Barbara (Hrsg.). *Nationaler Bildungsbericht 2012*, Band 2, S. 189-228. Online verfügbar unter: https://www.bifie.at/wp-content/uploads/2017/05/NBB2012_Band2_gesamt_20121217.pdf, Zugriff am 8.12.2019.
- Eder, Ferdinand, Neuweg, Georg Hans & Thonhauser, Josef (2009). *Leistungsfeststellung und Leistungsbeurteilung*. In Specht, Werner (Hrsg.). *Nationaler Bildungsbericht 2009*, Band 2, S. 247-267. Online verfügbar unter: https://www.bifie.at/wp-content/uploads/2017/05/2009-06-16_NBB-Band2.pdf, Zugriff am 22.5.2020.
- Friedrich, Nadja & Heim, Helena (2019). *Über die Gratwanderung von JunglehrerInnen beim Bewerten*. In Braun, Helga & Beer, Rudolf (Hrsg.). *Erziehung und Unterricht*, Heft 9-10/2019. S. 881-891. Wien: öbv.

- Geist, Sabine (2019). *Pädagogisches Diagnostizieren – alltäglicher Bestandteil schulischer Arbeit*. In Braun, Helga & Beer, Rudolf (Hrsg.). *Erziehung und Unterricht*, Heft 9-10/2019. S. 918-926. Wien: öbv.
- Gruber, Karl Heinz (2020). *Bildungspolitische Glühwürmchen*. Der Standard am 11.1.2020. Online verfügbar unter: <https://www.derstandard.de/story/2000113159783/bildungspolitische-gluehwuermchen>, Zugriff am 14.1.2020.
- Hämmerle, Klaus (2016). „So eine Matura gibt man nicht“. *Vorarlberger Nachrichten* am 31.5.2016. Online verfügbar unter: <https://epaper.vn.at/lokal/vorarlberg/2016/05/30/so-eine-matura-gibt-man-nicht.vn>, Zugriff am 20.12.2019.
- Kirchliche Pädagogische Hochschule Wien/Krems (2017). *Curriculum Bachelorstudium als Voraussetzung für ein Masterstudium zur Erlangung eines Lehramtes im Bereich der Primarstufe*. Nicht mehr verfügbar.
- Kronen Zeitung (2019). *Schüler frustriert. Mathe-Matura bekommt ein „Nicht Genügend“*. Kronen Zeitung am 13.5.2019. Online verfügbar unter: <https://www.krone.at/1920858>, Zugriff am 20.12.2019.
- LBVO (2020). *Rechtsvorschrift für Leistungsbeurteilungsverordnung, Fassung vom 22.12.2016 (RIS), §1 Abs. 2*. Online verfügbar unter: <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10009375&FassungVom=2016-12-22>, Zugriff am 11.5.2020.
- Neuweg, Georg Hans (2019). *Kompetenzorientierte Leistungsbeurteilung*. Linz: Trauner Verlag.
- Sacher, Werner (2014). *Leistungen entwickeln, überprüfen und beurteilen*. Bad Heilbrunn: Klinkhardt.
- Schmid, Christine, Paasch, Daniel & Katstaller, Michaela (2016). *Kompositionseffekte bei der Notenvergabe auf der 4. Schulstufe der österreichischen Volksschule*. *Zeitschrift für Bildungsforschung* 6 (2016) 3. S. 265-283.
- Schmid, Kurt (2004). *Regionale Bildungsströme in Österreich*. *ibw-Reihe Bildung und Wirtschaft*, Nr. 31. Online verfügbar unter: <http://www.forschungsnetzwerk.at/downloadpub/BW31.pdf>, Zugriff am 20.8.2020.
- Schreiner, Claudia & Kraler, Christian (2019). *Pädagogische Diagnostik und LehrerInnenbildung nach PISA*. In Braun, Helga & Beer, Rudolf (Hrsg.). *Erziehung und Unterricht*, Heft 9-10/2019. S. 861-871. Wien: öbv.
- Schreiner, Claudia & Wiesner, Christian (2017). *Genese der Bildungsstandards in Österreich*. *BIFIE-Journal* 1/2017. DOI: 10.7888/bifiejournal-1.2017-1-5

Akademischer Werdegang von HS-Prof.ⁱⁿ Mag.^a Dr.ⁱⁿ Eva Sattlberger

Eva Sattlberger wurde in Oberösterreich geboren, wo sie die Volksschule und das Gymnasium besuchte. 1995 schloss sie ein Lehramtsstudium für Mathematik und Physik ab (Diplomarbeit *Konstruktionsaufgaben*). Nach einem einjährigen Aufenthalt in San Diego (Kalifornien) begann sie 1996 in Wien als Lehrerin zu unterrichten. Ihre Dissertation zum Thema *College Preparatory Mathematics versus Offenes Lernen. Ein Vergleich zweier Mathematikprogramme hinsichtlich ihrer Eignung zur Vermittlung von Schlüsselqualifikationen basierend auf einer in Kalifornien durchgeführten Fragebogenuntersuchung* schloss sie 2001 ab. Bis 2003 absolvierte sie den ECHA-Ausbildungslehrgang Educating the Gifted an der Universität von Nijmegen.

Ihr beruflicher Weg führte sie vom Stadtschulrat für Wien (pädagogische Referentin) über die Universität Wien (wissenschaftliche Mitarbeiterin am Institut für Bildungswissenschaft/Lehrer*innenbildung) ans City College of New York (Juniorgastprofessur) und mit Lehraufträgen an verschiedene andere österreichische Universitäten. In der Zeit von 2012 bis 2019 leitete sie den Fachbereich Mathematik zur Standardisierten Reifeprüfung an AHS (zuerst BIFIE, dann Bundesministerium für Bildung, Wissenschaft und Forschung).

Seit 2004 ist sie Mitautorin der Lehrbuchreihen Sexl et al. *Physik* für die AHS-Oberstufe sowie Hanisch et al. *MatheFit* für die Sekundarstufe 1, seit 2019 Mitherausgeberin der Zeitschrift *SchulVerwaltung aktuell*, einer Fachzeitschrift für Schulentwicklung und Schulmanagement.

Ihre Tätigkeiten an verschiedenen Institutionen sind in zahlreichen Publikationen und Vorträgen dokumentiert und zeigen einen gelungenen Brückenschlag zwischen wissenschaftlicher Erkenntnis und schulpraktischer Erfahrung, den sie durch ihre fachdidaktische und bildungswissenschaftliche Expertise immer weiter ausbauen konnte.

Seit Februar 2019 ist Eva Sattlberger Hochschulprofessorin ph 1 für Schulpädagogik und Allgemeine Didaktik an der Kirchlichen Pädagogischen Hochschule Wien/Krems und ist in diesem Rahmen in Forschung und Lehre tätig. Ihre Forschungsschwerpunkte liegen in den Bereichen Leistungsbeurteilung, Genderforschung, Schulentwicklung und Lehrer*innen- und -fortbildung.

Ausgewählte Publikationen von HS-Prof.ⁱⁿ Mag.^a Dr.ⁱⁿ Eva Sattlberger

- Sattlberger, Eva (2006). *Kreativität im Mathematikunterricht oder Ein Plädoyer für Schokolade und Wasser*. In Schriftenreihe zur Didaktik der Mathematik der Österreichischen Mathematischen Gesellschaft (ÖMG). Heft 38. S. 129-139.
- Christof, Eveline & Sattlberger, Eva (2009). *Fortbildung an der Schnittstelle zwischen Theorie und Praxis. Der neu gestaltete Universitätskurs „Betreuungslehrer/in für die schulpraktische Ausbildung“ an der Universität Wien*. In Schrittmesser, Ilse (Hrsg.). Professionalität und Professionalisierung. Einige aktuelle Fragen und Ansätze der universitären LehrerInnenbildung. S. 79-94. Frankfurt am Main: Peter Lang.
- Sattlberger, Eva (2009). *Anmerkungen zur Qualität von Testungen im Bildungswesen*. In Bildungsqualität – Eine verdächtig selbstverständliche Forderung. Schriftenreihe Schulheft 136, 34. Jahrgang. S. 65-70. Innsbruck: Studienverlag.
- Aue, Vera, Frebort, Martina, Hohenwarter, Markus, Liebscher, Marlies, Sattlberger, Eva, Schirmer, Ingrid, Siller, Hans-Stefan, Vormayr, Günther, Weiß, Martina & Willau, Elisabeth (2013 und 2015). *Das Projekt „Standardisierte schriftliche Reifeprüfung in Mathematik“*. Online verfügbar unter: <https://www.matura.gv.at/index.php?eID=dumpFile&t=f&f=861&token=96731e3efb9cea71397b3063054a604396f213a7>.
- Sattlberger, Eva (2015). *Die standardisierte schriftliche Reifeprüfung – eine Systemänderung mit Konsequenzen*. In Das Ende der Schule, so wie wir sie kennen. Schriftenreihe Schulheft 160, 40. Jahrgang. S. 135-147. Innsbruck: Studienverlag.
- Siller, Hans-Stefan, Bruder, Regina, Hascher, Tina, Linnemann, Torsten, Steinfeld, Jan & Sattlberger, Eva (2016). *Kompetenzstufenmodell zu Reifeprüfungsaufgaben und deren Eignung für einen kompetenzorientierten Mathematikunterricht*. In Keller, Stefan, Reintjes, Christian (Hrsg.). Aufgaben als Schlüssel zur Kompetenz – Didaktische Herausforderung, wissenschaftliche Zugänge und empirische Befunde. S. 371-388. Münster: Waxmann.
- Siller, Hans-Stefan, Bruder, Regina, Linnemann, Torsten, Hascher, Tina, Steinfeld, Jan & Sattlberger, Eva (2016). *Competency level modeling for school leaving examination*. In Krainer, Konrad & Vondrova, Nikola (Hrsg.). CERME 9 – Ninth Congress of the European Society in Research in Mathematics Education. Feb 2015, Prague, Czech Republic. S. 2716-2723. Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education. Online verfügbar unter: <https://hal.archives-ouvertes.fr/hal-01289473>. Münster: WTM.
- Bartok, Larissa, Gewessler, Philipp, Steinfeld, Jan, Ponocny, Iva, Ponocny-Seliger, Elisabeth, Sattlberger, Eva & Themessl-Huber, Michael (2017). *Welchen Einfluss haben Persönlichkeitseigenschaften? Eine Analyse im Rahmen der Österreichischen Matura*. In Tagungsband für die 51. Tagung der Gesellschaft für Didaktik der Mathematik. WTM: Münster.
- Sattlberger, Eva, Steinfeld, Jan & Gewessler, Philipp (2019). *Geschlechtsspezifische Unterschiede in Mathematikleistungen: Welchen Einfluss haben Persönlichkeitseigenschaften auf die Lösungswahrscheinlichkeit von Maturaaufgaben?* In Schriftenreihe zur Didaktik der Mathematik der Österreichischen Mathematischen Gesellschaft (ÖMG), Heft 51. S. 67-78.



HS-Prof. in Mag. a Dr. in Eva Sattlberger

Leistungsbeurteilung ist sowohl im wissenschaftlichen Diskurs als auch in der schulischen Praxis ein vieldiskutiertes, sehr umfangreiches und in seinen Facetten insgesamt sehr komplexes Thema.

In der Antrittsvorlesung präsentiert Eva Sattlberger Schlüsse aus den ersten Durchgängen der standardisierten Reife- und Diplomprüfung in Mathematik hinsichtlich des gesellschaftlichen Anspruchs und setzt diese in Bezug zu aktuellen Entwicklungen im Bereich von Kompetenzfeststellungen im Schulwesen.

*Dabei definiert sie Ansatzpunkte für eine förderliche Beurteilung von Schüler*innenleistungen und skizziert die dafür nötigen Rahmenbedingungen. Inwiefern diese Thematik in der aktuellen Ausbildung von Studierenden des BA-Studiums Lehramt Primarstufe an der KPH Wien/Krems verankert ist, zeigt eine Analyse des Curriculums.*

Den Ausblick bilden geplante Forschungsvorhaben der Vortragenden und mögliche Entwicklungsfelder zur Professionalisierung von Studierenden und Lehrenden.

ISBN 978-3-904046-03-9



9 783904 046039