
Themenheft Nr. 45:

Pädagogisches Wissen im Lichte digitaler und datengestützter Selbstoptimierung

Herausgegeben von Estella Ferraro, Sabrina Schröder und Christiane Thompson

Vom pädagogischen Misstrauen gegen die Prüfung und dem Versprechen einer digitalen Problemlösung

Sabrina Schröder¹ 

¹ Universität Kassel

Zusammenfassung

Dieser Artikel lotet verschiedene Spielarten eines pädagogischen Misstrauens gegen die Prüfung aus. Hierfür werden zunächst schulpädagogische Auseinandersetzungen v. a. zu den Problematiken der Leistungsbeurteilung diskutiert. Dafür wird erstens auf die Selektionsfunktion der Schule Bezug genommen, zweitens die Differenz zwischen Standardisierung und Individualisierung der Leistungsmessung betrachtet und drittens das Problem der prognostischen Möglichkeiten von Noten in Bezug auf künftigen Lernerfolg thematisiert. Vor dem Hintergrund dieser drei Argumentationslinien verdichtet sich das Misstrauen gegen die Prüfung als ein unauflösbarer Streit, nicht nur in der Wahl um die angemessenen Instrumente der Leistungsbeurteilung, sondern auch in der Frage danach, wie pädagogisch angemessen mit der diagnostischen Aufgabe umzugehen ist. Das neue Forschungsfeld der Learning Analytics wird anschliessend als eine mögliche Antwort auf diese Frage thematisiert, und es wird gezeigt, wie mit digitalen Prüfungspraktiken das Versprechen einhergeht, auf die genannten Probleme zu reagieren und das schulische Lernen insgesamt zu optimieren. Gleichzeitig wird deutlich gemacht, dass auf diese Weise hinsichtlich des Lernsubjekts, des Lerngegenstands und der pädagogischen Beziehungen Verengungsdynamiken produziert werden. Der Beitrag schliesst mit der Frage danach, inwiefern das Auslagern der Prüfung an digitale Tools mit einem Verantwortungsverlust der pädagogisch Handelnden verbunden ist.

Remarks on the Educational Distrust of Examinations and the Promise of a Digital Problem Solving

Abstract

This article explores several forms of educational distrust towards examinations. In the first part of the paper, the discourse on school performance evaluation is scrutinized and three issues will be discussed: Firstly, reference is made to how school examinations serve as a selection mechanism. Secondly, the difference between standardization and

individualization of assessment is considered. Thirdly, the predictability of future success on the grounds of grades is discussed. These three arguments nourish the distrust of examination and bring about an unresolvable dispute with respect to the appropriate measures of performance evaluation and their choice. However, the dispute also concerns the question of how to perform diagnostics in a way that is educationally acceptable. Scholars from the new research field of Learning Analytics see digitalized practices of assessment as solutions to these problems, while also promising an optimization of processes of school-based learning. At the same time, however, it is obvious that these new forms produce reductions in complexity for both the subject and the object of learning, as well as the pedagogical relation in general. The article concludes with the question whether digitized examinations bring about a loss of pedagogical responsibility for those involved in education.

1. Einleitung

Am 29.01.2021 twitterte die *heute-show*: «Viele Schüler in Deutschland bekommen heute ihr #Zeugnis. Die Note in Sport ist dieses Mal noch ein bisschen willkürlicher als in den Jahren zuvor» (@heuteshow, 29. Januar 2021). Das Statement, welches sich auf die Leistungsbeurteilung unter den Bedingungen der aktuellen Coronakrise bezieht, zeugt von einer allgemein anerkannten Selbstverständlichkeit dahingehend, dass Noten etwas Willkürliches anhaftet – und das nicht nur bezogen auf den Sportunterricht. Daneben ist es auch kein Geheimnis, dass das Zensieren und Benoten zum «leidigen» Geschäft des Berufsalltags von Lehrpersonen gehört (Terhart 2009, 39). Das Prüfen und Benoten gilt vielen Pädagoginnen und Pädagogen selbst als «ungeliebt und <schmutzig>» (Breidenstein 2018, 317). Während auf Seiten der Schulpädagogik Probleme oder Fallstricke der Leistungsbewertung im Konkreten oft (selbst-)kritisch thematisiert werden, ist in allgemeinpädagogischer Perspektive gerade die langanhaltende Nichtthematisierung der Prüfung in einem weiteren Sinne auffällig (vgl. Loch 1963; Ricken und Reh 2017). Dies mag u. a. daran liegen, dass spätestens seit Foucaults Analysen zu den Macht- und Subjektivierungseffekten von Prüfungen ihr gegenüber «Vorsicht» geboten ist. In «Überwachen und Strafen» (Foucault 1994) wird die Prüfung als ein Kernelement der Disziplinarmacht beschrieben. Mit dem Verweis darauf, dass die Prüfung «die Techniken der überwachenden Hierarchie mit denjenigen der normierenden Sanktion» (ebd., 238) kombiniert, zeigt Foucault, dass die Möglichkeit, über die Prüfung «Wahrheit» – auch im Sinne eines «Wissens» über die Geprüften – zu ermitteln, stets an disziplinierende, normalisierende, subjektivierende Machteffekte gebunden ist. Die Schule bspw., die durch Prüfungspraktiken zu einem Ort der Erkenntnisbildung und der Hervorbringung von Wissen wird, wird gerade dadurch permanent auch zum Austragungsort von Disziplinierungspraktiken (vgl.

Foucault 2015, 83ff.). Es ist u. a. dieses disziplinierende Erbe, das der Prüfung nach wie vor anhaftet, was sie aus pädagogischer Perspektive problematisch erscheinen lässt.

Im Horizont dieser Beobachtungen artikuliert sich ein pädagogisches Misstrauen gegen die Prüfung (vgl. auch Schröder und Thompson 2015), das im Folgenden an einigen konkreten Linien nachgezeichnet werden soll. Ich werde dafür einen allgemeinpädagogischen Blick auf die diskursive Herstellung des Misstrauens, v. a. im Hinblick auf schulpädagogische Diskurslinien werfen (2). Anschliessend wird mit einem weiten Begriffsverständnis von Prüfung als einer ‹Objektivierung von Lernprozessen› argumentiert und am Beispiel von Learning Analytics gezeigt, inwiefern neue Formen der Digitalisierung die Optimierung von Lernen und Lernoutput versprechen und so Auswege aus dem Misstrauen möglich scheinen (3). Einer Problematisierung dieser Hoffnungen aus bildungstheoretischer Perspektive (4) folgt dann die Frage nach den pädagogischen Konsequenzen der Digitalisierungsversprechen (5).

2. Spielarten des pädagogischen Misstrauens gegen die Prüfung

Im Folgenden werden drei exemplarische Argumentationslinien im schulpädagogischen Diskurs nachgezeichnet, an denen sich ein Misstrauen gegen die Prüfung konkretisiert.

2.1 *Der gesellschaftliche Auftrag der Schule: Selektion*

An der Selektionsfunktion der Schule gewinnt das Misstrauen gegen das ‹disziplinierende Erbe der Prüfung› Kontur: So gut wie keine Publikation der schulpädagogischen Forschungsliteratur, die sich explizit auf Noten, Zensuren, Zeugnisse und Beurteilungspraktiken in Schulen bezieht, kommt ohne die Erläuterung der ‹Funktionen› von Leistungsbewertung aus (vgl. Sacher 2014; Meyer und Jansen 2016; Jürgens und Lissmann 2015). Eine oppositionale Grenze, die dabei stets in Anschlag gebracht wird, ist die zwischen der ‹pädagogischen Funktion›, die im Dienste der Förderung des Lernens stehe, einerseits und der ‹gesellschaftlichen Funktion›, die auf die Verteilung von Lebenschancen ziele, andererseits (vgl. Ingenkamp und Lissmann 2008, 20).

In einer historischen Genese des ‹Selektionstheorems› zeigt Breidenstein (2018), dass seit den 1970er Jahren die Vorstellung genährt wird, dass die Schule ‹für› die Gesellschaft – im Sinne einer Dienstleistung – Aufgaben erbringe, die in Differenz zu ihrem pädagogischen Auftrag stünden. Auf diese Weise könne sie sich distanzierend zum ‹Selektionsauftrag› positionieren. Indem Breidenstein aber deutlich macht, dass Selektion im Unterrichtsgeschehen selbst funktional ist, bspw. um Unterrichtsprobleme zu lösen, und somit meist für die Schule selbst statt für die Gesellschaft

selektiert wird (vgl. ebd., 314, 324; vgl. auch Tillmann 2008), äussert er den «Verdacht» (Breidenstein 2012, 14), dass die Unterscheidung zwischen gesellschaftlicher und pädagogischer Funktion der Prüfung zur Gewissensberuhigung diene: Mit ihrer Ausweichoption der «Zuliefererposition» für «die Gesellschaft» ist es möglich, die Notengebung «als «unpädagogisch» und dennoch notwendig» auszuweisen (Breidenstein 2018, 314).

Mit der Etablierung dieser höchst eigentümlichen Differenz und Gegenüberstellung zwischen «der Pädagogik» und «der Gesellschaft» kann der pädagogische Anspruch als frei von Zwängen, Disziplinierungen, Normalisierungen und gesellschaftlichen Anforderungen imaginiert und ein von disziplinierenden (Macht-)Praktiken befreiter Möglichkeitsraum geschaffen werden, in dem ganz bestimmte Effekte oder Funktionen von Prüfungen «eigentlich» nichts verloren haben.

Das Misstrauen, das die Narration, die Prüfung wäre «ein «unpädagogisches Verhalten», das mit der Erziehungsaufgabe im Grunde unvereinbar sei» (Loch 1963, 181) befeuert, ist funktional dafür, sich «von potentieller Verantwortung» zu entlasten (Breidenstein 2018, 325).

2.2 *Standardisierung oder Individualisierung der Leistungsmessung?*

Eine weitere Opposition, die in der schulpädagogischen Forschungslandschaft auftaucht und sich v. a. über die Wahl angemessener (Prüfungs-)Instrumente auf die Möglichkeit von gerechter und objektiver Leistungsmessung bezieht, ist die zwischen der Standardisierung und Individualisierung von Leistungsmessung. Die Differenz verläuft häufig zwischen psychologisch-diagnostischen Forderungen nach mehr Objektivität einerseits und einer an der Subjektivität der Lernenden ausgerichteten Bewertungspraxis andererseits (vgl. Rabenstein, Idel, und Ricken 2015).

Ausgangspunkt für Forderungen nach *standardisierten* Verfahren sind dabei häufig Probleme, die sich aus der Subjektivität der Lehrenden wie auch der Lernenden und aus damit einhergehenden Mess- und Urteilsfehlern ergeben, oder die Forderung nach allgemeingültigen Gütekriterien der Bewertung. Ingenkamp, der als Begründer der pädagogischen Diagnostik gilt, hatte in den 1960er Jahren gezeigt, dass die Unmöglichkeit, Beurteilungen auf objektive Kriterien zu stützen, dazu führt, dass Zensuren über die eigene Klasse hinweg keinen Vergleichswert besitzen (Ingenkamp 1971c, 161). Die Form der nicht-standardisierten Bewertung wird dann als ungerecht ausgewiesen, weil sie einerseits vom Zufall der Lehrkraft und andererseits von dem der Klassenzugehörigkeit abhängt (Ingenkamp 1968, 427; 1971c, 161). Eine Möglichkeit, diesen Fehlern beizukommen, bestehe demnach darin, strengere Regeln für die diagnostischen Verfahren zu etablieren und diese zu formalisieren.

Gegenläufig dazu gehen die Forderungen nach einer *individualisierten* Leistungsmessung davon aus, dass «Gerechtigkeit» bei der Bewertung nur dadurch zu erzeugen ist, dass die einzelnen Lernenden und ihre Entwicklungspotenziale selbst in den Blick gerückt werden. Es sei von einer grundsätzlichen Heterogenität der Schülerschaft auszugehen, die sich auch auf den Unterricht selbst auswirken müsse. Die Bewertung soll entsprechend das Individuum und seine Entwicklung selbst zum Massstab nehmen sowie gleichzeitig vom Individuum und seinem (evaluativen) Blick auf sich selbst ausgehen. Konkretisiert und realisiert wird dies bspw. im individualisierten Unterricht, etwa durch die Arbeit am Portfolio (vgl. Bohl 2009; Häcker 2011).

Auch in dieser «Opposition», d. h., im Streit um standardisierte (objektive) oder individualisierte (subjektive) Prüfungspraktiken zeigt sich das Misstrauen gegen die Prüfung: einerseits spezifisch in der Frage nach den angemessenen «Mess»-Instrumenten, andererseits in der permanent geführten Frage nach der Sinnhaftigkeit der Notengebung im Allgemeinen (vgl. Ingenkamp 1971a).

2.3 Prognostische Potenziale der Leistungsmessung

Als weiteres Problem der Leistungsmessung lässt sich die Infragestellung der Möglichkeit, aus aktuellem Lernerfolg einen künftigen Lernerfolg vorherzusagen, markieren. Nach Langfeldt gehöre es gar «zum Allgemeingut pädagogischen Wissens [...], daß Zensuren keine Grundlage für sichere Prognosen abgeben können» (Langfeldt 1984, 94). Auch hier zeigte sich mit den Studien des Bandes von Ingenkamp (1971a) v. a. die Problematik eines mangelhaft abgesicherten empirischen Wissens, weil es bis dato im deutschen Sprachraum aufgrund fehlender standardisierter Tests keine Langzeitstudien oder Vergleichsmöglichkeiten von aktuellem und späterem Schulerfolg gab, die Prognosen auf einen sicheren Boden hätten stellen können. Gleichwohl stellte Ingenkamp fest, dass es um den prognostischen Wert von Schulzensuren schlecht bestellt sei und es bspw. keinen Zusammenhang zwischen den Grundschulnoten und dem späteren Leistungserfolg auf dem Gymnasium gebe (vgl. Ingenkamp 1968, 415; Ingenkamp und Lissmann 2008).

Nach wie vor wird insbesondere für die Bereiche der Prognose und Vergleichbarkeit der Zensuren die Aktualität der Diagnose einer «Fragwürdigkeit der Zensurengebung» (vgl. Ingenkamp 1971a) betont (vgl. Breidenstein 2012, 14; Ingenkamp und Lissmann 2008), indem bspw. explizit eine Ziffern- und Zeugnisnoten ergänzende Beratung gefordert wird (vgl. Bohl 2009, 87).

Diese kurz angerissenen diskursiven Spuren zeigen, dass die Prüfung v. a. im pädagogischen Handlungsfeld Schule ein strittiges Thema ist – und die Frage ihres angemessenen Einsatzes noch mehr. Durch permanente Reform- und Alternativvorschläge bezüglich der Leistungsmessung zeigt sich eine Infragestellung ihrer Bedeutung

und Geltung. In entscheidendem Masse berührt also die (Aufgabe der) Prüfung die Frage nach der pädagogisch angemessenen Weise der ‹Objektivierung› und ‹Kontrolle› von Lernleistung.

3. Hoffnung Digitalisierung

Eine mögliche Antwort auf diese Frage bietet das ‹Digitalisierungsversprechen› bzw. der Einsatz von Learning Analytics (vgl. auch Schröder 2021). Learning Analytics (LA) lassen sich im weitesten Sinne als Techniken zur Akkumulation und Auswertung von Daten zu Lernverhalten verstehen. Der Forschungsbereich rund um LA weitet sich seit fünf Jahren stark aus (vgl. Ifenthaler und Drachslar 2020, 516) und mittlerweile existiert auch in Deutschland ein differenziertes Angebot (vgl. Hartong 2019a). Learning Analytics setzen mit einem simplen Versprechen an: Durch die Sammlung und Auswertung von Daten lassen sich pädagogisches Handeln und das Lernen insgesamt *optimieren*: Learning-Analytics-Anwendungen, die «eigenständige Softwarelösungen oder in Lernplattformen eingebettet» sind (Ifenthaler und Schumacher 2016, 177), sollen durch die Erhebung, Analyse und Auswertung von Daten der Lehrenden und Lernenden, des Unterrichts, des Klassenraums, der Lernumgebung etc. Einblicke in bislang kaum beobachtbare Zusammenhänge ermöglichen und auf diese Weise (Lern-)Prozesse optimieren (vgl. Larusson und White 2014, 1).¹

Durch Datengenerierung und die Auswertung der lernbezogenen Daten über Analysealgorithmen können entsprechende Rückmeldungen zur Optimierung von Lernprozessen und -umgebungen in Echtzeit gegeben werden (vgl. Ifenthaler und Drachslar 2020, 517). Hier könnte von einer zweigliedrigen Informationsbasis für die Rückmeldung zu Lernprozessen gesprochen werden: Daten zur aktuellen Performance und zum Lernprozess werden gesammelt und analysiert und «in Bezug zu lernpsychologischen Erkenntnissen gesetzt» (Ifenthaler und Schumacher 2016, 177). Auf dieser Basis werden algorithmisiert Prognosen generiert und Möglichkeiten zur Optimierung des Lernverlaufs oder der Lernumgebungen etc. nahegelegt.

Diese Optimierung bezieht sich dabei auf alle am Prozess Beteiligten gleichermaßen.² *Lernende* sollen individuelle Informationen rund um ihre eigenen (Lern-)Bedürfnisse geliefert und Wege aufgezeigt bekommen, die eigenen Lernanstrengungen zu verbessern. Learning Analytics sind darauf ausgelegt, Lernenden zu ermöglichen, «selbstverantwortlich mit den Informationen eines Feedback-Systems (z. B. Learning-Analytics-Dashboard) umzugehen und das bisherige Verhalten bei Bedarf

1 Als derlei Einblicke nennen Ifenthaler und Schumacher (2016) u. a. Folgendes: «Welche zur Verfügung gestellten Dateien werden wann genutzt, wie lange hält der Lernende sich in der Lernumgebung auf, welche Pfade verwendet er, welche Beiträge verfasst er in Diskussionen und wie ist der individuelle Lernfortschritt im Vergleich zur Gruppe oder zum Lernziel?» (ebd., 176).

2 Bemerkenswert ist, dass LA auch auf bildungspolitischer Ebene wirksam werden sollen (vgl. Ifenthaler und Schumacher 2016, 167).

anzupassen» (Ifenthaler und Drachslers 2020, 524). Vonseiten der Bildungsinstitutionen können massgeschneiderte Lernangebote zur Verfügung gestellt werden. *Lehrende* können auf diese Weise nicht nur die Lernenden zielgerichtet unterstützen, sondern auch ihr Lehrangebot an die Lernenden anpassen (vgl. Ifenthaler und Schumacher 2016, 177). Den Lehrenden wird es durch die errechneten lernbezogenen Prognosen möglich, «[g]efährdete Lernende [zu] identifizieren» (Ifenthaler und Drachslers 2020, 525). Learning-Analytics-Anwendungen versprechen auf diese Weise «Antworten» auf die vorhin im Zusammenhang mit Prüfung genannten Probleme, wie im Folgenden gezeigt werden soll.

3.1 Prognostische Potenziale von Learning Analytics

Pardo (vgl. 2014) kennzeichnet fünf Stadien der Funktionsweisen von LA, von denen eine die der *prediction* ist. Diese wird, im Anschluss an die erste Sammlung der Daten und ihrer Weitergabe an die verschiedenen Stakeholder – wie bspw. Lehrende –, als dritte Phase im Prozess der LA-Intervention gekennzeichnet. Prognosemöglichkeiten können sich bei LA auf vielfältige Weise auf das «Wissen» der Lernenden oder auch auf deren «affect» (Baker und Inventado 2014, 63) beziehen, so bspw. darauf, wie wahrscheinlich es ist, dass eine Schülerin oder ein Schüler einen spezifischen Kurs besteht. Alternativ lässt sich basierend auf den aktuellen Daten bestimmen, an welchen Stellen im Lernprozess voraussichtlich Schwierigkeiten auftauchen werden. Die folgende Phase, *act*, ist entsprechend dadurch gekennzeichnet, auf diese Prognose zu reagieren und sie im besten Falle umzukehren. Es geht also darum, «a student's future outcomes» vorherzusagen (ebd.) und potenziell zu intervenieren. Diese Phase der *prediction* ist für alle Stakeholder von besonderer Bedeutung, weil hier die LA auf automatisierte Weise je spezifische «Entscheidungen» oder «Interventionsmöglichkeiten» nahelegen und den jeweiligen Stakeholdern übermitteln:

«Prediction algorithms may help students anticipate difficulties during an experience, help instructors to identify students that are not performing as expected, and help administrators to anticipate complications in a course» (Pardo 2014, 17).

Die Daten sollen damit die Anpassung, Optimierung und kontrollierte Planung des Lerngeschehens auf unterschiedlichen Ebenen ermöglichen. Den vorhergesagten Problemen kann präventiv sowohl auf individueller Ebene als auch bspw. auf der Ebene der Unterrichtsentwicklung begegnet werden. Ein Anliegen der LA besteht also in einer «vorwegnehmenden Zukunft» möglicher Szenarien, die ausgehend von gegenwärtigen Szenarien errechnet und prognostiziert werden: «the true power of this data is to derive models capable of anticipating events that will occur in the future» (ebd., 29).

3.2 *Standardisierung und Individualisierung*

Die Möglichkeiten, Leistungen zu prüfen, gestalten sich mittels LA als eine Mischung aus standardisierten Vorgaben und individuellen Lösungsmöglichkeiten. Sie beanspruchen dabei zugleich, eine Objektivität bei der Messung herzustellen wie auch auf das individuelle Entwicklungspotenzial der Lernenden Bezug nehmen zu können. Deutlich wird dies am Beispiel der Einbindung von LA in Cloud-Systeme an Schulen.³

Höhne, Karcher und Voss (2020) analysieren die pädagogischen Konsequenzen der Wissenstransformation und der Art und Weise der Vermittlung und Aneignung von Wissen durch die Einführung schulischer Cloud-Systeme. Mit der Cloud als ‹Wissensdatenbank› der Schule verbindet sich sowohl der Anspruch, durch Digitalisierung ‹Wissen› und Möglichkeiten des Zugriffs auf Wissen für alle gleichermassen zentral zur Verfügung zu stellen, als auch der, es den einzelnen Lernenden zu ermöglichen, individuell und selbstgesteuert darauf zuzugreifen.

Es ist gerade diese Mischung zwischen einer Standardisierung des Wissens unter der Ägide des New Public Managements und der Forderung nach einer individuellen Aneignung von Kompetenzen, die Höhne, Karcher und Voss herausstellen: Die Daten in der Schul-Cloud lassen sich als ‹standardisierte Informationseinheiten zur Bearbeitung durch die individuellen LernerInnen› (ebd., 326) begreifen. Die Arbeit mit der Cloud ist so gleichermassen Ausdruck der auf Standards, Kompetenzentwicklung und Outputsteuerung angelegten ‹Bildungsreform nach PISA› (ebd.), wie sie zugeschnitten ist auf individuelle Lernwege, Problemlösungen und (Lern-)Entwicklungspotenziale, die es ‹eigenverantwortlich› zu entfalten gilt. Der Umgang mit der Schul-Cloud, die standardisierte Bausteine zur Kompetenzaneignung liefert, liesse sich also als eine Form unterrichtlicher individualisierter Praxis verstehen (vgl. ebd., 328).

Dabei sind es v. a. die LA-Anwendungen und das mit ihnen verbundene Feedbacksystem, welche ‹Möglichkeiten der Individualisierung› (Ifenthaler und Schumacher 2016, 178) versprechen. Die Informationen, die LA über das Lernverhalten der Lernenden herausgibt und die Weise, wie damit ein ‹Wissen über das eigene Lernen› (Ifenthaler und Drachler 2020, 524) produziert wird, verstehen sich als Anreize zur Reflexion, um Lernprozesse selbstgesteuert zu optimieren. Damit operiert diese Art des subjektiven Reflexionsanreizes sehr viel stärker über das Versprechen eines ‹objektiven Urteils›, als das bspw. das Portfolio kann: Die algorithmisch berechnete Rückmeldung zum Lernverhalten kann sich über ihre ‹Objektivitätsaura› (Heintz 2007, 79) legitimieren.⁴

3 Höhne, Karcher und Voss zeigen in ihrer Analyse, dass von den Apologetinnen und Apologeten der Schul-Clouds befürwortet wird, LA ‹zu einem wichtigen Teil des Schul-Cloud-Projekts› zu machen (Höhne, Karcher, und Voss 2020, 332).

4 Der Begriff ‹Aura› verweist darauf, dass Objektivität selbst eine normative Zurechnung ist (vgl. Heintz 2007, 80). Allerdings ist eine Dekonstruktion der Objektivität als Zurechnung voraussetzungsvoll, weil es dafür meist ‹alternative Zahlen resp. ein Wissen darüber [braucht], auf welche Weise sie zustande kamen› (ebd., 78).

Der Problematik eines ›subjektiven Urteils‹ soll dadurch begegnet werden, dass LA Möglichkeiten einer quantifizierbaren Feedbackkultur schaffen, die Reflexionsanreize, Lernprozessanalysen und die Optimierung des Lernens nicht mehr (nur) an die Lernenden selbst oder an die Rückmeldung durch Lehrende bindet, sondern das Verhältnis des Subjekts zu seinen Daten in den Vordergrund rückt.

Die Nutzung von LA-Anwendungen stellt dabei in Aussicht, ein möglichst ›objektives‹ Bild des Lerngeschehens zu erhalten, frei von subjektiven Einflüssen oder Störfaktoren (vgl. Pardo 2014, 15). Die Objektivität wird vor allem durch die Tools selbst hergestellt: Sie versprechen eine ›Neutralisierung‹ von ›Subjektivität‹ und die Auflösung von Urteilsfehlern durch Lehrende (vgl. Paradies, Wester, und Greving 2009) über algorithmisch legitimierte Ordnungskriterien. Querliegend zu z. B. aktuellen Diskursen rund um die Standortgebundenheit von Wissen, wird so ein ›view from above, from nowhere‹ (Haraway 1988, 589) beschworen. Die datenerzeugenden Technologien werden dabei selbst als ›neutral‹, d. h. ›ideologiefrei‹ dargestellt – erst ihre Nutzung durch verschiedene Stakeholder und deren Interessen würden eine spezifische (normative) Wirkung zeitigen: ›Surely these neutral tools can be used well or badly depending on who is involved‹ (Selwyn 2019, 11).

Man kann sich denken, dass es von Seiten der Pädagogik kritische ›Anfragen‹ an diese Versprechen zur Optimierung des Lernens durch LA gibt, angefangen bei der Neutralität oder Harmlosigkeit der Tools, die v. a. aus machtkritischer Perspektive zurückgewiesen wird (vgl. Hörtnagl 2019). Hartong problematisiert, dass die Tools und die Weise ihrer Wissensproduktion normativ aufgeladen sind und Ungleichheitsverhältnisse stabilisieren, indem sie kategorisieren und selektieren (vgl. Hartong 2019b, 2020). Von einer Neutralität der Tools liesse sich so nur unter Ausblendung gesellschaftlicher, soziokultureller Verhältnisse ausgehen.⁵ Wie sich u. a. an der genannten Analyse der Veränderung von Wissensvermittlung in der Schule durch Clouds zeigen lässt (vgl. Höhne, Karcher, und Voss 2020), sind Daten nicht einfach ›Abbild von Realität‹, sondern sie stellen diese her und üben daher ›eine oftmals unterschätzte Macht‹ aus, die weder neutral noch unpolitisch sein kann (Hartong 2019b, 7). Es stellt sich daher die Frage nach den ›ideologies and politics of the technologies we find ourselves using‹ (Selwyn 2019, 12).

Diese Frage, so lässt sich mit Bezug auf das Misstrauen gegen die Prüfung sagen, ist eine genuin pädagogische. Wenn in kritischer Manier danach gefragt wird, ›Welche Logik(en) der Selektion, der In-Formation und der In-Wert-Setzung von Lernen – oder gar Bildung – finden wir hier?‹ (Hartong 2019b, 10), eröffnet sich das Feld der (pädagogischen) Auseinandersetzungen darüber, was Lernen und Bildung bedeuten

⁵ Ein sprechendes Beispiel dafür ist, dass anhand algorithmischer Profilbildungen über zugeschriebene Kategorien wie *class*, *race* und *gender* Personen diskriminiert und selektiert werden oder Chancen wie Risiken und bspw. biopolitische Massnahmen an errechnete ›Identitäten‹ von Nutzerinnen und Nutzern und damit verbundene Zuschreibungen an ›Gruppenidentitäten‹ gebunden werden (vgl. Someh et al. 2019, 726).

sollen. Dass solche Diskussionen im Bereich der LA-Anwendungen überhaupt keine Rolle zu spielen scheinen, lässt nicht nur das den LA inhärente Lernverständnis als fraglich erscheinen.

Gleichermassen diskussionswürdig sind auch die pädagogischen Konsequenzen, die aus der Einsicht zu ziehen sind, dass die Verheissungen auf problematischen Voraussetzungen basieren.

4. Verengungsdynamiken

Aus bildungstheoretischer Perspektive ist die Frage nach einer möglichen reduktionistischen Auffassung von Lern- und Bildungsprozessen v. a. über die Umstrukturierung des Bildungswesens in Formen der Neuen Steuerung und des New Public Managements problematisiert worden. Auch ohne Bezugnahme auf die Digitalisierung des Lernens verweist Ruhloff (2009) auf die Umstellung pädagogischer Prozesse hin zu einer «Rationalität der Datenerhebung und -verarbeitung», durch die eine Relativierung von Bildung und Lernen stattfindet. Als relevant erscheine nur mehr das, was «testfähig» sei (ebd., 199). Aus Ruhloffs Perspektive ist eine auf «Bildung» ausgerichtete pädagogische Praxis im Sinne einer Eröffnung von Problemhorizonten zu verstehen, die Denken und Reflexion allererst ermöglicht. Im Sinne der Digitalisierungsversprechen wird demgegenüber die Überwindung von Problemen als Zielpunkt des Lernens anvisiert. Momente der Entzogenheit, Unkontrollierbarkeit oder auch der Störung werden auf diese Weise aus den Lernprozessen getilgt.

Es ist daher ein Unterschied ums Ganze, ob Unterrichtsinhalte und Prüfungsformate darauf ausgerichtet sind, Fragehorizonte zu eröffnen, und von der Schwierigkeit der Planbarkeit von pädagogischen Prozessen ausgehen oder ob sie vor dem Hintergrund einer Steuerungsrationaltät agieren, die Lernprozesse als identifizierbar, kontrollierbar und vorhersagbar versteht. Auch wenn es bei LA gerade um die Mobilisierung der Selbststeuerungspotenziale und Reflexion der Lernenden geht, beschreibt dies eine Aufforderung, ein Verhältnis zu «seinen» vergangenen Lerndaten einerseits und zu den probabilistisch errechneten zukünftigen Lernvollzügen andererseits einzunehmen. Es geht dabei wohl weniger um Selbstaussetzung, Entzug und Fremdheit; vielmehr wird Reflexion zu einem als «transparent» imaginierten Verhältnis zwischen Gelernt-Haben und Noch-zu-Lernen. Die LA produzieren also durch ihr permanentes Evaluations-, Feedbackschleifen-, Ranking-, Prüfungssystem zwar eine riesige Dynamik mit einem anhaltenden Anreiz zur Reflexion, aber es ist eine «Verengungsdynamik» (ebd.) mit einer «mehr oder weniger leeren Reflexivität» (ebd., 201f.). Einige der Konsequenzen, die aus einer solchen Verengung des Blickfeldes folgen, möchte ich im Folgenden aufzeigen.

4.1 Verengungen des Lernsubjekts

Die proklamierte Verengungsdynamik arbeitet mit Steuerungsphantasien, die das Bild einer «Machbarkeit des Menschen» (Ruhloff 2009, 198) zur Grundlage haben. Karcher zufolge liegt den LA-Anwendungen die implizite Annahme einer kybernetischen Anthropologie zugrunde: Die Lernenden werden als «Summe der über sie erhobenen Daten» (Karcher 2020, 162) verstanden. Es sind gerade die Ideen einer «völlige[n] Lesbarkeit des Selbst in Daten» (Bächle 2016, 172) und die einer «technisch steuerbaren Lernanregung» (Hartong 2019a, 426, Herv. im Original), welche die Vorstellung pädagogischer Steuerbarkeit von Lern- und Bildungsprozessen nähren und Interventionsmassnahmen als technologisch herstellbar imaginieren.

Die Möglichkeit zur Prognose ergibt sich aus der Annahme, dass aktuelle Daten Muster für zukünftige Daten liefern, d. h., dass sich Vorhersagen zu künftigen Lernleistungen aus aktuellen Performances berechnen lassen. Es ist dies die Annahme, «dass die Muster der Vergangenheit sich in Zukunft wiederholen werden» (O’Neil 2018, 57). Die prognostischen Urteile über Potenziale von Lernenden basieren somit auf der Vorstellung einer relativ störungsfreien «Kontinuität» im Lernverhalten einerseits und auf den von Hause aus begrenzten Möglichkeiten der Tools andererseits – begrenzt deshalb, weil für die algorithmischen Berechnungen nur ein begrenztes Set an Daten und Parametern zur Verfügung steht.⁶

4.2 Verengungen des Lernbegriffs

Die LA-Anwendungen operieren über behavioristische Steuerungen, die sich vielfach den Techniken des *Nudging* bedienen:

«Rather than working to support <learning> per se, learning analytics is essentially a form of individually focused behaviour management – sometimes termed <nudging> people’s decision-making and action» (Selwyn 2019, 15).

Über das *Nudging* («anstossen», «schubsen»), d. h. über die Schaffung von Anreizen zur Verhaltensänderung, wird v. a. Wert auf Lust, Vergnügen und Spass im Lernprozess gelegt und Momente der schmerzhaften und irritierenden Fremdheitserfahrung werden aus dem Lernen ausgelagert (vgl. auch Hartong 2019b). Das von den LA normalisierte Lernen suggeriert also Abenteuer, Lebensnähe und Alltagsrelevanz, die sich über Anreize wie Punktvergaben oder Levelsysteme ausgestalten (vgl. bspw. Williamson (2017), der das ClassDojo analysiert).

⁶ Daher sehen bspw. Greller und Drachsler eine Gefahr darin, dass Lernpotenziale unterminiert, statt unterstützt werden könnten: «For example, not every learner who has difficulties mastering subject level two, will automatically not master level three» (Greller und Drachsler 2012, 48). Im Zweifel werden auf diese Weise die Lernenden nicht nur auf eine bestimmte Gegenwart, sondern auch auf eine bestimmte Zukunft hin «festlegt» (vgl. Foucault 1994, 247).

Es sind gerade die ›kindgerechten‹ Lernplattformen, die auf die Differenz zwischen einer Normalität der langweiligen und trockenen Wissensvermittlung in der Schule, in der v. a. Bücher und Reaktivität in Bezug auf das Unterrichtsgeschehen vorherrschten, und dem Bild einer aufregenden und das Kind autonomisierenden, aktiven, alltagsnahen Lernumgebung setzen. Der Unterricht soll ›gamifiziert‹ werden – eine Transformation, die sich als problematische Reaktualisierung reformpädagogischer Narrative verstehen lässt (vgl. Buck 2017, 268) und einmal mehr deutlich macht, dass die ›Freisetzung‹ einer kindlichen Lust am Lernen durch gamifizierte Lernumgebungen keineswegs die Verabschiedung von pädagogischen Disziplinierungen bedeutet.⁷

Hinter der auf Selbststeuerung und Nudging ausgelegten Lenkung verbirgt sich eine Infantilisierung der Lernenden (vgl. Selwyn 2019, 15), wenn ihre ›Arbeit‹ auf das begrenzt wird, was sie als nächstes anklicken sollen, mit welcher Übung, welchem Test das nächste Level erreicht werden kann. Im Sinne eines «teaching to the algorithm» (ebd., 13) spricht Selwyn davon, dass sowohl auf der Seite der Lernenden als auch auf der der Lehrenden vor allem die Dinge im Fokus stünden, die im Blickpunkt der Datenanalyse wären:

«Students will act in ways that ›please‹ the indicators, or else attempt to ›game‹ the system logic. Teachers will work to ›second guess‹ what the system will reward» (Selwyn 2019, 13).

Besonders problematisch erscheint dies vor dem Hintergrund, dass die «indicators» keineswegs durch *pädagogische Theorie* abgesichert sind: Aus Perspektive der Wirtschaftspädagogik verweisen Ifenthaler und Schumacher bspw. darauf, dass Nachholbedarf bestehe hinsichtlich der

«empirisch validierten Auswahl von Indikatoren für die Analysealgorithmen [...], da Lernen häufig mit Variablen wie Login-Häufigkeiten, Anzahl der Gruppendiskussionsbeiträge oder der termingerechten Bearbeitung gleichgesetzt wird. Doch nicht nur die Menge der Beiträge oder die Nutzungsdauer, sondern vor allem der Inhalt ist ein Indikator für Lernen» (Ifenthaler und Schumacher 2016, 181).

⁷ Im Gegenteil: Eine Besonderheit der LA besteht darin, permanente, d. h. räumlich und zeitlich entgrenzte, Beobachtungsverhältnisse zu etablieren, über die die Akkumulation grosser Datenmengen überhaupt erst möglich wird – die Daten werden im Sinne der Dataveillance (vgl. Bächle 2016) nicht mehr (nur) durch klassische Prüfungen, Testverfahren oder Lernstandserhebungen etc. erfasst, sondern vielmehr «beiläufig während der Interaktion mit den digitalen Bildungsmedien» (Karcher 2020, 157). Weil also im Modus der Gamification eine Art Compliance der Kinder in Bezug auf ihre Unterwerfung geschaffen (vgl. Williamson 2017, 7) und es Lehrenden wie Eltern möglich wird, eine permanente Kontrolle des Lernprozesses der Kinder zu realisieren (Williamson spricht daher von einer «ludic surveillance»; ebd., 6), liesse sich von einer entpolitizierenden Praxis sprechen, die nicht die Disziplinierungen, sondern vielmehr die Möglichkeiten zum Einspruch gegen ebensolche abschafft (vgl. Buck 2017).

Dass ernsthaft daran erinnert werden muss, dass nicht nur die Verweildauer innerhalb der Anwendung, sondern auch ‹der Inhalt› als Indikator von Lernen zu beachten wäre, ist pädagogisch höchst bedenklich.

4.3 Verengungen der Lernbeziehung

Der fehlende Bezug zur Pädagogik in den LA-Konstruktionen hat auch Auswirkungen auf die Rolle der Lehrenden: Ein Vorteil von LA wird darin gesehen, dass sie neue Formen des Feedbacks ermöglichen, die nicht zwangsläufig an Interaktionen gebunden sind:

«Learning Analytics bieten durch die Analyse großer Datenmengen eine weit- aus differenziertere Informationsbasis als das in klassischen Lehr-Lern-Situati- onen durch eine einzelne Lehrperson möglich wäre» (Ifenthaler und Drachsler 2020, 524).

Der Blick auf schulpädagogische Auseinandersetzungen ist an dieser Stelle wie- derum interessant, weil hier schon seit Jahren anhaltend die Frage der Aussage- kraft einer Ziffernote oder Zeugniszensur diskutiert wird. Jürgens spricht von ei- nem Versagen der Note im Hinblick auf ihre ‹Rückmeldefunktion› für die Lernenden: Die Zensur könne eine Auskunft darüber, «wie weitergelernt werden kann, welche Schwächen noch überwunden werden müssen» *nicht* geben, denn dazu benötige es «vielschichtigere Leistungsbeschreibungen» (Jürgens 2010, 66).

Während für Ingenkamp, der sich bekanntermassen für eine stärkere Standardi- sierung von Leistungsmessungen aussprach, noch galt: «Kein diagnostisches Hilfs- mittel kann den Diagnostiker ersetzen» (Ingenkamp 1968, 423), scheint die Diskus- sion um die ‹Eigentümlichkeit› der diagnostischen Aufgabe in den Diskussionen um LA völlig abhandengekommen zu sein. Die durch Zahlen und Daten erzeugte Evidenz beansprucht, den pädagogischen Verhandlungsraum um das angemessene Gestal- ten von Rückmeldungen zu Lernprozessen technologiebasiert zu schliessen.

Mit LA wandelt sich nicht nur das diagnostische, sondern überhaupt das pädä- gogische Handeln zu einer distanzierten Beobachtung und einem datengestützten Intervenieren auf potenzielle Zukünfte hin (vgl. Karcher 2020, 164). Die direkten päd- agogischen Interaktionen zwischen Lernenden und Lehrenden werden dabei zuneh- mend überflüssig (vgl. Hartong 2019a, 440).⁸ Es zeigt sich also: Die Hoffnungen, dass sich das ‹pädagogische Problem Prüfung› mit dem ‹Digitalisierungsversprechen› in den Griff bekommen lässt, schafft neues Unbehagen – eines, welches das pädagogi- sche Verhältnis selbst betrifft.

⁸ Dramatisch und aufschlussreich sind diesbezüglich aktuelle Bemühungen, die die unbeobachtete und ‹un- mittelbare› pädagogische Interaktion als ein klassenspezifisches Privileg ausweisen (vgl. Pooth 2019).

5. Prüfung und pädagogische Verantwortung

Mit Ingenkamp ist der ‹pädagogische Auftrag› der diagnostisch Handelnden bereits angeklungen: Als ‹Anwalt des Kindes› (Ingenkamp 1971b, 27) können diese nur auftreten, wenn sie nicht zugunsten der Illusion, dass die Daten für sich selbst sprechen (vgl. kritisch dazu Kitchin 2014, 5), aus der Prüfung ausgelagert werden. Mit Blick auf die Differenz von Gerechtigkeit und Objektivität plädiert Ingenkamp mitnichten für einen standardisierten Blick ‹out of nowhere›, sondern für ein professionelles Urteil der Lehrenden, das der Machtblindheit standardisierter Testverfahren etwas entgegensetzen kann.

Die Möglichkeit zu diesem Einspruch verweist nicht nur darauf, dass das Prüfungsgeschehen als ein relationales Verhältnis zwischen verschiedenen Akteurinnen und Akteuren zu verstehen ist. Sie zeigt zugleich an, dass die Prüfung als ein Verantwortungsverhältnis gefasst werden muss, insofern man hier ‹für sich, für einen anderen oder für anderes eine bestimmte Verpflichtung eingeht› (Meyer-Drawe 1992, 14). In bildungstheoretischer Perspektive basiert diese Verpflichtung auf der ‹Beziehung zur Singularität und radikalen Andersheit des Anderen› (Wimmer 1996, 441) und wird damit in den Kontext des Gerechtigkeitsbegriffs gerückt. In einem solchen Verhältnis stehen die pädagogisch Handelnden in der Verantwortung: Sie müssen antworten, ‹ohne vorher wissen zu können› (ebd., 442).

Im Folgenden kann keine begriffliche Explikation des Verantwortungsbegriffes vorgenommen werden (vgl. hierfür Kuhlmann 2021), es soll aber mit Fenwick und Edwards (2016) danach gefragt werden, wie sich das pädagogische Verantwortungsverhältnis, das mit der Prüfung evoziert wird, mit dem Einsatz digitaler Tools verändert. Gerade der bildungstheoretische Fokus auf ‹das Antworten› bildet hier einen Ausgangspunkt, von dem her ersichtlich wird, dass sich mit LA Transformationen dahingehend vollziehen, ‹wie eine Antwort hervorgerufen wird, von wem oder wovon, welche Formen sie annimmt, und worin ihre Konsequenzen bestehen› (ebd., 8).

In geisteswissenschaftlicher Tradition hat Loch (1963) die Unverfügbarkeit und Fremdheit der Anderen als Ausgangspunkt für das pädagogische Verantwortungsverhältnis innerhalb der Prüfung thematisiert: Die Prüfung wird dadurch zu einem Ort, an dem nicht nur die Geprüften, sondern auch die Prüfenden selbst bzw. ihr pädagogisches Handeln infrage gestellt werden (vgl. ebd., 181). Unter Rückgriff auf Herbarts Bildsamkeitsbegriff konturiert Loch die Prüfung als unmittelbar an das Erziehungsgeschehen selbst gebunden – wenn man so will, als Teil einer ‹Antwort› auf die Bildsamkeit der zu Erziehenden.

Eine solche (Ver-)antwortung zielt dann nicht einfach auf das ‹Ergebnis› der Prüfung. Die Prüfung kann zwar die konkrete Leistung der erziehenden Person infrage stellen; gleichwohl können Erziehende ‹die Rechtfertigung [ihrer; SaS] Erziehung nicht allein vom Ausgang einer Prüfung abhängig machen [...]. Gleich wie die Prüfungen ausgehen: die Aufgabe der Erziehung bleibt davon prinzipiell unabhängig› (ebd.,

182). Die Prüfung ist damit gewissermassen Anlass einer Verhandlung und Auseinandersetzung über das erzieherische Handeln. Sie ist ein möglicher Ausgangspunkt, von dem her das pädagogische Handeln selbst befragt werden kann.

Unmittelbar einsichtig dürfte sein, dass es mir an der Stelle keineswegs um eine Idealisierung der geisteswissenschaftlichen Pädagogik geht. Vielmehr bietet der Ausgangspunkt, die Prüfung als ein Antwortgeschehen zu begreifen, die Möglichkeit, den Verantwortungsbegriff als ein analytisch-kritisches Instrument in Stellung zu bringen. Mit diesem rückt nämlich die bildungstheoretische Frage «Wie dem Anderen gerecht werden?» (Wimmer 2007) als eine Aufgabe der pädagogischen Beziehung in den Vordergrund. In der Prüfung sind alle Beteiligten der radikalen Fremdheit der Anderen und deren Ansprüchen ausgesetzt: «Das Antworten *lässt sich ein* auf Anforderungen, Anforderungen, Herausforderungen und ähnliches» (Waldenfels 1992, 141, Herv. SaS) und formuliert auf diese Weise einen radikalen Anspruch, auch in Prüfungszusammenhängen.

Dieser radikale Anspruch ist nun aber gerade nicht Ausgangspunkt der digitalen Lerntools: Indem sie, wie oben beschrieben, «Muster der Vergangenheit» identifizieren bzw. wiederholen und «Erwartungskreisläufe» (Fenwick und Edwards 2016, 15) herstellen, «antworten» sie eben nicht. Ohne mindestens das zusätzliche pädagogische Urteil, das mit den rückgekoppelten Daten «etwas anfangen muss», kann wohl von einem «Einlassen» auf den Anspruch der Anderen nur sehr begrenzt die Rede sein. Indem der «Anspruch der Anderen» in eine digitale Form gebracht wird, werden diese qua Technologie zu identifizierten Anderen. Auch wenn mit Wimmer (2007) davon auszugehen ist, dass das pädagogische Handeln *stets* daran scheitert, den Anderen gerecht zu werden, ist die algorithmische «Erfassung» nicht gleichzusetzen mit der Vorstellung eines «Einlassens» oder «Antwortens» auf das Ungewisse. Denn, die digitalen Tools sind darauf ausgerichtet, zielgenaue Ergebnisse zu formulieren, um daran anschliessend «angemessen» zu intervenieren. Die Intervention basiert auf der Berechenbarkeit des Verhältnisses aktueller und künftiger Prüfungs- oder Lernergebnisse und ermöglicht eine Vereindeutigung dessen, was jetzt zu tun wäre. Weil ihr Ausgangspunkt eben nicht die prinzipiell unzugängliche Andersheit, sondern die Identifikation ist, versprechen LA die Verfügung über das Fremde: Mit Waldenfels liesse sich davon sprechen, dass die Prüfung sich auf diese Weise einem Zustand annähert, «wo es nichts weiter mehr gäbe als mehr oder weniger konditionierte und programmierte Antworten, Programmverbesserungen eingeschlossen» (Waldenfels 1992, 141).

Wenn also pädagogische Verantwortung in der Prüfung einerseits an die Möglichkeiten und Grenzen der Daten und Tools sowie andererseits an die Lernenden selbst abgegeben wird, wird das pädagogische Verhältnis zunehmend von der Vorstellung einer technischen Beherrschbarkeit durchsetzt. Eine Konsequenz, die daraus erwachsen könnte, ist die Durchkreuzung der Vielgestaltigkeit und Heterogenität pädagogischer «Antwortgeschehen» und damit die Infragestellung gegenseitiger Verpflichtungen.

Einen Hinweis darauf bietet Thompson, die in einem Essay (2013) die Transformationen im pädagogischen Verhältnis von Lehrenden und Studierenden durch die umfassende Einführung von Evaluationsinstrumenten in den Seminaralltag beschrieben hat. Bezugnehmend auf Roland Barthes' Text «An das Seminar» (vgl. 2010) verweist Thompson auf die Problematik, das pädagogische Verhältnis zwischen Studierenden und Lehrenden in eine Input-Output-Relation der «Wissensübertragung» und «Wissensaufnahme» zu reduzieren. Mit Barthes liesse sich die Seminarsituation stattdessen als ein «being-with» (Thompson 2013, 293) beschreiben – als eine Beziehung unter Lernenden, die an der Herausbildung eines gemeinsamen Gegenstands teilhaben. Dieser geteilte Raum, in dem gemeinsam eine Sache von Belang verhandelt wird und der dadurch entsteht, dass man sich ihm und den anderen gegenüber verpflichtet fühlt, steht nicht nur im Kontrast zur «language of evaluation» (ebd., 294), wie sie Thompson beschreibt, sondern ebenso zur hier aufgezeigten «language of LA».

Auch die Diskussionen zur Dataveillance bieten zum Problem der Verantwortung wichtige Impulse: Ausgehend von der Frage nach einer «Ethik des Überwachens» diskutierten Bauman und Lyon (2013), dass das Handeln qua Technologie und Datafizierung eine Unterbrechung der Unmittelbarkeit und somit eine «Distanzierung des Akteurs von den Folgen seines Handelns» ermöglicht (ebd., 165) – was unmittelbar einsichtig ist am Beispiel des Einsatzes von Drohnen zur Kriegsführung. Bauman und Lyon sprechen hier von der Gefahr einer «Exklusion des zerteilten und neuarrangierten Objekts aus der Klasse der moralisch bedeutsamen Entitäten und dem Universum ethischer Verpflichtungen» (ebd., 168). Es ist dieser Raum einer gemeinsamen ethischen Verpflichtung pädagogischer Art, der durch LA auf dem Spiel zu stehen scheint.

Man könnte nun wohl von einem neuen Unbehagen sprechen, was sich auf die «Umweghaftigkeit» des pädagogischen Verhältnisses bezüglich der Prüfung als eines datafizierten Regulierungsinstruments bezieht und möglicherweise auf eine «decomposition» of the pedagogical relation» (Thompson 2013, 284) hinausläuft: Die «Aufforderung zur Selbsttätigkeit», wie sie Benner (2001) im Rekurs auf Herbarts Bildsamsamkeitsbegriff formuliert hat und wie sie jeder Prüfung inhärent ist, wäre dann nicht länger der pädagogischen Interaktion verpflichtet. Wenn sich die Reaktionen auf die Anstrengungen der Einzelnen vollständig über die Entscheidungs- und Auswertungsprozesse digitaler Tools strukturieren, steht infrage, inwiefern hier überhaupt noch von einer pädagogischen «Verpflichtung» gesprochen werden kann.

In diesem Sinne, so das abschliessende Plädoyer, ist grosses Misstrauen angebracht, wenn die Prüfung als «das Andere der Pädagogik» missverstanden (Breidenstein und Thompson 2014, 106; Ricken und Reh 2017, 248) und aus dem Raum des Pädagogischen ausgelagert wird.

Literatur

- Bächle, Thomas Christian. 2016. *Digitales Wissen, Daten und Überwachung zur Einführung*. Hamburg: Junius.
- Baker, Ryan Shaun, und Paul Salvador Inventado. 2014. «Educational Data Mining and Learning Analytics». In *Learning Analytics. From Research to Practice*, herausgegeben von Johann Ari Larusson, und Brandon White, 61–75. New York/Heidelberg/Dordrecht/London: Springer. <https://doi.org/10.1007/978-1-4614-3305-7>.
- Barthes, Roland. 2010. «An das Seminar». In *Was ist Universität? Texte und Positionen zu einer Idee*, herausgegeben von Johanna-Charlotte Horst, Johannes Kagerer, Regina Karl, Vera Kaulbarsch, Johannes Kleinbeck, Elias Kreuzmair, Anouk Luhn, Adrian Renner, Anna Sailer, Tillmann Severin, Hanna Sohns, und Jennifer Sréter (Unbedingte Universitäten 2010), 17–26. Zürich: diaphanes.
- Bauman, Zygmunt, und David Lyon. 2013. *Daten, Drohnen, Disziplin. Ein Gespräch über flüchtige Überwachung*. Berlin: Suhrkamp.
- Benner, Dietrich. 2001. *Allgemeine Pädagogik. Eine systematisch-problemgeschichtliche Einführung in die Grundstruktur pädagogischen Denken und Handelns*. Weinheim/München: Juventa.
- Bohl, Thorsten. 2009. *Prüfen und Bewerten im Offenen Unterricht*. Weinheim/Basel: Beltz.
- Breidenstein, Georg. 2012. *Zeugnisnotenbesprechung. Zur Analyse der Praxis schulischer Leistungsbewertung*. Opladen/Berlin/Toronto: Barbara Budrich.
- Breidenstein, Georg. 2018. «Das Theorem der <Selektionsfunktion der Schule> und die Praxis der Leistungsbewertung». In *Leistung als Paradigma. Zur Entstehung und Transformation eines pädagogischen Konzepts*, herausgegeben von Sabine Reh, und Norbert Ricken, 307–327. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-15799-9>.
- Breidenstein, Georg, und Christiane Thompson. 2014. «Schulische Leistungsbewertung als Praxis der Subjektivierung». In *Interferenzen. Perspektiven kulturwissenschaftlicher Bildungsforschung*, herausgegeben von Christiane Thompson, Kerstin Jergus, und Georg Breidenstein, 89–109. Weilerswist: Velbrück.
- Buck, Marc Fabian. 2017. «Gamification von Unterricht als Destruktion von Schule und Lehrberuf». *Vierteljahrsschrift für wissenschaftliche Pädagogik* 93 (2): 268–282. <https://doi.org/10.1163/25890581-093-02-90000005>.
- Fenwick, Tara, und Richard Edwards. 2016. «Die Auswirkungen digitaler Technologien auf professionelle Verantwortung und Ausbildung». *Berliner Debatte Initial* 27 (1): 6–21.
- Foucault, Michel. 1994. *Überwachen und Strafen. Die Geburt des Gefängnisses*. Frankfurt a. M.: Suhrkamp.
- Foucault, Michel. 2015. *Die Wahrheit und die juristischen Formen*. Frankfurt a. M.: Suhrkamp.
- Greller, Wolfgang, und Hendrik Drachslers. 2012. «Translating Learning into Numbers: A Generic Framework for Learning Analytics». *Educational Technology & Society* 15 (3): 42–57. <https://www.jstor.org/stable/jeductechsoci.15.3.42>.

- Häcker, Thomas. 2011. «Portfolioarbeit – Ein Konzept zur Wiedergewinnung der Leistungsbeurteilung für die pädagogische Aufgabe der Schule». In *Diagnose und Beurteilung von Schülerleistungen – Grundlagen und Reformansätze*, herausgegeben von Werner Sacher, und Felix Winter, 217–230. Hohengehren: Schneider.
- Haraway, Donna. 1988. «Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective». *Feminist Studies* 14 (3): 575–599. <http://www.jstor.org/stable/3178066>.
- Hartong, Sigrid. 2019a. «Bildung 4.0? Kritische Überlegungen zur Digitalisierung von Bildung als erziehungswissenschaftliches Forschungsfeld». *Zeitschrift für Pädagogik* 65 (3): 424–443. <https://doi.org/10.3262/ZP1903424>.
- Hartong, Sigrid. 2019b. «Learning Analytics und Big Data in der Bildung: Zur notwendigen Entwicklung eines datenpolitischen Alternativprogramms». *GEW Broschüre Bildung in der digitalen Welt*, 7–27. http://www.aufwach-s-en.de/wp-content/uploads/2019/12/hartong_learning-analytics_2019_web.pdf.
- Hartong, Sigrid. 2020. «Zum Optimierungsdrang des Bildungsmonitorings». *Zeitschrift für Pädagogik* 66 (1): 64–71. <https://doi.org/10.3262/ZP2001064>.
- Heintz, Bettina. 2007. «Zahlen, Wissen, Objektivität: Wissenschaftssoziologische Perspektiven». In *Zahlenwerk. Kalkulation, Organisation und Gesellschaft*, herausgegeben von Andrea Mennicken, und Hendrik Vollmer, 65–85. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-531-90449-8>.
- Höhne, Thomas, Martin Karcher, und Christin Voss. 2020. «Wolkige Verheißungen. Die Schul-Cloud als Mittel der Technologisierung von Schule und Lernen». *Zeitschrift für Pädagogik* 66 (3): 324–340. <https://doi.org/10.3262/ZP2003324>.
- Hörtnagl, Jakob. 2019. «Subjektivierung in datafizierten Gesellschaften – Dividualisierung als Perspektive auf kommunikative Aushandlungsprozesse in datengetriebenen Zeiten». In *Das vergessene Subjekt, Medien. Kultur. Kommunikation*, herausgegeben von Peter Gentzel, Friedrich Krotz, Jeffrey Wimmer, und Rainer Winter, 135–156. Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-23936-7_7.
- Ifenthaler, Dirk, und Hendrik Drachsler. 2020. «Learning Analytics. Spezielle Forschungsmethoden in der Bildungstechnologie». In *Handbuch Bildungstechnologie. Konzeption und Einsatz digitaler Lernumgebungen*, herausgegeben von Helmut Niegemann, und Armin Weinberger, 515–534. Berlin: Springer. <https://doi.org/10.1007/978-3-662-54368-9>.
- Ifenthaler, Dirk, und Clara Schumacher. 2016. «Learning Analytics im Hochschulkontext». *WiSt – Wirtschaftswissenschaftliches Studium* 45 (4): 176–181. <https://doi.org/10.15358/0340-1650-2016-4-176>.
- Ingenkamp, Karlheinz. 1968. «Möglichkeiten und Grenzen des Lehrerurteils und der Schultests». In *Begabung und Lernen*, herausgegeben von Heinrich Roth, 407–431. Stuttgart: Klett.
- Ingenkamp, Karlheinz. 1971a. *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte*. Weinheim/Berlin/Basel: Beltz.

- Ingenkamp, Karlheinz. 1971b. «Einführung in den Themenkreis». In *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte*, herausgegeben von Karlheinz Ingenkamp, 9–33. Weinheim/Berlin/Basel: Beltz.
- Ingenkamp, Karlheinz. 1971c. «Sind Zensuren aus verschiedenen Klassen vergleichbar?». In *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte*, herausgegeben von Karlheinz Ingenkamp, 156–163. Weinheim/Berlin/Basel: Beltz.
- Ingenkamp, Karlheinz, und Urban Lissmann. 2008. *Lehrbuch der Pädagogischen Diagnostik*. Weinheim/Basel: Beltz.
- Jürgens, Eiko. 2010. *Leistung und Beurteilung in der Schule. Eine Einführung in Leistungs- und Bewertungsfragen aus pädagogischer Sicht*. Sankt Augustin: Academia.
- Jürgens, Eiko, und Urban Lissmann. 2015. *Pädagogische Diagnostik. Grundlagen und Methoden der Leistungsbeurteilung in der Schule*. Weinheim/Basel: Beltz.
- Karcher, Martin. 2020. «Die (kybernetische) Bändigung des Zufalls. Dataveillance und Learning Analytics als Herausforderung erziehungswissenschaftlicher Reflexion: Versuch einer Technikfolgenabschätzung». In *Neue Steuerung – Renaissance der Kybernetik?*, herausgegeben von Detlef Fickermann, Veronika Manitus, und Martin Karcher, 151–168. Münster: Waxmann. <https://doi.org/10.31244/9783830991618>.
- Kitchin, Rob. 2014. «Big Data, New Epistemologies and Paradigm Shifts». *Big Data and Society* 1 (1): 1–12. <https://doi.org/10.1177/2053951714528481>.
- Kuhlmann, Nele. 2021. «On the Power of the Concept of «Responsibility»». *Educational Theory*. i. V.
- Langfeldt, Hans-Peter. 1984. «Die klassische Testtheorie als Grundlage normorientierter (standardisierter) Schulleistungstests». In *Leistungsdiagnostik in der Schule*, herausgegeben von Kurt A. Heller, 65–98. Bern/Stuttgart/Toronto: Hans Huber.
- Larusson, Johann Ari, und Brandon White. 2014. «Introduction». In *Learning Analytics. From Research to Practice*, herausgegeben von Johann Ari Larusson, und Brandon White, 1–14. New York/Heidelberg/Dordrecht/London: Springer. <https://doi.org/10.1007/978-1-4614-3305-7>.
- Loch, Werner. 1963. «Historische Vorbetrachtungen zu einer Pädagogik der Prüfung». *Bildung und Erziehung* 16: 180–193.
- Meyer-Drawe, Käte. 1992. «Nachdenken über Verantwortung». *Verantwortung. Friedrich Jahresheft* (X): 14–16.
- Meyer, Markus, und Christian Jansen. 2016. *Schulische Diagnostik. Ein Studien- und Arbeitsbuch*. Bad Heilbrunn: Klinkhardt. <https://doi.org/10.36198/9783838545325>.
- O’Neil, Cathy. 2018. *Angriff der Algorithmen. Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden*. Bonn: bpb.
- Paradies, Liane, Franz Wester, und Johannes Greving. 2009. *Leistungsmessung und -bewertung*. Berlin: Cornelsen.

- Pardo, Abelardo. 2014. «Designing Learning Analytics Experiences». In *Learning Analytics. From Research to Practice*, herausgegeben von Johann Ari Larusson, und Brandon White, 15–38. New York/Heidelberg/Dordrecht/London: Springer. <https://doi.org/10.1007/978-1-4614-3305-7>.
- Pooth, Laura. 2019. «Digitalpakt Schule – der falsche Ansatz.» *Neue Gesellschaft. Frankfurter Hefte*. 05.07.2019. <https://www.frankfurter-hefte.de/artikel/digitalpakt-schule-der-falsche-ansatz-2767/>.
- Rabenstein, Kerstin, Till-Sebastian Idel, und Norbert Ricken. 2015. «Zur Verschiebung von Leistung im individualisierten Unterricht. Empirische und theoretische Befunde zur schulischen Leitdifferenz». In *Heterogenitätsforschung. Empirische und theoretische Perspektiven*, herausgegeben von Jürgen Budde, Nina Blasse, Andrea Bossen, und Georg Rißler, 241–258. Weinheim und Basel: Beltz Juventa.
- Ricken, Norbert, und Sabine Reh. 2017. «Prüfungen – Systematische Perspektiven der Geschichte einer pädagogischen Praxis. Einführung in den Thementeil». *Zeitschrift für Pädagogik* 63 (3): 247–258. <https://doi.org/10.3262/ZP1703247>.
- Ruhloff, Jörg. 2009. «Die Tradition humanistischer Bildung seit der Renaissance und die gegenwärtige Neudefinition von ‹Bildung›». In «*Was den Menschen eigentlich zum Menschen macht...*». *Klassische Texte einer Philosophie der Bildung*, herausgegeben von Hans-Ulrich Lessing, und Volker Steenblock, 183–202. Freiburg im Breisgau: Karl Alber.
- Sacher, Werner. 2014. *Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe*. Bad Heilbronn: Julius Klinkhardt. <https://doi.org/10.35468/9783781553644>.
- Schröder, Sabrina, und Christiane Thompson. 2015. «A matter of exposition: examination and education». *Ethics and Education* 10 (2): 152–162. <https://doi.org/10.1080/17449642.2015.1039273>.
- Schröder, Sabrina. 2021. «Die Vermessung des Lernens. Objektivierung und Subjektivierung in digitalen Lernplattformen». *Pädagogische Korrespondenz. Zeitschrift für kritische Zeitdiagnostik in Pädagogik und Gesellschaft* 63: 85–110.
- Selwyn, Neil. 2019. «What’s the Problem with Learning Analytics?» *Journal of Learning Analytics* 6 (3): 11–19. <https://doi.org/10.18608/jla.2019.63.3>.
- Someh, Ida, Michael Davern, Christoph F. Breidbach, und Graeme Shanks. 2019. «Ethical Issues in Big Data Analytics: A Stakeholder Perspective.» *Communications of the Association for Information Systems* 44: 718–747. <https://doi.org/10.17705/1CAIS.04434>.
- Terhart, Ewald. 2009. «Schüler beurteilen – Zensuren geben. Wie Lehrerinnen und Lehrer mit einem leidigen, aber unausweichlichen Element ihres Berufsalltags umgehen». In *Leistung ermitteln und bewerten*, herausgegeben von Silvia-Iris Beutel, und Witlof Vollstädt, 39–50. Hamburg: Bergmann + Helbig.
- Thompson, Christiane. 2013. «Evaluations and the Forgetfulness of Pedagogical Relations: Remarks on Educational Authority». *Educational Theory* 63 (3): 283–298. <https://doi.org/10.1111/edth.12024>.

- Tillmann, Klaus-Jürgen. 2008. «Viel Selektion – wenig Leistung: Der PISA-Blick auf Erfolg und Scheitern in deutschen Schulen». In *Von der Delegation zur Kooperation. Bildung in Familie, Schule, Kinder- und Jugendhilfe*, herausgegeben von Karin Böllert, 47–66. Wiesbaden: VS. <https://doi.org/10.1007/978-3-531-90846-5>.
- Waldenfels, Bernhard. 1992. «Antwort und Verantwortung». *Verantwortung. Friedrich Jahresheft (X)*: 139–141.
- Williamson, Ben. 2017. «Decoding ClassDojo: psycho-policy, social-emotional learning, and persuasive educational technologies». *Learning, Media and Technology*: 1–17. <https://doi.org/10.1080/17439884.2017.1278020>.
- Wimmer, Michael. 1996. «Zerfall des Allgemeinen – Wiederkehr des Singulären. Pädagogische Professionalität und der Wert des Wissens». In *Pädagogische Professionalität. Untersuchungen zum Typus pädagogischen Handelns*, herausgegeben von Arno Combe, und Werner Helsper, 404–447. Frankfurt a. M.: Suhrkamp.
- Wimmer, Michael. 2007. «Wie dem Anderen gerecht werden? Herausforderungen für Denken, Wissen und Handeln». In *Kindliche Fremdheit und pädagogische Gerechtigkeit*, herausgegeben von Alfred Schäfer, 155–184. Paderborn/München/Wien/Zürich: Schöningh.
- ZDF heute-show (@heuteshow). 2021. «Viele Schüler in Deutschland bekommen heute ihr #Zeugnis. Die Note in Sport ist dieses Mal noch ein bisschen willkürlicher als in den Jahren zuvor». 29.01.2021. <https://twitter.com/heuteshow/status/1355122640546127872>.