



KLEINE KIELER  
BEITRÄGE ZU

# Künstlicher Intelligenz



BAND 7



| VK:KIWA

KLEINE KIELER BEITRÄGE ZU

# **Künstlicher Intelligenz**

Pädagogische und fachdidaktische Perspektiven  
auf technologische Entwicklungen

BAND 7



Herausgeben von

Kirsten Schindler, Nicolaus Wilder, Virtuelles Kompetenzzentrum:

Künstliche Intelligenz in Bildung, Wissenschaft & Arbeitswelt (VK:KIWA)

Marc Fabian Buck

# **Prompt zurückgewiesen, Prüfung bestanden?**


Pädagogische Prüfungen in Zeiten  
generativer KI

## KLEINE KIELER BEITRÄGE ZU KÜNSTLICHER INTELLIGENZ | BAND 7

eISSN: 3053-4488

### HERAUSGEGEBEN VON


**Prof.in Dr. Kirsten Schindler** , Bergische Universität Wuppertal

**Dr. Nicolaus Wilder** , Christian-Albrechts-Universität zu Kiel

**Virtuelles Kompetenzzentrum: Künstliche Intelligenz in Bildung,  
Wissenschaft & Arbeitswelt**

<https://www.vkkiwa.de/>

### AUTOR

**Marc Fabian Buck** , Dr. phil.,

absolvierte sein Grundstudium der Pädagogik und Soziologie an der CAU und wechselte dann für das Hauptstudium und die Promotion im Fach Erziehungswissenschaft an die HU Berlin. Er vertritt derzeit die Professur für Allgemeine Pädagogik an der Universität Koblenz. Zu seinen zentralen Lehr- und Forschungsinteressen zählen Transformationen der Gesellschaft und Bildung unter besonderer Berücksichtigung von Ökonomisierungs- und Digitalisierungsprozessen. In den vergangenen Jahren hat er zu diversen Phänomenen und Problemen der Digitalisierung der Bildung publiziert und im vergangenen Jahr eine Einführung in die Digitalisierung der Bildung (Buck, 2025b) veröffentlicht. Weitere Informationen zu Publikationen, Vorträgen und Forschungsprojekten finden sich auf seiner persönlichen Homepage: <https://www.marcfabianbuck.net>.

---

### Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie. Detaillierte bibliographische Daten sind über <https://dnb.d-nb.de> abrufbar.

### Open Access



Das Werk ist unter der Creative-Commons-Lizenz Namensnennung-4.0 International veröffentlicht. Den Vertragstext finden Sie unter: <https://creativecommons.org/licenses/by/4.0/>.

Bitte beachten Sie, dass einzelne, entsprechend gekennzeichnete Teile des Werks von der genannten Lizenz ausgenommen sein bzw. anderen urheberrechtlichen Bedingungen unterliegen können.

Das Werk ist auf dem Open-Access-Publikationsserver MACAU der Universitätsbibliothek Kiel (<https://macau.uni-kiel.de>) frei verfügbar: <https://doi.org/10.38071/2026-00523-2>.

### Kiel, 2026

Selbstverlag des Virtuellen Kompetenzzentrums für Künstliche Intelligenz und Wissenschaftliches Arbeiten (VK:KIWA) im Hosting-Service der UB Kiel.

**Titelbild:** Erstellt mit Recraft (<https://www.recraft.ai/>).

## Geleitwort zur Einordnung

Kaum ein Thema im Feld von Künstlicher Intelligenz und Bildungssystem wird seit der Verfügbarmachung von ChatGPT so kontrovers – und mitunter so hilflos – verhandelt wie die Frage nach *zeitgemäßen* Prüfungen. Wer die gegenwärtige Debatte verfolgt, erkennt: Sie wird weithin getragen von soziologischen, psychologischen und ökonomischen Perspektiven, die zweifellos Erklärungs- und Sensibilisierungskraft besitzen, aber im Kern oft an dem vorbeireden, was Prüfen im Bildungskontext eigentlich ist: eine spezifische Form pädagogischen Handelns, die sich nicht in Messlogiken, Steuerungsphantasien oder Effizienz Narrativen erschöpft.

Der Beitrag von Marc Fabian Buck setzt genau hier an – mit einer Klarheit, Präzision und Strenge, die in dieser Debattenlage selten geworden ist. Er markiert präzise das Defizit, das viele Reform-Rufe begleitet: die systematische Verfehlung des Pädagogischen. Buck insistiert auf einer begründeten Rückgewinnung der Fragehoheit: Was heißt es, pädagogisch zu prüfen – und wer oder was wird dabei eigentlich geprüft? Indem er Prüfungen historisch-systematisch als Handlungen rekonstruiert und die „theoretische Leerstelle“ des Prüfens im erziehungswissenschaftlichen Kanon ausleuchtet, erinnert er an Vergessenes, das sich im Lärm des Gegenwartsdiskurses erstaunlich gut überhören ließ.

Die Stärke dieses Textes liegt nicht nur in seiner Diagnose, sondern im entschiedenen Perspektivwechsel, den er vornimmt. Buck liest die technische Verfügbarkeit von LLMs nicht primär als Bedrohung, die schriftliche Prüfungen obsolet mache, sondern als Anlass, die pädagogische Eigenlogik des Prüfens wieder sichtbar zu machen. Damit wird KI nicht zum Schicksal, dem sich Prüfungskulturen zu fügen hätten, sondern zum Stresstest für jene Selbstverständlichkeiten, die ohnehin schon lange erodierten: die stillschweigende Gleichsetzung von Prüfung mit Outputkontrolle, die Überhöhung von Standardisierung als Gerechtigkeitsersatz, die Delegation pädagogischer Urteilsfragen an vermeintlich objektive Verfahren.

Besonders überzeugend ist, wie Buck die dominanten Deutungsangebote der Gegenwart – Selektion/Allokation, Metakognition, datenförmige

Dauerbeobachtung, EdTech-Interessen – nicht einfach negiert, sondern als parapädagogische Überformungen identifiziert: als Logiken, die sich in Bildungsinstitutionen einschleichen, weil sie sich als alternativlos, modern, effizient ausgeben. In dieser Perspektive wird sichtbar, dass *zeitgemäß* in vielen Reformpapieren weniger ein pädagogisches Urteil als ein politisch-ökonomischer Code ist: Es meint Anschlussfähigkeit an Steuerbarkeit, Vergleichbarkeit und Skalierbarkeit. Buck arbeitet heraus, dass gerade diese Trias – Standardisierung, Ökonomisierung, technologische Überformung – nicht zufällig miteinander harmoniert. Und er benennt, worum es dabei geht: um Deutungshoheit über Bildungszwecke, um die Verschiebung von pädagogischem Urteil hin zu administrativ-juristisch kontrollierbaren Prozeduren, um die stille Ersetzung von Bildung durch Zertifizierung.

Dass der Text dabei nicht in Kulturpessimismus verfällt, ist eine zweite, nicht minder wichtige Qualität. Buck formuliert einen Gegenentwurf, der sich in seiner vermeintlichen Rückwärtsgewandtheit hochgradig erfrischend und hoffnungsvoll liest: die Re-Fokussierung auf den individuellen Menschen mit seiner je eigenen Lebensgeschichte, Originalität und der Fähigkeit des Denkenkönnens. Seine bildungstheoretisch informierte Rekonstruktion der Prüfung der Originalität verschiebt den Maßstab von vermeintlich objektiver Outputbewertung hin zu Fairness im Sinne einer kontextsensiblen Urteilsbildung. Das wirkt radikal realistisch: LLMs operieren im Modus von Text ohne Kontext; pädagogisches Prüfen, sofern es diesen Namen verdient, ist unauflöslich Kontextarbeit. In einer Zeit, in der Objektivität gern als moralisch überlegene Abkürzung gehandelt wird, insistiert Buck darauf, dass gerade im pädagogischen Feld Fairness der anspruchsvollere und demokratischere Maßstab ist – weil er die Person nicht in einem standardisierten Idealbild verschwinden lässt.

Damit wird auch die politische Dimension des Textes deutlich, ohne dass er sich in Parolen erschöpft. Bucks Perspektive auf Prüfungen als pädagogische Handlungen positioniert sich klar gegen Standardisierungssehnsüchte, gegen eine Ökonomisierung im Gewand von Effizienz und Employability, und gegen tech-oligarchisches Wunschdenken.

*Nicolaus Wilder & Kirsten Schindler*

# Inhalt

<b>Geleitwort zur Einordnung</b>	<b>5</b>
<b>1 Einleitung und Problemaufriss</b>	<b>8</b>
<b>2 Prüfungen als pädagogische Handlungen</b>	<b>15</b>
2.1 Loch: Prüfungen als Lernaufgaben	18
2.2 Ruhloff: Drei Formen von Prüfungen	21
<b>3 Prüfungen post KI im Nexus von Digitalisierung und Ökonomisierung</b>	<b>26</b>
<b>4 Die Prüfung der Originalität als pädagogisches und demokratisches Projekt</b>	<b>35</b>
<b>Literaturverzeichnis</b>	<b>44</b>

# 1 Einleitung und Problemaufriss

Am 21. November 2025 forderte Papst Leo XIV. im Rahmen der US-amerikanischen *National Catholic Youth Conference* in Indianapolis die dort Versammelten dazu auf, im (hoch-)schulischen Umgang mit Künstlicher Intelligenz Vorsicht walten zu lassen. Nötig sei »serious reflection and ongoing discussion on the inherently ethical dimension of AI, as well as its responsible governance.« (BBC, 2025) Das Magazin *Vulture* titelte entsprechend »Even God Is Worried About ChatGPT« (*Vulture*, 2025).<sup>1</sup>

Es darf Zweifel am Erfolg des moralischen Appells des katholischen Kirchenoberhauptes angemeldet werden, wofür zwei Gründe zu nennen sind: Erstens die Geschwindigkeit, mit der sich Large Language Models (LLMs) und deren popularisierte Anwendungen wie ChatGPT, Gemini, Copilot entwickeln; zweitens der Niedergang der Breitenwirkung moralischer Belehrung durch die Kirche. Die Normalisierung von ChatGPT und Co. in der Gesellschaft scheint zu großen Teilen vollzogen und nur schwer aufzuhalten.<sup>2</sup>

Welche Effekte generative KI auf Lern- und Bildungserfahrungen Heranwachsender, auf soziale Verhältnisse in pädagogischen Einrichtungen und das »soziale Band« (Bedorf & Herrmann, 2016) der Gesellschaft, auf das Lehren als Praxis und pädagogische Professionen hat, ist bisher weder in der Breite empirisch erforscht noch im Theoriebestand berücksichtigt worden

- 
- 1 Der vorliegende Text geht zurück auf einen Vortrag, den ich unter dem Titel »Die Funktion der Prüfung und KI. Eine Neubesinnung« im Rahmen des Projektes »It's all about the prompts« am 28.11.2025 an der Universität Bern habe halten dürfen. Für zahlreiche konstruktive Hinweise zum Manuskript danke ich Gaudenz Welti, Robin Gerstgrasser, Stefan Emmenegger, Thomas Mikhail, Ann-Kathrin Hoffmann, Anne Pesch, Henning Pätzold und Jirko Piberger ausdrücklich. Ebenso möchte ich Kirsten Schindler und Nicolaus Wilder für das gründliche Review, produktive Verbesserungsvorschläge und die Aufnahme in die *Kleinen Kieler Beiträge zu Künstlicher Intelligenz* herzlich danken.
  - 2 Die Gleichsetzung von KI mit LLM-Anwendungen ist genauso technisch falsch, wie sie sich im allgemeinen Sprachgebrauch dennoch durchgesetzt hat. Im Sinne der besseren Lesbarkeit folgt dieser Beitrag dem berühmten Wittgenstein'schen Diktum: »Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache« (Wittgenstein, 2025, S. 80).

(vgl. Buck, 2025b, S. 127ff.) – das ist einigermaßen erstaunlich, da zum Zeitpunkt der Abfassung dieses Beitrags die Veröffentlichung von ChatGPT 2.0 im November 2022 bereits über drei Jahre und Versionsnummern zurückliegt.<sup>3</sup>

Diese Diagnose gilt auch für die Veränderung von Prüfungen in pädagogischen Kontexten (im Folgenden: pädagogische Prüfungen) unter dem Einfluss voranschreitender KI-Entwicklung, die sich von allgemeinbildenden über berufliche Schulen und Weiterbildungseinrichtungen bis zur Universität über sämtliche Bildungsbereiche erstreckt und als solche ausgewiesen ist. Die damit verbundene Vermutung lautet, dass primär schriftliche Prüfungen durch die fortschreitend leistungsfähiger werdenden Sprachmodelle obsolet werden, da jene in absehbarer Zukunft Texte zu produzieren imstande sind, die es Prüfenden verunmöglicht, die damit verbundene Leistung als persönliche zu erkennen und zu verbuchen. Um es vorwegzunehmen: Diese Befürchtung wird hier nicht geteilt. Im Gegenteil: Die stete Verbesserung von LLMs fordert zu einem Nachdenken darüber auf, *was* und *wer* in pädagogischen Prüfungen *wie* geprüft wird.

Von Prüfungen zu differenzieren sind ebenfalls in pädagogischen Kontexten vorfindliche Praktiken der Beobachtung und Diagnostik, in denen Beobachtete und Diagnostizierte gegebenenfalls gar nicht wissen, dass sie, ihr Verhalten oder ihre Konstitution Gegenstand einer möglicherweise folgenschweren Beurteilung sind. Auch sind sie von außerpädagogischen Prüfungen unseres Alltags zu differenzieren, wie sie etwa im Sinne der Verkehrssicherheit, des Brandschutzes, der Gesundheitsvorsorge usw. institutionalisiert sind. Dazu zählen etwa die regelmäßigen Hauptuntersuchungen von Kraftfahrzeugen (›TÜV-Prüfung‹) und die Feuerstättenschau (›Kaminprüfung‹). Ebenfalls zu unterscheiden sind pädagogische Prüfungen von solchen, die in unserer Praxis zwar an Bedeutung gewonnen haben, aber

---

3 Die Gründe dafür müssen vorerst spekulativ bleiben. So lässt sich etwa die Vermutung anstellen, dass in der politischen Ökonomie der Forschungspraxis zunächst solche Projekte als förderwürdig eingestuft werden, die belastbare Daten über Effekte und Steuerungsmöglichkeiten von technischen Innovationen in der Lehre in Aussicht stellen; nicht also primär solche, deren Erkenntnisinteresse auf unintendierte Folgen und Nebenwirkungen von Transformationen und deren Theoretisierung gerichtet ist (vgl. Radtke, 2003; Reichenbach, 2016).

externe Erfordernisse (bildungs-)politischer Bedürfnisbefriedigung spiegeln: Vergleichsarbeiten und *International Large Scale Assessments* (ILSAs) wie PISA, IGLU oder ICILS (Addey et al., 2025) sowie psychologische und medizinische Diagnostik, wie sie teilweise in Einschulungsprozessen zur Anwendung kommt.

Prüfungen in ihrer spezifisch pädagogischen Form heben auf Fähigkeiten und Kenntnisse einer bestimmten Person ab, nicht die eines Aggregats, wie einer Schulklasse oder gar eines Schulsystems. Sie verfügen über meist unausgesprochene Besonder- und Eigenheiten, die es im Folgenden herauszuarbeiten gilt und für deren Neubesinnung generative KI den willkommenen Anlass bietet. Dabei ist es zunächst nachrangig, ob mit ihnen *kleine* Prüfungen wie Vokabeltests signifiziert sind oder solche, die potenziell enormen biographischen Einfluss ausüben, also etwa Abiturprüfungen, Modulabschlussprüfungen, staatliche Examina usw. Für alle gilt: Es scheint, als würde KI in der Praxis primär als Bedrohung für pädagogische Prüfungen, wie wir sie bisher kannten, wahrgenommen.

An den Beispielen Schule und Hochschule lässt sich dies leicht illustrieren. So fordert die KMK in ihrer Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit KI in schulischen Bildungsprozessen eine »zeitgemäße[] Aufgaben- und Prüfungskultur« (KMK, 2024, S. 6 f.).<sup>4</sup> Allerdings sei ein »allgemeines Verbot von KI zur Bearbeitung von in häuslicher Arbeit anzufertigenden Produkten [...] weder zielführend und wünschenswert noch durchhaltbar« (KMK 2024, S. 6). Zugleich sollen alle »Prüfungsformate, die juristisch nicht einwandfrei die in der Aufgabe geforderte eigenständige

---

4 Die Forderung nach *Zeitgemäßheit* ist freilich eine diskursverweigernde und unwissenschaftliche, da sie von vornherein Alterität ausschließt und somit ihre meist unausgesprochene Normativität immunisiert (vgl. Buck, 2025a, S. 189 f.). Sie ist genau betrachtet eine nach Ahistorizität und Ent-Differenzierung pädagogischer Praxis, da nicht die historisch-systematische Reflexion auf das Phänomen als mögliche Folge aus veränderten Handlungsbedingungen problematisiert, sondern die (niemals zu antizipierende) Innovation eingefordert wird (vgl. affirmativ etwa Langela-Bickenbach et al., 2024). Es ließe sich gar vermuten, dass es sich bei der Forderung nach Zeitgemäßheit um das handelt, was Johannes Bilstein (1990) und Roland Reichenbach (2003) als pädagogischen Kitsch bezeichnen – das Einschleifen differenter Praktiken zugunsten einer neuen Übersichtlichkeit.

Leistung gewährleisten [...] abgeschafft oder grundlegend weiterentwickelt [werden] « (ebd., S. 7). Es gelte

»im Sinne des formativen Assessments die Bewertung weg von der Produkt- hin zur Prozessorientierung zu lenken, wobei der Reflexion durch die Lernenden eine entscheidende Rolle zu gesprochen wird« (ebd.).

Damit sind der Ton und das Maß gesetzt, die zumindest im schulpolitischen Kontext eine dringende und tiefgreifende Reformnotwendigkeit markieren, ohne selbstredend konkrete Lösungen liefern zu können – und zugleich die Wissenschaft als reflexive, forschende, ggf. gar sinngebende Instanz ausklammern. Juristische Bedenken, nicht pädagogische Notwendigkeiten, sind im Sinne der KMK ausschlaggebende Gründe für eine Prüfungsreform mit verbindlichen Regelungen. Ähnlich lautende Forderungen lassen sich auch seitens der Schulbuchverlage (jeweils exemplarisch: Cornelsen, 2024), Professionsverbände (BLLV, 2025) und sich an Lehrkräfte richtende Medienportale (Haverkamp, 2024) finden.<sup>5</sup>

---

5 Für die Präzisierung dieser Fußnote bin ich Nicolaus Wilder zu großem Dank verpflichtet: Dass im *Deutschen Schulportal* Hendrik Haverkamp – Mitgründer der Feedback-KI *Fiete.ai* bzw. inzwischen *FelloFish* – als Gastautor, »Lehrer und KI-Experte« (Haverkamp, 2024) zu Wort kommt, ohne auf wirtschaftliche Interessenkonflikte hinzuweisen sagt viel über den journalistischen Anspruch des Mediums und den laufenden Digitalisierungsdiskurs aus. Sogar mit den besten Absichten einer Verbesserung schulischer Praxis ist es in der fortschreitenden Verwicklung von Ökonomisierungs- und Digitalisierungstransformationen (Buck, 2023; 2026) kaum noch möglich sich der Verschmelzung von Lern- und Konsumpraktiken zu entziehen (vgl. Karcher, Voß & Höhne, 2020, S. 330). Die Vermarktlichung von Schule in Form der Privatisierung ehemals staatlicher Aufgaben im Zuge der Digitalisierung lässt sich auf das Strategiepapier „Bildung in der digitalen Welt“ (KMK, 2016) zurückführen (vgl. Prüfungen post KI im Nexus von Digitalisierung und Ökonomisierung) und sorgt insbesondere hinsichtlich ressourcenintensiver und somit hohe Kosten verursachender KI-Applikationen dafür, dass mit BigTech-Unternehmen wie OpenAI, Microsoft oder Google nur noch solche Initiativen konkurrieren können, die ihre Dienstleistungen kostenpflichtig anbieten – und somit unfreiwillig der Neoliberalisierung öffentlicher Erziehung und Bildung in die Hände spielen (Höhne, 2015).

Ein möglicher Gegenentwurf für eine Rekommunalisierung schulischer Infrastruktur (Radtke, 1997) liegt im Aufbau staatlicher Infrastrukturen, die den Betrieb quelloffener, sicherer und für Schulen kostenfreier KI-Anwendungen ermöglichen – analog zu den größtenteils verschwundenen Laborschulen. Parallel spielt die dritte Phase der Lehramtsausbildung in diesem Zusammenhang eine

Die derzeitige Diskussion um Prüfungen an Hochschulen verläuft sehr viel differenzierter, aber nicht weniger radikal. Inzwischen gehen vereinzelte Hochschulen dazu über, Bachelor-Arbeiten durch mündliche Prüfungen zu substituieren (wie an der Wirtschaftsuniversität in Prag; DLF, 2023) oder umzudeuten in Dokumente zur Protokollierung von eigens durchgeführten Forschungsprojekten (wie im Studienbereich Management und Entrepreneurship der FH Wien; ORF, 2024). Ein solcher Wegfall oder die Umdeutung schriftlicher Prüfungen kann als pädagogischer Offenbarungseid gedeutet werden, der anzeigt, dass sich höhere Bildungseinrichtungen nicht mehr willens oder imstande sehen, die akademischen Kulturtechniken des Recherchierens, Argumentierens und wissenschaftlichen Schreibens adäquat zu lehren und zu prüfen. Auch der in der taz(!) unterbreitete Vorschlag zur Einrichtung von Computerräumen an Universitäten, an denen unter Aufsicht Hausarbeiten erstellt werden sollen (vgl. taz, 2026), bezeugt zunächst einen Vertrauensverlust gegenüber der pädagogischen Zielgruppe.

Anderenorts werden von Studierenden Erklärungen zur Nutzung von KI für die Erstellung schriftlicher Arbeiten eingefordert, wie etwa an den Universitäten Koblenz und Erfurt. Es bleibt zu klären, inwiefern hier genau genommen eine Gewissensprüfung, keine am Gegenstand orientierte erfolgt.<sup>6</sup> An der Universität zu Köln (2025) wird indes empfohlen, ChatGPT und andere

---

Schlüsselrolle, um die unterschiedlichen Wissensstände und Zuständigkeiten innerhalb der Kollegien zu egalisieren. Insbesondere im Kontext digitaler Souveränität scheint eine solche Strategie zunehmend viabler.

- 6 Eine Bewahrheitung dieser Vermutung hätte vor allem deswegen ein Geschmäckle, weil auch prüferseitig vermehrt LLM-Anwendungen in der Lehre, zur Korrektur von Klausuren und Hausarbeiten, zur Begutachtung von Bachelor- und Master-Arbeiten oder zur Erstellung anderer Gutachten, etwa für Begabtenförderungswerke, zum Einsatz kommen (vgl. exemplarisch Willatt & Ossa, 2026). In der norwegischen Universität in Bergen kam es im Herbstsemester 2025 zu Beschwerden von Medizinstudierenden, deren Prüfungen durch eine KI vorbeurteilt wurde. Zugleich haben einer britischen Untersuchung zufolge 22 % der Studierenden ungebührlichen Gebrauch von KI im Studium angegeben (vgl. Khrono, 2026a; b). Insofern hängt der vermehrte Einsatz von KI und LLMs an Hochschulen ggf. mit einem möglichen Wandel wissenschaftlicher Bereichsethik und den ungeklärten Fragen nach ihrem legitimen Gebrauch zusammen, was derzeit bspw. in der Forschungsgruppe »Ethik und kritische Theorien der KI« an der Uni Osnabrück erforscht wird: <https://www.uni-osnabrueck.de/fb8/ikw/forschung/arbeitsgruppen/ethik-der-ki>.

LLMs als Quelle bzw. Autor aufzuführen.<sup>7</sup> Es lässt sich als aktuelles Problem festhalten, dass derzeit und offenkundig große Unsicherheit und Uneinigkeit hinsichtlich des Einflusses von KI auf pädagogische Prüfungen und die adäquaten Reaktionen darauf vorherrschen.

Weder sollen jedoch im Folgenden die Zwecke diskutiert werden, die Prüfungen in pädagogischen Kontexten entfalten – insbesondere, wenn es sich dabei um außerpädagogische wie etwa Employability oder die Legitimierung von Ungleichheit in einer Meritokratie handelt –, noch soll einer zweckrationalen Perspektive auf KI das Wort geredet werden, die zur Klärung abhebt, inwiefern der Einsatz von LLMs pädagogischen Prüfungen dienlich ist oder zu verhindern wäre. Auch soll es nicht um die »Fragwürdigkeit der Zensurenggebung« (Ingenkamp, 1989) gehen, also die Reliabilität und externe Validität pädagogischer Prüfungsergebnisse und ihrer Eignung als Prädiktoren für die weitere Bildung- und berufliche Karriere von Geprüften. In Abgrenzung von solch funktionalistischen Fragen soll vielmehr diese im Zentrum stehen, wie sich spezifisch pädagogische Eigenheiten von Prüfungen historisch-systematisch begründen lassen.<sup>8</sup>

Nur wenn man dieser Fragerichtung folgt, lässt man sich nicht das pädagogische Heft aus der Hand nehmen in der Deutungshoheit darüber, nach welchen Kriterien Prüfungen und auf sie vorbereitende Lehre, Unterricht usw. zu gestalten sind. Die diesem Beitrag zugrundeliegende Vermutung lautet, dass im Zuge voranschreitender Neoliberalisierung des Bildungssystems (Radtke, 1997; Bellmann & Weiß, 2009; Höhne, 2015; Buck, 2023) und der Etablierung eines »Testregimes« post PISA (Brinkmann, 2009) das

---

7 Damit bricht die Universität Köln radikal mit einer wissenschaftlichen Diskursethik, die den Ausweis der Herkunft eines Argumentes als Fundament des akademischen Handwerks anerkennt. Gerade in Anbetracht der Opazität bzw. Irreversibilität von LLMs – im Gegensatz zu Algorithmen lassen sich KI-Outputs nicht zurückrechnen, weil uns entzogen ist, wie sie arbeiten (vgl. Buck 2025a) – scheint dieses Beispiel in besonderer Weise Ohnmacht zu bezeugen.

8 Hans-Werner Prahl (1980a) differenziert aus sozialisationstheoretischer Perspektive acht Funktionen von Prüfungen: Die Schaffung von Identität/Selbstkonzept, Motivierung, Orientierung, Sozialisierung, Disziplinierung, Rückmeldung, Evaluation und Selektion/Statuszuweisung (ebd., S. 133 ff.). Erst später wird die Funktion der Initiation ergänzt (Prah, 1983; vgl. 2.1).

Nachdenken über pädagogische Prüfungen sui generis außer Mode geraten ist und daher einer Neubesinnung bedarf.

Folglich erscheint eine historisch und theoretisch informierte Thematisierung von Prüfungen als pädagogische Handlungen notwendig (2). Anschließend soll an einem Fallbeispiel illustriert werden, welche programmatischen und paradigmatischen Transformationen mit Prüfungen in Folge der Etablierung von KI und im Nexus von Digitalisierung und Ökonomisierung verbunden sind (3). Abschließend wird mit der Prüfung der Originalität ein bildungstheoretisch fundierter Vorschlag unterbreitet, Prüfungen auf Individuen und deren Bildungsprozesse zu re-fokussieren (4).

## 2 Prüfungen als pädagogische Handlungen

Fragt man Helmut Fends »Neue Theorie der Schule« (Fend 2008) um Rat, so ist die Frage nach der pädagogischen Funktion von Prüfungen schnell beantwortet. Sie sollen in erster Linie nicht personalisieren oder qualifizieren, sie sollen auch nicht integrieren, sondern *selektieren* und *allozieren* (ebd., S. 51). Mit anderen Worten: Die Prüfung dient aus soziologischer Warte als Instrument für die Legitimation sozialer Ungleichheit, die sich aus der Leistungserbringung der Geprüften ableitet. Damit verbunden ist die gesellschaftliche Anerkennung der Prüfungspraxis pädagogischer Institutionen, sofern diese nicht einfach allen Prüflingen eine Bestnote verschafften, sondern (dem äußeren Anschein nach) individuelle Leistungen an fachlichen Kriterien messen und differenzierend beurteilen. Im Sinne Fends wird bis heute die pädagogische Leistungsbeurteilung von der Gesellschaft in Form von Zeugnisnoten und (beruflichen, akademischen, schulischen) Abschlusszeugnissen weitgehend als gültig anerkannt, obschon etwa im tertiären Bildungssektor Eingangsprüfungen tendenziell häufiger werden.<sup>9</sup>

Möchte man jedoch schulische Handlungen in ihrer Pädagogizität verstehen und von solchen psychologischer, soziologischer oder kulturwissenschaftlicher Provenienz theoretisch fundiert differenzieren können (vgl. Kraft, 2021), so muss das folglich auch für die Praxis der pädagogischen Prüfung möglich sein. Die Breite des Wortgebrauchs allerdings erschwert eine Theoretisierung dieses pädagogischen Phänomens: Vom Vokabeltest in der Mittelstufe bis zur Disputatio einer Doktorarbeit kommen Prüfungen sehr

---

9 Fend setzt hier das Ideal einer meritokratischen Gesellschaft voraus, die mit Hilfe ihrer (schulischen) Institutionen zielsicher Individuen nach ihren Leistungsvermögen an unterschiedliche und verdiente Positionen zuzuteilen vermag. Es darf vor dem Hintergrund der Ungleichheits- und Eliteforschung von Bourdieu & Passeron (1971) über Hartmann (2002) bis hin zu Krüger et al. (2011) Zweifel an der Tragfähigkeit der Prämisse angemeldet werden (vgl. Die Prüfung der Originalität als pädagogisches und demokratisches Projekt).

unterschiedliches Gewicht in den Bildungsbiographien ihrer Prüflinge zu.<sup>10</sup> Zudem erzeugen Prüfungen unerwünschte Nebenwirkungen, die in die entsprechende Theoriebildung zu integrieren wäre: Phänomene wie Prüfungsangst (vgl. Prah, 1977a) und Prüfungswillkür (vgl. Kvale, 1972) sind als unerwünschte Effekte lange und wohlbekannt, auch die Tendenzen zur fortschreitenden Verrechtlichung, Bürokratisierung und Verschulung (vgl. Prah, 1980a, S. 81 ff.; 1980b) sind dokumentiert.

Reichweitenstark waren Prüfungen zuletzt 2017 in der Zeitschrift für Pädagogik Thema. In der Einleitung des entsprechenden Thementeils konstatieren Norbert Ricken und Sabine Reh, dass Prüfungen »offenkundig schlechte Reputation im pädagogischen und erziehungswissenschaftlichen Diskurs« (Ricken & Reh, 2017, S. 247) genießen. Sie sind möglicherweise genau deswegen größtenteils »Leerstellen pädagogischer Reflexion« (ebd.) bzw. »zu den Tabubereichen gehör[end]« (Prah, 1977b, S. 1) – oder gar »Stiefkinder[] der pädagogischen Theorie« (Loch 1963, S. 180).

Die Situation der Prüfung als erziehungswissenschaftliches Anathema trotz ihrer Allgegenwart im pädagogischen Alltag spiegelt sich im kanonisierten Wissen wider. Wir finden keinen Eintrag zu Prüfungen im »Historischen Wörterbuch der Pädagogik« (Benner & Oelkers, 2010), auch nicht im »Handbuch Erziehung« (Sandfuchs et al., 2012), im »Grundwissen Pädagogik« (Kron, Jürgens & Standop, 2013), im »Handwörterbuch Erziehungswissenschaft« (Andresen et al., 2009) oder in den »Schlüsselbegriffe[n] der Allgemeinen Erziehungswissenschaft« (Feldmann et al., 2024). Im immerhin 800-seitigen »Beltz Lexikon Pädagogik« heißt es spärlich:

»Prüfung: Leistungsbewertung im Rahmen eines Bildungsgangs/einer Ausbildung, die unter standardisierten, juristisch kontrollierbaren Bedingungen stattfindet und über die Vergabe eines Abschlusszertifikats entscheidet; kann als Abschlussprüfung oder ausbildungsbegleitend in

---

10 Hinzu kommt die semantische Dimension einer wissenschaftlichen Prüfung, nicht nur von Thesen im Rahmen empirischer Forschung, sondern auch formallogischer Art, nämlich von Aussagen und Prämissen auf ihre Widerspruchsfreiheit, Konsistenz etc., wie sie bis heute in der transzendental-kritischen Pädagogik gepflegt wird (exemplarisch: Petzelt, 2018).

mehreren Schritten durchgeführt werden.« (Tenorth & Tippelt, 2007, S. 583, Sp. 2)

Und im Wörterbuch der Pädagogik (Böhm & Seichter 2025, S. 391) wird vermerkt, dass Prüfungen »arrangierte Situationen [sind], in denen bestimmte Leistungsvollzüge provoziert werden, um diesen zugrundeliegende Fähigkeiten, Wissen, Können und Dispositionen zu messen« (ebd.). Dieser Setzung entgegen steht die auf Noam Chomsky zurückgehende These, dass Prüfungen gerade nicht Kompetenz, sondern Performanz messen (vgl. Radtke, 2003). Lediglich in den nicht mehr aufgelegten »Pädagogische[n] Grundbegriffe[n]« (Lenzen & Rost, 2007) finden wir einen Eintrag, der auf die Etymologie des Wortes eingeht:

»Das lateinische Verb ›probare‹ bezeichnet bereits einen der heutigen Bedeutungen des Wortes Prüfung entsprechenden Vorgang: auf Echtheit und Güte untersuchten, als gut erkennen, billigen, erproben. Im Mittelhochdeutschen wird ›prüevunge‹ im Sinne von Beweisführung, Bewährung, Erprobung verwendet. Aus dem Verb ›prouven‹ wird im Neuhochdeutschen ›prüfen‹, mit der Bedeutungsdekonzentrierung von ›examinare‹ (heraustreiben), das dann in erster Linie in der ›Schulsprache‹ Anwendung findet.« (Bichler & Brauner, 2007, S. 1285)

Thomas Bichler und Klaus Brauner vermengen in ihrem Beitrag funktionalistische Positionen soziologischer und psychologischer Art und bleiben eine Erläuterung des Pädagogischen der pädagogischen Prüfungen schuldig. Sie stehen damit stellvertretend für den Umgang der Erziehungswissenschaft mit der Praxis der Prüfung, die meist wie ein Fremdkörper administrativer oder psychologischer Zuständigkeit behandelt wird.

Diese Situation gibt Norbert Ricken & Sabine Reh, Werner Loch sowie Hans-Werner Prahle Recht in ihren gleichlautenden Mutmaßungen über den Status von Prüfungen im erziehungswissenschaftlichen Diskurs.<sup>11</sup> Der schlechte Ruf der Prüfung mag ebenso erklären, warum das Prüfen auch nicht als elementare (vgl. Prange & Strobel-Eisele, 2014) bzw. Grundform

---

11 In den vergangenen Jahren sind Praktiken, Formen und Techniken des Prüfens vermehrt in das Interesse bildungshistorischer Forschung gerückt, insbesondere mit Blick auf (zentrale) Abiturprüfungen und IQ-Tests. Vgl. hierzu exemplarisch Engelhardt, 2013; Kämper van-den-Boogaart et al., 2013; Gelhard, 2018.

pädagogischen Handelns (vgl. Giesecke, 2015) geführt wird, folgt man entsprechend lautenden Sammlungen.

## 2.1 Loch: Prüfungen als Lernaufgaben

Erstaunlich ist der genannte Befund deswegen, da das Prüfen untrennbar zum pädagogischen Geschäft gehört, wie auch Werner Loch (1963, S. 183) vermerkt: »Die Arbeit des Erziehens ist nicht ohne die Arbeit des Prüfens möglich.« Er erinnert daran, dass das Prüfen bereits in der Form des sokratischen Gesprächs bildende Funktionen erfüllt, da die Elenktik – das fragende Prüfen – nur in Verbindung mit der Mäeutik als pädagogische Form der Geburtshilfe der Erkenntnis zu verstehen sei. Während erstere in universeller Absicht Wahrheit zu suchen beabsichtigt, weist letztere auf das geprüfte Individuum und die durch die Prüfung evozierte Selbsterkenntnis im Medium der Wahrheit zurück (vgl. ebd., S. 184) – Γνωθι σεαυτόν!<sup>12</sup>

Das Fragen ist in der Geschichte der Pädagogik die Urform der Prüfung, ohne dass mit dieser Feststellung bereits Frageabsicht, Erwartungshorizont und Verhältnis von Frage und Antwort hinreichend geklärt wären. Didaktisch entsteht aus ihr das Gespräch als »grundlegende Unterrichtsform [in] prüfende[r] Funktion« (ebd.), wie sie bei Eugen Fink (1970) im Begriff der *Fragegemeinschaft* pointiert wird und an die notwendige Sozialität pädagogischer Prüfungen erinnert.<sup>13</sup> In pädagogischen Prüfungen wird im Modus des Fragens jedenfalls immer zweierlei geprüft: das gefragte Fachwissen oder Können (etwa in musisch-künstlerischen Fächern, im

---

12 Die Aufforderung des Gnothi seauton, am ehesten bekannt als Inschrift einer Säule des Apollon-Tempels in Delphi, geht ursprünglich auf Heraklit zurück. Der Imperativ des »erkenne Dich selbst« steht zugleich für die höchste Bildungsaufgabe des Menschen und die Erinnerung an die Schwierigkeit, diese zu bewältigen. Prüfungen sind die elenktische Hilfe, der Selbsterkenntnis überhaupt näher kommen zu können.

13 Sowohl bei Kerschensteiner als auch bei Montessori finden wir Formen vermeintlicher Selbstprüfung (Loch, 1963, S. 190; Montessori, 1996, S. 112f.), die bei genauerer Beschau jedoch keine pädagogischen Handlungen darstellen, sondern Selbsttechniken, deren Gelingen final durch pädagogische Autorität bestätigt werden müssen.

Chemiepraktikum, in beruflichen Schulen usw.) und die jeweiligen Träger des Wissens und Könnens.

Aus der Ableitung der Tradition der Katechetik, so Loch weiter, wird die (nicht beliebige) Gliederung des Prüfungsstoffs und die Verbindung von Lehre und Prüfung plausibel, während das Leistungsprinzip als Einfluss nicht-pädagogischer Interessen zugleich auf das gesellschaftliche Vertrauen in ihre pädagogische wie sachliche Angemessenheit verweist (vgl. Loch 1963, S. 188 f.).<sup>14</sup>

Mit Werner Loch lassen sich demnach soziale, erzieherische und didaktische Dimensionen von Prüfungen diskutieren, die diese weder ausschließlich funktionalisieren, noch einseitig auf Gängelung und Machtausübung zurückführen,<sup>15</sup> sondern bildende und erziehende Momente im Modus der Prüfung freizulegen erlauben. Zugleich würden Prüfungen jedoch der Bildungsidee entgegenlaufen, da sie Geprüfte »immer nur in einer bestimmten Hinsicht« ansprechen und somit Gefahr liefen, »den Menschen zu versachlichen und damit zu erniedrigen« (Loch, 1963, S. 181). Als pädagogische Handlungen erfordern Prüfungen folglich ein gewisses Fingerspitzengefühl im Umgang mit den Prüflingen und dem Prüfungsgegenstand, was Johann Friedrich Herbarts Figur des *pädagogischen Taktes* (Herbart 1982; vgl. auch Muth 1962; Burghardt, Krinninger & Seichter 2015) über das Unterrichtsgeschehen hinausweisen lässt.

---

14 Von Heinz-Elmar Tenorth (2024) habe ich das Argument geborgt, dass die Abiturprüfung ein wichtiger Grund für die Legitimation von Schulen als Prüfinstanz von Studierfähigkeit darstellt, was sich an der Funktion von Abiturnoten als hartem Kriterium für die Zuteilung von Studienplätzen mit Evidenz versehen lässt. Erstaunlicherweise stimmt das noch immer, obschon Hochschulen vermehrt auf Eingangstest als ergänzendes oder ersetzendes Kriterium zurückgreifen, wenngleich hierzulande sehr viel seltener als bspw. im anglophonen Raum.

15 In ihrer Doktorarbeit benennt Sabrina Schröder in poststrukturalistischen Annäherung Prüfungen als »Knotenpunkt pädagogischer Disziplinierung und Kontrolle, der dem Anspruch einer kritischen oder emanzipierenden Bildungsidee im Wege zu stehen scheint« (Schröder, 2026, S. 15), die zudem als Medium wirksam werden, in dem sich die Subjektivierung von Schülerinnen und Schülern vollzieht (und demnach auch produktive Wirkung entfalten). Diese theoretische Perspektive, auf das von Michel Foucault bereits benannte Desiderat einer Geschichte der Prüfung aufbauend, wird hier aufgrund der paradigmatischen Differenzen im Denken nicht weiterverfolgt.

Allein: Auch diese Anmerkung wird zum normativen Lastenheft des pädagogischen Handelns addiert, ohne dass damit eine Methode beschrieben wäre, wie man denn das Prüfen als Handlung erlernt. Es verdichten sich mit dem pädagogischen Takt als Bezugsgröße die Hinweise darauf, dass die Voraussetzung für genuin pädagogisches Prüfen im Ethos bzw. der Professionalität zu finden ist. Werner Lochs letzter Hinweis verweist ebenfalls auf dieses Argument, denn pädagogische Prüfungen unterscheiden sich in genau diesem Aspekt von juristischen, medizinischen oder solchen anderer Zünfte: Die Anerkennung der Lernfähigkeit bzw. Bildsamkeit Lernender bei gleichzeitiger Anerkennung ihrer Würde bleibt auch trotz misslungenen Prüfungen gegeben. Dieses pädagogische Wohlwollen, so fruchtlos es in manchen Fällen sein mag, ist ethische Voraussetzung unserer hoffnungsvollen Praxis. Insofern sind pädagogische Prüfungen Teil des Geschäfts, ohne dass sie zugleich im Zweck und Ziel des pädagogischen Handelns allgemein aufgingen. In den Worten Werner Lochs:

»Gleich wie die Prüfungen ausgehen: die Aufgabe der Erziehung bleibt davon prinzipiell unabhängig. Ein Prüfungsergebnis kann deshalb weder über die Richtigkeit einer pädagogischen Praxis noch darüber definitiv befinden, ob sie tatsächlich ihre Arbeit vollkommen verwirklicht hat.« (Ebd., S. 182)

Werner Loch zufolge sind Prüfungen Teil jeder Lernaufgabe, weil sie zur Beherrschung bestimmter Kenntnisse und Fähigkeiten spätestens zum Prüfungszeitpunkt auffordern und zugleich potenzieller Ort einsetzender Lernhilfen sind. Damit integriert Loch Prüfungen in sein erziehungstheoretisch fundiertes »Kreuz der Erziehung«, mit Hilfe dessen sich sämtliche erzieherische Handlungen in einem Fünfschritt fassen lassen. Nach 1) der Voraussetzung der Lernfähigkeit einer bestimmten Sache werden 2) Lernaufgaben gestellt, bei deren Bearbeitung es zwangsläufig zu 3) Lernhemmungen kommt. Diese zu erkennen und ihnen in Form von 4) Lernhilfen abzuhelpfen, ist zentrale pädagogische Aufgabe. Nach (meist mehrmaligem) Scheitern stellt sich dann 5) ein Lernerfolg ein (vgl. Buck, 2012).

Pädagogische Prüfungen sind demnach elementarer Bestandteil jeglicher Lernaufgaben, insofern sie einen didaktischen Zweck verfolgen, keinen testend-diagnostizierenden oder gar vergleichenden. Diese Differenz zwischen

Test- und didaktischen Aufgaben scheint sich im vergangenen Vierteljahrhundert seit der ersten PISA-Studie zunehmend eingeschliffen zu haben (vgl. Brinkmann, 2009) – und es scheint sinnvoll, sie wieder zu betonen.

Geprüft wird im Sinne Lochs, was als Lernaufgabe beigegeben wird, wobei es Prüfungen gibt, die für den gesamten Lebenslauf bedeutsam werden. Von diesen *curricularen Lernaufgaben* arbeitet Loch insgesamt 22 aus, wobei der erfolgreichen Prüfung der elften curricularen Lernaufgabe eine besondere Bedeutung zukommt. Das Ergebnis dieser ist die *Initiation* in die vorliegende Gesellschaft – für heutige, westliche Gesellschaften in Form eines schulischen Abschlusses. Zu erlernen ist auf dieser Stufe im Modus der Prüfung die Selbstdarstellungsfähigkeit, deren Lernhemmungen unterschiedlich in Erscheinung treten, nämlich als

»Angst zur Selbstdarstellung, einem Mangel an Zivilcourage, Scham oder Ähnliches eine Nötigung empfunden wird oder wenn die Aufmerksamkeit Anderer nur zur Befriedigung der eigenen Eitelkeit dient.« (Buck, 2012, S. 96)

Gelingt das Erlernen der Selbstdarstellungsfähigkeit, wird das Individuum schließlich gesellschaftlich mündig gesprochen. Die Sozialität pädagogischer Prüfungen zeigt sich jedoch nicht nur in der gesellschaftlichen Ab- oder Anerkennung, sondern auch in der direkten Interaktion im Prüfungsgespräch. Wie uns u. a. Klaus Prange und Gabriele Strobel-Eisele (2014, S. 84ff.) erinnern, verweisen Prüfungen in der Form des reaktiven Zeigens immer auch auf die Eignung von Prüfenden hinsichtlich ihrer (unterrichtlichen, lehrenden) Vorbereitung und (fragenden, weisenden) Durchführung von Prüfungen: Sie wirken geradezu »informierend und aufklärend auf Lehrende und Lernende zurück[]« (Böhm & Seichter, 2025, S. 391).

## 2.2 Ruhloff: Drei Formen von Prüfungen

In einer weiteren Perspektive differenziert Jörg Ruhloff (2008) drei Prüfungsformen voneinander und mit Blick auf ihre gesellschaftliche Relevanz. Diese werden von ihm in eine chronologische, nicht aber sachlogische Abfolge zueinander gestellt und bezeugen den Bedeutungswandel von Prüfungen für die jeweilige Gesellschaftsform. Die *sokratisch-tauglichkeitskritischen*

*Prüfungen* der Freigeborenen der attischen Polis, wie sie von Ruhloff als Ausgangspunkt des pädagogischen Prüfungswesens genannt werden, hätten allein bildenden Wert. Die sokratische Prüfung qualifiziert für nichts, sie bilanziert auch nicht. Ihr Anspruch liegt in der universalistischen Erkenntnis im Angesicht der Aporie und im Modus spekulativer Philosophie (vgl. ebd., S. 17). Das sokratische Gespräch erfüllt eine sozial-moralische Funktion insofern als sie Bescheidenheit evoziert – in ihm wurde jede und jeder zielsicher in die Aporie getrieben, was ein hybrides Selbstbild zu verhindern verstand.

Dieser Prüfung von Argumentations- und Fehlbarkeitseinsichtsfähigkeit folgten ab der Frühmoderne *emanzipatorische Prüfungen* wie bspw. das preußische Genie-Examen, das talentierten Söhnen(!) von Handwerkern und Bauern erlaubte, sich aufgrund besonderer Eignung vom Kriegsdienst freustellen zu lassen und den üblicherweise anderen Ständen vorbehaltenen Besuch von Schulen und Hochschulen zu erlauben (vgl. ebd., S. 11f.). Diese Form der Prüfung hat primär sozialpolitische Funktionen, für die Fachlichkeit Mittel zum Zweck darstellt, nicht aber deren erklärtes Ziel.

Schließlich reüssierte in den vergangenen Jahrzehnten die *Qualitätskontroll-Prüfung*, deren Ursprung Ruhloff auf die 1.500 Jahre zurückliegende Missionierung des heutigen Großbritanniens zurückführt. Aus einer »kirchengeschichtlichen Praxis« (ebd., S. 21), nämlich der Glaubensprüfung potentieller Konvertiten, ihrer *certitudo*, habe sich daraus die Institutionalisierung eines Zertifizierungsglaubens entwickelt, der inzwischen in fast alle Lebensbereiche kapitalistischer Gesellschaften übergegangen ist.<sup>16</sup> Nach erfolgreicher Hauptuntersuchung erhalten Kraftfahrzeuge neue HU-Plaketten, eine erfolgreich absolvierte Gesellenprüfung wird mit einem entsprechenden Zertifikat quittiert, erfolgreich akkreditierte Studiengänge werden mit einem entsprechenden Siegel der jeweiligen Drei-bis-fünf-Buchstaben-Agenturen versehen.

---

16 Auch das *Examen*, darauf weisen Ricken und Reh (2017, S. 249) hin, lässt sich begrifflich auf die katholische Gewissenserforschung zurückführen. Die stete Prüfung durch Gott in der jüdischen und christlichen Theologie und ihren Einfluss auf pietistische Grundlegung pädagogischer Praxis (etwa bei Francke) illustriert Loch (1963, S. 185f.) anschaulich.

Der Zertifizierung eingeschrieben bleibt ihre »soziale Funktion, den Zusammenhalt der Kirchenmacht durch Überprüfung der Rechtgläubigkeit und der Glaubens-Zuverlässigkeit zu sichern« (ebd., S. 26). Es ist ein »Säkularisat« (ebd., S. 21) einer »unablässig prüfenden und geprüften Weltgesellschaft« (ebd., S. 7) in vergeblicher – da selbsttäuschender – Absicht, Haltungen, Fähigkeiten und Fertigkeiten zuverlässig attestieren zu können und das, »obwohl es doch nicht gerade schön und angenehm ist, andauernd zu prüfen und geprüft zu werden« (ebd., S. 22). Trotz dessen hat sich offenbar die Zertifizierungsprüfung als eine omnipräsente durchgesetzt, möglicherweise wegen ihrer komplexitätsreduzierenden Funktion in sukzessive steigenden Unübersichtlichkeitsbedingungen der Spätmoderne.

Ruhloff bedient sich in seiner Argumentation dreier Frühschriften Johann Friedrich Herbarts, dessen Name unfreiwillig »mit der flächendeckenden Durchsetzung des öffentlichen Disziplinar- und Kontrollschulwesens verbunden« (Ruhloff, 2008, S. 11) wird und der klugerweise vermerkt, dass bei einer Prüfung mindestens drei Elemente zu beachten seien: Gegenstand, Geprüfter und Prüfender – was sie als triadische, didaktische Praktik ausweist. Dem Anspruch einer nach bildungstheoretisch fruchtbaren Momenten fragenden Analyse von Prüfungen entsprechend, soll folgende These entfaltet werden: Trotz der sozial-regulativen Funktion von heute dominanten Qualitätssicherungsprüfungen erfüllen (hoch-)schulische Prüfungen bis heute pädagogische Funktionen. Drei davon sollen in Kürze genannt werden:

- 1) Prüfungen sind nicht nur Anlass und Auslöser für mehr oder minder aufwändige, sorgfältige oder gelungene Rekapitulationen vorangegangenen Unterrichts, auch verfertigt sich in ihnen eine systematische Ordnung von Wissen, was all jene bezeugen können, die in mündlichen Prüfungen Mäeutik betreiben oder in Klausuren, Hausarbeiten oder anderen schriftlichen Zeugnissen dem verschriftlichten Denken ihrer Geprüften nachfolgen können. In ihnen vollzieht sich folglich das, was Herbart (2022, S. 50 ff.) als *Besinnung* in den Formalstufen *System* und *Methode* kennzeichnet, die Überführung disparater Wissensbestände in eine theoretische Ordnung, denn »System fordert niemand von der Erfahrung« (ebd., S. 51) und »[d]er Fortschritt der Besinnung ist Methode« (ebd., S. 50). Beide bräuchte es, um zur »höchste[n] Gelenkigkeit des

Denkens« (ebd., S. 134) zu gelangen. Prüfungen sind demnach im Sinne Herbarts willkommene Anlässe und Fokalfpunkte zur Synthese von Wissen und Können.

- 2) Der von Ruhloff vorgebrachte Aspekt einer zertifizierenden, qualitätssichernden Prüfung zur Regulation sozialer Zugehörigkeit wird auch dann plausibel, denkt man an die Inauguration in die Wissenschaft mit Bestehen der Promotion. Die weihevoll Aufnahme in den erlauchten Kreis derer, die erfolgreich eine Prüfung bestanden haben, vollzieht sich analog mit der Beendigung schulischer und hochschulischer Bildung im Akt der Zeugnisübergabe. Als zeremonieller Akt der Bezeugung vor anderen wird abermals das soziale Moment von Prüfungen deutlich. In gleicher Weise gilt diese Form der Anerkennung auch für kleinere bestandene schulische wie außerschulische Prüfungen wie Fahrradführerscheine, Klausuren, Trainerscheine usw. Nach Werner Loch bleiben pädagogische Prüfungen in ihrer Form jedoch paradoxal, da

»ohne sie die Selbstdarstellung nicht wirklich gelernt werden kann, obwohl sie ihrerseits dasjenige Erziehungsmittel ist, das in Gestalt von Prüfungsangst, Mangel an Selbstvertrauen, schlechtem Gewissen wegen mangelnder Vorbereitung die stärksten Lernhemmungen hervorruft. Prüfungen sind jedoch unvermeidlich. Wer den schulischen Prüfungen ausweicht, muß in den Prüfungen des Lebens dafür bezahlen.« (Loch, 1996, S. 175)

- 3) Freude, die man als Prüfender oder Lehrender während oder nach erfolgreichen Prüfungen verspürt, ist niemals nur auf eine angemessene Prüfungsvorbereitung oder das Funktionieren von Aufgabenformaten zurückzuführen. Mit jeder erfolgreich vollzogenen Prüfung erfolgt auch eine Reproduktion des Fachs bzw. der Disziplin. Erziehungstheoretisch gewendet: Jedes Zeigen beinhaltet eine chronologische Dimension insofern, als es darauf abzielt, dass das Gezeigte vom Lernenden wieder-gezeigt werden kann (vgl. Prange, 2012, S. 73). Dies bezieht sich nicht nur auf individualisiertes Wissen, sondern auch auf die gemeinsame Aushandlung von Wissen als Gegenstand disziplinären Bildungsinhalts, wie es Wolfgang Klafki (1964, S. 9 ff.) treffend formulierte.

Insofern qualifizieren, integrieren und enkulturieren Prüfungen im Fend'schen Sinne durchaus (aber nicht immer und zwangsläufig!), weil durch sie nicht nur ausschließlich Leistung bezeugt wird, sondern auch, weil sie Geprüfte zum systematischen Denken auffordern, ihnen darüber hinaus zeigen, was ihnen leicht von der Hand geht und was nicht, wo deren Neigungen und Interessen liegen und schließlich, weil Prüfungen zugleich fachliche Bestände und disziplinäres Wissen und dessen Ordnung tradieren.

Vor dem Hintergrund dieser allgemeinen Rekapitulation pädagogischer Prüfungen stellt sich die Frage, inwiefern die fortlaufende KI-Entwicklung als Teil fortschreitender Digitalisierung und Ökonomisierung pädagogischen Denkens und Handelns uns dazu auffordert oder gar nötigt, anders über Prüfungen nachzudenken, ohne in reaktionärer Weise auf die Wiederkehr einer sokratisch-tauglichkeitskritischen zu insistieren.

### 3 Prüfungen post KI im Nexus von Digitalisierung und Ökonomisierung

Der Dschinn ist aus der Flasche. Selbstverständlich kann es kein Zurück mehr geben in eine Welt ohne KI und die Allgegenwart von LLMs und deren Anwendungen, die nicht immer auf wohlwollende Rezeption stoßen. Im Englischen ist von *AI slop* die Rede, also sinngemäß von KI-Schrott. Alternativ spricht man von *AI pollution*, also der Verschmutzung der Umwelt durch KI.<sup>17</sup> Eine Abschaffung von Prüfungen aus Angst vor Täuschung durch unerklärte Hilfsmittel jedenfalls kann nicht die Antwort sein, dafür sind sie gesellschaftlich wie pädagogisch zu zentral, wie sich ausweislich des Dargelegten plausibilisieren lässt – und zwar trotz der Tatsache, dass in der marxistisch-anarchokapitalistischen (Illich, 1971; vgl. Zenke, 2022) wie vulgären (Precht, 2013) Schulkritik diese Forderung nach Abschaffung von Prüfungen regelmäßig Wiederkehr feiert. Es ist uns vielmehr aufgetragen, eine Neubesinnung im Angesicht der Transformationen von Prüfungen in Zeiten von KI vor dem Hintergrund ihrer pädagogischen Zwecke vorzunehmen. Diese soll in drei Annäherungen geschehen: empirisch, programmatisch, paradigmatisch.

*Empirisch* soll uns ein Social-Media-Beitrag erhellen, der Aufschluss gibt über die Inneneinsichten eines mutmaßlich 18-jährigen Schülers (Reddit, 2025). Dieser berichtet davon, dass es in seiner Klasse gängig geworden sei, »ganze Klausur[en] bei Chatgpt ein[zugeben]«, da sie in 90 % der Fälle damit

---

17 Diese Analogie ist deswegen bemerkenswert, weil die Assoziationen zur Umweltverschmutzung impliziert, dass von einer reinen oder unverschmutzten Ausgangslage ausgegangen wurde, die nun verunreinigt zu werden droht. Es stimmt zwar, dass der durch KI enorm gesteigerte Rechenbedarf tatsächlich zur Umweltverschmutzung beiträgt. Der weltweite Stromverbrauch steigt dadurch Schätzungen zufolge jährlich(!) um 16 % (mittleres Szenario; Gröger et al., 2025, S. 15). Eine solche Perspektive reiht sich in das ein, was bei Klafki als epochaltypische Schlüsselprobleme verhandelt wird (Münzinger & Klafki, 1995), allerdings nicht als reines Umweltproblem. Die Prämisse hingegen darf und sollte dennoch kritisiert werden: Auch vor ChatGPT und Co. gab es mittelgute bis schlechte Haus- und Abschlussarbeiten, Referate etc.

durchkämen, »und selbst wenn sie erwischt werden, dann sind es halt einmal 0 Punkte und sonst immer was zwischen 10-15 Punkten« (ebd.).

Neben dieser Bewertungsungleichheit seien es jedoch zentrifugal wirkende Effekte im Sozialen, die durch LLM-Nutzung entstünden:

»Vor allem in Ethik bereitet es mir Probleme, da wir anspruchsvolle philosophische Texte lesen und verstehen müssen, jedoch manche Schüler den Text einfach von Chatgpt vereinfachen oder sogar zusammenfassen lassen und sich innerhalb von 2min melden, um den Inhalt eins zu eins so wiederzugeben« (ebd.).

Dennoch wolle er »niemanden verpetzen, weil man dann als Außenseiter dasteht, der anderen keine guten Noten gönnt« (ebd.). Der letztgenannte Punkt scheint besonders thematisierungswürdig: Die KI-Nutzung verändert das soziale Gefüge an pädagogischen Institutionen. Am genannten Beispiel der Nutzung in der Schule wird diese These plausibel.

Aber auch an Hochschulen gerieren LLMs zu *Kontingenzerzeugungsmaschinen*, die Dozierende mitunter ein Misstrauen gegenüber der aufrichtigen Arbeitsweise ihrer Studierenden entwickeln lässt, während jene befürchten, fälschlicherweise einer illegitimen bis illegalen KI-Nutzung bezichtigt zu werden. Das hat schwerwiegende Folgen. Aus Commiltonibus, die gemeinsam für die Wahrheit kämpfen, werden sich einander kritisch beäugende Gruppen. Was die genannte Unsicherheit im Sozialen noch verschärfen mag, ist die seit etwa einem Jahr andauernde *Charmeoffensive* der LLMs. Wird jede kritische Prompt-Rückfrage als genial oder bahnbrechend markiert, kann ein falsches Selbstbild von Prüflingen über die eigene Denk- und Leistungsfähigkeit emergieren. Der Effekt ist dann der Gegenteilige einer sokratischen Prüfung: nicht Bescheidenheit, sondern ein hybrides Selbstbild der eigenen Denkfähigkeit, bezeugt durch eine – *naja* – künstliche Intelligenz.

Es zeigt sich bereits jetzt und anhand solcher ausschnitthaft bleibenden Einzelempirien, dass KI weit über ihre technische Dimension hinaus Folgen nach sich zieht. Wie Melvin Kranzberg in dessen ersten berühmten Gesetz postuliert: »Technology is neither good nor bad; nor is it neutral« (Kranzberg, 1986, S. 545), was uns in der Folge dazu auffordert, auf die mittelbaren

Effekte und Nebenwirkungen ihrer Implementation zu blicken, nämlich: »environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves« (ebd.).<sup>18</sup>

In *programmatischer Hinsicht* vermutet Martin Karcher (2021) eine Dauersituation der Prüfung im Zuge der digitalen Transformation. Karcher schließt hier an Ruhloffs Argument der Dauerprüfung an, allerdings mit einem zentralen Unterschied. So nähme nicht die Schlagzahl von Prüfungen als Ereignisse einer Bildungskarriere zu, vielmehr würden diese substituiert durch eine Dauerprüfung im Kontext von (educational) Big Data. Genauer: Durch die beiläufige, sublimale Erhebung von Daten – als Substitut pädagogischer Prüfungen – wird eben diesen Daten Autorität zugebilligt, die eine Schließung pädagogischer Praxis als Ort »hermeneutische[r] Anstrengungen und politische[n] Gestaltungswillen[s]« (ebd., S. 92) mit sich bringe. Eine automatisierte Wissensproduktion als Prüfungsersatz mit »fehlende[r] Gastfreundschaft des neuen Daten-Wissens gegenüber anderen Wissensarten« (ebd., S. 85) beinhaltet somit auch immer die Gefahr einer Deprofessionalisierung bei gleichzeitiger Externalisierung der Prüfkriterien.

Die entgrenzte Dauerhaftigkeit dieses Ersatzes, so Karcher unter Bezug auf Andreas Gelhard weiter, erzeuge allerdings einen paradoxalen Effekt hinsichtlich der Zeitdimension der Prüfung. Waren pädagogische Prüfungen traditionell zugleich Bestandsaufnahme und Prädiktor für zukünftige Prüfungen, so gingen mit deren Dauerhaftigkeit zentrale Merkmale verloren, nämlich der der Entlastung von und Entlassung aus ihnen – also von der Prüfungssituation wie auch den damit verbundenen Erwartungen (ebd., S. 88). Auch, das lässt sich der Dauerhaftigkeit als Problemdimension begeben, erscheint lehrseitig eine Zergliederung vormals zusammenhängender, systematisch geordneter Wissensbestände spätestens seit der Bologna-Reform ein ausgeprägter Affirmationswille zu dazu analogen Prüfungsformaten vorzuliegen, etwa in Form von Portfolios (statt Hausarbeiten),

---

18 In den sog. *critical platform studies* würden solche Effekte unter der Dimension *beyond the platform* verhandelt werden (vgl. Decuypere, 2021). Wegen der kaum zu unterschätzenden Wirkmächtigkeit von und bereits jetzt vorfindlichen Durchsetzung gesellschaftlicher Praxen mit KI scheint deren Thematisierung als Plattform als epistemologische Einschränkung.

formativen Assessments (statt Klausuren am Ende einer Unterrichtsreihe, eines Semesters oder Seminars) und zahlreichen Einzelnoten jedes Studienmoduls (statt vormals gewichtiger Prüfungen nach dem Grund- und Hauptstudium). Das eingangs referierte KMK-Papier legt davon Zeugnis ab.

Es steht zu vermuten, dass aus bisher ungeklärten Motivlagen dauerhafte, automatisierte Prüfungen an der Grenze zur Überwachung nicht nur epistemische Verschiebungen erzeugen (ebd., S. 93), sondern auch durch den steigenden Einfluss von EdTech-Unternehmen auf pädagogische Praxis (vgl. Buck, 2025b, S. 23ff.) ohne große Widerrede salonfähig werden und somit der fortschreitenden Ökonomisierung der Bildung (vgl. Höhne, 2015; Buck, 2023) Vorschub leisten. Ob das allseitige Entlastungsversprechen in Form milderer Korrekturaufwands eine Rolle in der Implementation von KI-Werkzeugen an (Hoch-)Schulen spielt, muss in diesem Zusammenhang erst noch empirisch erforscht werden.

*Paradigmatisch* ist ein abermaliges Erstarken des psychologischen Metakognitionsdiskurses der 1980er-Jahre festzustellen, wie er bereits in den frühen 2000er-Jahren über Franz Weinert und die erste PISA-Studie Aufschwung innerhalb der Erziehungswissenschaft erfuhr. Dieser Befund mag insofern als fortlaufendes Phänomen der Psychologisierung pädagogischer Theorie und Praxis gedeutet werden (vgl. Reichenbach & Oser, 2002), als sie die soziale Dimension von Lern- und Bildungsprozessen zwar nicht komplett ausblendet, wohl aber als nachrangig behandelt und das Lernen tendenziell auf einen vermeintlich solipsistischen Prozess reduziert, auf Lernergebnisse fokussiert und diese (oder ihr Ausbleiben) den Lernenden zuschreibt. Auch im Kontext der Prüfung als triadische Struktur im Anschluss an Herbart ist dieser Reduktionismus problematisch, da er einer technisch-solipsistischen Rationalität pädagogischer Praxis das Wort redet.

Es nimmt nicht wunder, dass auch die KMK (2024) das inzwischen normalisierte Vokabular und Denkbesteck psychologisch-empirischer Forschung anlegt und im Zuge dessen Lernende verstärkt in den Blick nimmt, über

ihnen die Reflexion der eigenen Lernprozesse zusätzlich zum Lerninhalt aufzuerlegen:<sup>19</sup>

»Schülerinnen und Schüler können in ihrem schulischen und außerschulischen Lernen durch KI-Anwendungen, bspw. als persönlichen Lerntutor, als Intelligentes Tutorielles System[sic] oder adaptive Lernumgebung unterstützt werden. Gleichzeitig findet schulische Bildung immer im sozialen Raum und in zwischenmenschlicher Interaktion statt. Deswegen soll der Einsatz von KI-Anwendungen nicht zur Abschwächung des gemeinsamen Lernens führen.« (KMK, 2024, S. 3)

Bildung und Lernen werden im Papier der KMK gleichgesetzt, der unauflösbare Widerspruch zwischen Vereinzelung und dem Wunsch nach sozialem Lernen wird als Nicht-soll-Bestimmung hinausdefiniert. Nicht nur die Antwort, sogar die Frage danach, wie denn die Abschwächung des gemeinsamen Lernens verhindert werden soll, bleibt das Papier schuldig.<sup>20</sup> Es steht paradigmatisch für ein neoliberales Denken, das als bildungspolitisches Programm bereits seit Längerem besteht. So fordert sie im Zuge der Digitalisierung des Bildungsbereichs ein erhöhtes »privates Engagement« (KMK 2016, S. 44) in Form von Public-Private-Partnerships, die sich tatsächlich im Zuge digitaler Transformationen durchzusetzen scheinen, wenngleich mitunter verschleiert und vermittelt durch Stiftungen, Bildungsinitiativen etc. (vgl. Höhne & Striebing, 2020; Förschler, 2021).

Mit diesem Rückzug des Staates aus seinem eigenen Verantwortungsbereich wird zugleich die demokratische Öffentlichkeit um das Recht ihrer

---

19 Der wissenschaftliche Anstand gebietet es offenzulegen, dass sich Teile der folgenden Absätze aus dem Manteltext meiner unveröffentlichten Habilitationsschrift mit dem Titel »Digitalisierende und ökonomisierende Transformationen der Pädagogik« speisen.

20 Besonders das Problem steigender Abhängigkeit von LLM-Outputs für alltägliche Lebensfragen bis hin zur Imagination mit der KI soziale Beziehungen einzugehen, wie es v. a. unter Jugendlichen ein Trend zu sein scheint, wird hier in gefährlicher Weise vernachlässigt. Die Psychologie verwendet dafür den Begriff der *Parasozialität*, pädagogisch gibt es hier deutlichen Nachholbedarf in einer philosophisch wie sozialwissenschaftlich orientierten Theoretisierung und empirischen Erforschung des Phänomens (vgl. Schweiger & Zichy, 2023). Dieses scheint schon dort zu beginnen, wo ChatGPT et al. als Assistenz, Lern-, Schreib- oder Diskurspartner usw. bezeichnet oder von ihren Usern mit Spitznamen (*Chatty* usw.) versehen wird.

Mitbestimmung am Umgang mit öffentlichen Gütern beraubt (lat. *privare*). Digital-ökonomische Transformationen bedeuten demnach *potenziell* und *wahrscheinlich* Ent-demokratisierung, da sie ob ihrer Medialität und Opazität Möglichkeiten des Ein- und Widerspruchs vorgängig und technisch verhindern und somit die demokratische Idee der Deliberation technisch eliminieren. Das bedeutet jedoch nicht, dass *sämtliche* digitalisierende Transformationen des Pädagogischen diese Folge nach sich ziehen. In der Abwendung von der Idee einer relativ isolierten pädagogischen Praxis, ist eine Erklärung dafür zu finden, warum ökonomisierende und digitalisierende Transformationen einerseits relativ konfliktlos miteinander harmonieren und andererseits sich gemeinsam gegenüber anderen gesellschaftlichen Transformationen als vermeintlich *notwendige* behaupten können (vgl. Buck, 2026). Wird bildungspolitisches Handeln auf die Vermittlung von Zielgrößen reduziert, können private Intermediäre und deren Kundinnen – KiTas, Schulen, Volkshochschulen usw. – nicht anders als ihr Handeln an der Erfüllung staatlicher Zielvorgaben, d. h. zweckrational, auszurichten. Der Staat schließlich

»übersetzt transnationale Leistungsstandards (Kompetenzstufen), Qualifikationsstandards (Europäischer und Nationale Qualifikationsrahmen) oder ökonomische Vorgaben wie ›employability‹ zur Akkreditierung von Studiengängen [...] in die jeweiligen Kontexte« (Höhne, 2020, S. 1196)

unterschiedlicher Bereiche des Bildungssystems. Pädagogische Praxis sieht sich somit stärker als zuvor mit dem Problem konfrontiert, ihre eigenen Vorstellungen gelungener Erziehung, anspruchsvoller Inhalte usw. mit der Erreichung einer bestimmten Quote von Kompetenzen und Abschlüssen zu vermitteln. Nicht nur führt das zu dem Phänomen der sogenannten »Noten-inflation« (vgl. exemplarisch Grözinger & Baillet, 2015) und somit zum Sinken der Validität unserer eigenen Maßstäbe, auch steigt die Gefahr stark, dass sich parapädagogische Überformungen pädagogischer Strukturen und Handlungen – wie etwa von Prüfungen – vollziehen. Parapädagogisch sind sie dann, wenn sie ein »Denken und Handeln« repräsentieren, in dem »pädagogische Absichten eine neben- oder gar untergeordnete Rolle spielen« (Mikhail, 2016, S. 211).

Damit ist nicht die Selbstverständlichkeit kritisiert, mit der Brüche zwischen institutionellen, gesellschaftlich-politischen, professionellen und individuellen Normen im pädagogischen Tagesgeschäft dauerhaft ausgehandelt werden. Vielmehr wird mit der Fischer'schen bzw. Ruhloff'schen Figur des *Parapädagogischen* eine heuristische Unterscheidung möglich, die Transformationen auf legitime und illegitime Folgen für pädagogisches Handeln und seine strukturellen Rahmenbedingungen zu befragen erlaubt. Illegitim wären Transformationen dann, wenn sie zu einer Verdrängung des maßgeblich Pädagogischen durch das Parapädagogische führen, also bspw. das Erziehungsziel der Mündigkeit in der Substitution des argumentativen Austauschs durch Überredung verunmöglicht wird. Akte der Manipulation sind schließlich »teleologisch nicht auf die Anerkennung von sittlichen Gültigkeiten gerichtet«, sondern »[haben] die Abhängigkeit des Gegenübers zum Ziel« (ebd., S. 228).

Das vermeintliche Zuvor pädagogischer Praxis in ihrer rhetorischen Vielfalt (graue Theorie, pauken, Kreidezeit, fensterloser Gang durch die Geschichte, Erziehung als Zumutung etc. pp.) wird als Grund für Reformnotwendigkeit behauptet, ohne damit pädagogische Argumente vorzulegen. Wohl aber wird diese neue Form des Sprechens über Pädagogik durch die Praxis und Öffentlichkeit zunehmend affirmiert. Das diskursive Problem des öffentlichen Sprechens über Erziehung – zwischen politischem Handeln, der Ökonomisierung aller Lebensbereiche und dem Umbau des Wissenschaftssystems – zeigt sich auch darin, dass von Erziehung nur noch sehr ungern gesprochen wird, vor allem in der Erwachsenen- und Politischen Bildung (vgl. Fuhr, 2022; Nohl, 2025). Idealerweise überall und stattdessen soll sich pädagogisches Handeln in Formen der *Lernbegleitung* und des *Coaching* manifestieren. Diese Form der Rede über Pädagogik hat sich inzwischen in den für das Lehramt qualifizierenden Studiengängen normalisiert, worunter ein bereits affirmiertes Selbstverständnis von Pädagogik als Dienstleistung verstanden werden kann, nicht als legitime Zumutung (vgl. Reichenbach 2013; Nohl, 2025), die notwendigerweise mitunter normativ-paternalistisch begründet ist (vgl. Drerup, 2016; Fuhr, 2022).

Die durch Digitalisierung und Ökonomisierung angestoßene Transformation des Pädagogischen verbleibt offenbar nicht allein auf der sprachlichen Ebene. Das kommodifizierende und automatisierende Denken in Märkten

und Tauschverhältnissen trägt sich nicht nur in der arbeitsmarktpolitischen Praxis der Zuteilung von Bildungsgutscheinen oder der bildungspolitischen Auflösung von Schulspengeln und daraus folgenden Schaffung von Quasi-Märkten ab (vgl. Radtke & Weiß, 2000). Sie zeigt sich auch in schulischen Praktiken und didaktischer Theoriebildung. In der Schule werden Klassenverträge geschlossen, mit Hilfe derer ein kontraktuelles, transaktionales Denken ein- und ausgeübt wird. Im Lehramtsstudium wird Fends bzw. Helmkes Angebot-Nutzungs-Modell und dessen Weiterentwicklung (vgl. Seidel, 2014) ebenso gelehrt wie Classroom-Management-Modelle. Auch der Einsatz von generativer KI ohne didaktisches Korrektiv ist hierzu zu zählen.

Damit geht nicht nur potentiell eine gewisse Deprofessionalisierung vonstatten, wie sich schultheoretisch plausibilisieren lässt, auch werden bereits *in* Einrichtungen der öffentlichen Erziehung und Bildung Sozialformen eingeübt, die nicht vom Solidaritäts-, sondern vom Wettbewerbsgedanken getragen sind. Diese marktbedingte Denkform des Sozialen zeitigt Effekte, die mit jeder sich inklusiv begreifenden Pädagogik konfligiert (vgl. Buck, 2025b, S. 62ff.), da sie, vermittelt über den Leistungsgedanken und seine meritokratische Manifestation, Abweichung von *marktfähiger* Existenz als pädagogisch unerwünscht markieren muss. Die notwendige Konsequenz daraus ist eine neue Stratifizierung des Bildungssystems (vgl. Helsper, Krüger & Lüdemann, 2019).

Wie Frank-Olaf Radtke bereits 2003 erkannte, hat sich im Zuge der zweiten Empirisierungswelle der Erziehungswissenschaft eine Praxis der Forschung etabliert, die der politischen Ökonomie weitgehend entspricht. Er stellt zum Imperativ der Effizienzsteigerung fest:

»Solange sich in dieser Koalition aus Wirtschaft und Politik (noch) niemand traut, öffentlich eine quantitative Einschränkung des Bildungsangebots zu fordern, bleibt nur die Intervention auf der Kostenseite.« (Radtke, 2003, S. 113)

Folglich kommt es auch in der Erziehungswissenschaft zu einer (empiristischen) Transformation. Es wird solche Forschung honoriert, die die fehlende Kausalität von Intervention und Lernerfolg mit empirischen Modellen in bildungspolitisches Steuerungswissen zu überführen sucht, was angesichts

der prinzipiellen »Nicht-Zurechenbarkeit von Ursachen und Wirkungen« (ebd., S. 125) unserer Praxis jedoch nur scheitern kann.

Im Zuge der zu großen Teilen privatisierten digitalen Transformationen verstärkt sich ein zweckrationales Bild von pädagogischer Praxis als Wettbewerb und marktförmiger Agonalität, wie sich auch an den hier gezeigten phänomenalen Erscheinungsformen und Effekten von KI in Prüfungen an Schulen und Hochschulen zeigen lässt.

## 4 Die Prüfung der Originalität als pädagogisches und demokratisches Projekt

LLMs sind der sprichwörtliche Elefant im pädagogischen Porzellanladen. Sie sind disruptives Moment im Wortsinne, da sie die pädagogische Praxis zur Reorganisation nötigen, ohne produktive Gegenvorschläge zu offerieren. Das gilt nicht nur für die Praxis des Unterrichts (Buck, 2025a; b, S. 107ff.), sondern auch für die der Prüfungen. Jene als Ort der Systematisierung des eigenen Denkens wiederzugewinnen, muss Aufgabe jeder Prüfungskultur sein, die sich als pädagogisch versteht und zu lange Phänomene wie das landläufig bekannte *Bulimie-Lernen* als Vorbereitung auf gehäufte Prüfungen ignorierte. Das hier Dargelegte zeigt, wie weit die pädagogische Dimension der Prüfung über das hinausgeht, was ihr bildungssoziologisch zugesprochen und von ihr erwartet wird. Am Beispiel der schulischen Anforderungsbereiche (AFB) soll hier skizziert werden, wie sich der Grenzbereich des Einsatzes von KI in Prüfungen kartographieren lässt.

Größtenteils unbedenklich scheint der Einsatz von KI zur Prüfung von reproduktiver (AFB 1) und reorganisierender (AFB 2) Lern- und Denkleistung, sofern Datenschutz und informationelle Selbstbestimmung der Prüflinge (vgl. Pötzsch & Buck, 2023) sichergestellt sind. Kenntnisse wie Vokabeln, termini technici und Jahreszahlen wie auch Fertigkeiten des Feilens, Blutabnehmens und Pipettierens erfordern keine interpretative Denkleistung – sie sind schlichtweg einzuüben und können anhand eindeutiger Kriterien in ihren Ergebnissen gemessen werden. Am AFB 3 jedoch, der Reflexion und Problemlösung abprüft, zeigt sich die Grenze der Leistungsfähigkeit künstlicher Intelligenz. Dort, wo individuelle und biographische Kontexte in die pädagogische Beurteilung der Leistung einfließen, wo der Prozess des Denkens in individueller Hinsicht geprüft wird, kann finalistische Objektivität kein gangbarer Maßstab für die Differenzierung von Denkleistungen sein.

In der Differenz von Text und Kontext ist die hermeneutische Grundlegung des pädagogischen Urteils zu finden. Während LLMs stets größere

Datenmengen in ihren Trainingskorpus aufnehmen, können sie keine Relationierung vornehmen zwischen Aussagen und ihren biographischen und individuellen Voraussetzungen. Für eine Bildungsaufsteigerin, die zeitweise in der Jugendwohngruppe lebte und im häuslichen Umfeld mit Armut, Gewalt und Substanzmissbrauch konfrontiert ist, bedeutet die Abiturprüfung eine weitaus größere Leistung als für privilegierte Mitschülerinnen und Mitschüler, die wohlbehütet aufwuchsen und sich unter Zuhilfenahme von Nachhilfe haben vorbereiten können. Diese Aspekte einer Leistungsbeurteilung vor dem Hintergrund der jeweiligen Biographien sind, das ist keine steile These, institutionalisierter Teil lehrseitiger Praxis. Künstliche Intelligenz aber bleibt für das zu prüfende Individuum und seine Lebensgeschichte (vgl. Fuchs, 2011) blind, sie verfügt allein über Text ohne Kontext. Gerade *weil* Objektivität demnach in pädagogischen Prüfungen des Reflexionsvermögens und vor dem Hintergrund des Geschilderten unmöglich zu erreichen ist, ist ein Maßstab anzulegen, der die hermeneutischen Fähigkeiten menschlicher Prüferinnen und Prüfer nutzt, um die subjektive Dimension der Prüfung als Entfaltungsmöglichkeit der *Originalität* zu ihrem Recht kommen zu lassen. Die Beurteilung von Prüfungen wäre nicht aus dem Output abzuleiten, sondern aus den (prozessualen) Bildungserfahrungen der Prüflinge, was im Folgenden skizziert werden soll.

Dieser Maßstab wäre dann weder als objektiv noch als subjektiv, sondern als *fair* zu kennzeichnen. Faire Prüfungen zertifizieren keinen Glauben (in den *richtigen* Denkstil oder das *richtige* Paradigma), sie testen auch nicht das Gewissen oder die Ehrlichkeit (aus Anlass ausgewiesener oder verschwiegener KI-Nutzung), sondern suchen individuelle Denkfähigkeit anlässlich eines gegenständlichen Problems zu prüfen. Somit bewehrt sich eine pädagogische Idee der Prüfung gegen eine solche der Passförmigkeit, sondern fixiert individuelle, kontextabhängige Denkleistung als ihren Maßstab.<sup>21</sup>

---

21 Unstrittig scheint es, am Leistungsbegriff im Sinne einer Überwindung innerer (Denk-)Widerstände als Prüfungskriterium festzuhalten, dieses demnach nicht außerpädagogischen Anforderungen auszuliefern. Zur Nerowski-Berkemeyer-Tenorth-Kontroverse über Leistung als Kriterium der Bildungsgerechtigkeit vgl. Nerowski, 2017; Berkemeyer, 2020a; Tenorth, 2020; Berkemeyer, 2020b; Nerowski 2020. Die notwendige Konsequenz, Denkleistungen nur als Performanz, niemals

Äquivalent zum Unterricht lassen sich pädagogische Prüfungen mit Johann Friedrich Herbart als triadische Struktur ausweisen, deren Ziel die Würdigung einer im Wissen begründeten und begründbaren Haltung zu einem Thema ist und in denen die »Gelenkigkeit des Denkens« (Herbart, 2022, S. 134) der Geprüften zur Geltung kommt. Das bedeutet selbstredend keine Forderung nach einer Rückkehr zu einer Prüfungspraxis, die auf die Bekundung der Leistungsfähigkeit durch Arrivierte reduziert wird, wie es lange Zeit in der Vergabe von Doktorgraden Praxis war, bevor sie im Prozess der Disputatio in die Gemeinschaft der Scholaren aufgenommen wurden (vgl. Reble, 2004, S. 63 f.).

Wenn sich nun plausibel behaupten lässt, dass ein externer, objektiver Maßstab weder möglich noch erwünscht ist für die Erteilung eines pädagogischen Urteils im Rahmen einer Prüfung, so lässt sich daraus folgern, dass nur eigene Kriterien dem Bildungsanspruch von Schulen und Hochschulen zu genügen vermögen. Analog zum Vorschlag *genetischen Zeigens* als Praxis einer Lehre, die die Genese von Geltung bzw. Wahrheit thematisierten Wissens in der Schule lehr- und lernseitig ein- und ausübt (Buck, 2025a), ist die Form eines *genetischen Prüfens* zu entwickeln, die auf das Originalitätskriterium im Denken der Geprüften abhebt.

Der Originalitätsgedanke findet sich an mehreren Stellen in Wilhelm von Humboldts Schriften.<sup>22</sup> In dessen 1792 erschienenen Text mit dem sprechenden Titel »Wie weit darf sich die Sorgfalt des Staats um das Wohl seiner Bürger erstrecken?« (Humboldt, 1988), später ausgebaut zu den »Ideen zu einem Versuch die Grenzen der Wirksamkeit des Staats zu bestimmen« (Humboldt, 1980c), aus der die berühmte Formel des wahren Zweckes des Menschen als »höchste und proportionirlichste Bildung seiner Kräfte zu

---

als Kompetenz ausweisen zu können, mahnt an das bereits fortgeschrittene Denken in »Kausalitätsmythen in der Erziehung« (Radtke, 2003, S. 117).

22 Eine exakte chronologische Rekonstruktion des Humboldt'schen Wortgebrauchs schuldig bleibend, wird hier die vorsichtige Vermutung entfaltet, dass sich in der Transformation der Nomenklatur von der *Eigentümlichkeit* zur *Originalität* ein Bildungsprozess Wilhelm von Humboldts nachzeichnen lässt, der auf die intensive Lektüre französischsprachiger Literatur (u.a. Jean-Jacques Rousseaus, mit dem sich Humboldt die Liebe zur Sprachphilosophie teilte) zurückzuführen ist.

einem Ganzen« (Humboldt 1988, S. 242) stammt, vermerkt Wilhelm von Humboldt zum »bildenden Nutzen« (ebd., S. 243) von Freundschaft:

»Denn wenn ohne diese Innigkeit der eine den ändern nicht genug aufzufassen vermag, so ist die Selbstständigkeit nothwendig, um das Aufgefasste gleichsam in das eigne Wesen zu verwandeln. Beides aber erfordert Kraft der Individuen, und eine Verschiedenheit, die, nicht zu gross, damit einer den ändern aufzufassen vermöge, auch nicht zu klein ist, um einige Bewundrung dessen, was der andre besitzt, und den Wunsch rege zu machen, es auch in sich überzutragen. Diese Kraft nun und diese mannigfaltige Verschiedenheit vereinen sich in der *Originalität*, und das also, worauf die ganze Grösse des Menschen zuletzt beruht, wonach der einzelne Mensch ewig ringen muss, und was der, welcher auf Menschen wirken will, nie aus den Augen verlieren darf, ist Eigentümlichkeit der Kraft und der Bildung. Wie diese Eigenthümlichkeit durch Freiheit des Handelns und Mannigfaltigkeit des Handelnden gewirkt wird; so bringt sie beides wiederum hervor.« (Ebd., S. 244; Hervorhebung durch MFB)

In dieser *Eigentümlichkeit* findet sich der Ausdruck von Bildung, stets gebunden an sinnliche Empfindung einerseits und Selbsttätigkeit andererseits. Im Gegensatz dazu war Humboldt das Einförmige, Uniforme suspekt. So formuliert er mit unverkennbarem Bezug dazu in seinem »Fragment zur Theorie der Bildung des Menschen«:

»Allein nur, indem man dies schrittweise verfolgt und am Ende im Ganzen überschaut, gelangt man dahin, sich vollkommne Rechenschaft abzulegen, wie die Bildung des Menschen durch ein regelmässiges Fortschreiten Dauer gewinnt, ohne doch in die Einförmigkeit auszuarten, mit welcher die körperliche Natur, ohne jemals etwas Neues hervorzubringen, immer nur von neuem dieselben Umwandlungen durchgeht.« (Humboldt, 1980a, S. 240)

An anderer Stelle, in der Schrift »Über den Geist der Menschheit«, iteriert Humboldt die Idee, dass ein *innerer Maßstab* zur Beurteilung von Welt untrennbar mit gelingender Bildung verbunden sei. Nur eine »entschiedene und originelle Individualität« (Humboldt, 1980b, S. 512) ermögliche Vollkommenheit.

In dieser *Originalität* – nicht zu verwechseln mit reduzierten Identitätskonzepten – wären demnach Bildungserfahrungen zu suchen, die einer schriftlichen oder mündlichen Prüfung einen bildungstheoretisch fundierten Anspruch verleihen können, der nicht einfach im Sinne des Anforderungsbereich 3 den Transfer oder die Anwendung von Wissen abprüft, sondern im Herbart'schen Dreieck den Prüfungsgegenstand mit Prüfling und Prüfendem zu relationieren versteht und diesen erlaubt, ihrer Originalität im Denken Ausdruck zu verleihen.

Nur dort zeigt sich, dass und wie menschliche Denkleistung das übersteigt, wozu LLMs aufgrund ihrer Funktionsweise der nach Wahrscheinlichkeiten operierenden Zusammenstellung vorhandener Texte nicht imstande sind. Diese können entgegen den Behauptungen ihrer Schöpfer nicht *denken*, auch können sie, wie oben beschrieben, keinen individuellen Kontext berücksichtigen und folglich keine hermeneutischen Akte vollziehen – Rora et al. (2025) sprechen treffend von »künstlicher Verstehenssimulation«. Sie können abstrakte Ideen nicht auf individuelle lebensweltliche Probleme beziehen oder in lebensweltnaher Sprache ausdrücken, die nicht Teil des Trainingskorpus ist.<sup>23</sup> Aber: Menschen und somit auch Prüflinge können es.

Prüfungen, die darauf abzielen, Denkbewegungen und Argumente nicht einfach nur zu reproduzieren, sondern das *Denken* erst selbst zu lernen und einzuüben und später abzuprüfen, wären solche, denen gegenüber vermeintlich objektiven Prüfungen das Prädikat *fair* anhaftet.<sup>24</sup> Faire Prüfungen

---

23 Diese Einsicht zieht zwei mögliche Umgänge damit nach sich: 1) die radikale Datafizierung des Lebens und die Bereitstellung dieser Daten für KI oder 2) der Verzicht auf KI-Nutzung in Fragen der Erörterung, Abwägung, Beurteilung, Interpretation. Die erste Variante entspricht den politischen Fantasien der Silicon-Valley-Oligarchen bzw. sog. Tech-Bros großer und politisch einflussreicher Tech-Unternehmen, die zweite verweist auf das Erbe der Aufklärung als un abgeschlossenes Projekt der Moderne (vgl. Habermas, 1994) und die anthropologische Differenz als Denkfigur.

24 Eine solche Anmerkung kommt selbstredend nicht ohne einen Verweis auf Martin Heideggers Vorlesung aus dem Wintersemester 1951/52 mit dem Titel »Was heißt denken?« (Heidegger, 2002) aus. Heidegger skizziert darin zunächst, dass das Denken gelernt werden müsse, da es sich weder aus der Denkfähigkeit noch aus der Begabung des Menschen notwendigerweise ableiten ließe. Entlang der *Vorstellung* als Denkakt entfaltet Heidegger mit Hilfe der nietzeanischen Figur des Übermenschen überpositivistische Denkbewegungen, die Ontos, Zeit und

diskriminieren nicht zwischen überzeugten KI-nutzenden Prüflingen und solchen, die primär auf ihre eigene Urteils- und Artikulationsfähigkeit vertrauen. Vielmehr sind sie so gestaltet, dass sie das für Prüflinge Denkwürdige zum Thema ihrer Prüfung machen. Der Maßstab ihrer Beurteilung wäre demnach kein objektiver – als Einfallstor für diskriminierende Urteile entlang normativ aufgeladener Maßstäbe wie bspw. einem bestimmten akademischen Habitus –, sondern ein intersubjektiv-individueller, der die Fähigkeit des Denkenkönnens in Abhängigkeit vom Ausgangspunkt der Prüflinge bemisst. Eine solche Prüfungsform erfordert ein hohes Maß an Fachkenntnis, Wissen um diskriminierende Praktiken im Kontext der Schule (und wie man diese vermeidet), ein technisches Verständnis der Effekte und Limitationen von KI sowie Empathie für seine Prüflinge und deren Lebenslage. Sie verweigert sich einer digitalen Überformung von Prüflingen als imaginiertem und standardisiertem Ideal (vgl. Walgenbach, 2023), sondern nutzt ihr fachliches Verständnis, um im Modus wohlwollender Interpretation den Prüflingen zur Entäußerung ihrer individuellen Denkfähigkeit zu verhelfen. Kurzum: In ihr kulminiert ein Professionsverständnis wie es für das Lehramt so lange gültig ist, wie es sich als *erziehende* Profession versteht (vgl. Rucker, 2024) und ihren fachlichen Anspruch nicht einer Standardisierung und Automatisierung preisgibt.

Wie auch in Unterricht und Lehre ein enger Zusammenhang von Gegenstand und Methode besteht, lassen sich Prüfungen mit dem Anspruch des Fairen nur in ausgewiesener Kenntnis der jeweiligen Fachdidaktik durchführen. Die Prüfung ist, wie auch die Lehre, eine Kunstform, die die Antinomien und Paradoxien schulischer Praxis (vgl. Helsper, 2016) nicht ignoriert, sondern anerkennt und, wo möglich, produktiv wendet. In Hinblick auf die Prüfung scheint die *Sachantinomie* als vordergründiges Problem, die Lehrkräfte zur Vermittlung von Sachinhalt und „konkrete[n], spezifische[n], partikuläre[n] Hintergründe[n], Erfahrungen und Lebensbedingungen“ (ebd., S. 56) auffordert.

---

Anerkennung bzw. Andenken des *Anwesenden* berücksichtigen. So komplex Heideggers darin entfaltete Gedanken sind und so sehr sie über die Erwartungshorizonte der meisten Prüfungen hinausgehen, so wertvoll ist dessen darin enthaltener antiessentialistischer und zugleich fundamentalontologischer Impetus, der für *faire* Prüfungen geradezu notwendig scheint.

Faire Prüfungen, wie sie hier vorgeschlagen werden, verabschieden sich vom ohnehin nicht erreichbaren Objektivitätskriterium der Prüfung, ohne es zugleich für das Objektivitätskriterium des Wissens zu tun. Stattdessen erfolgt eine Berücksichtigung der zu prüfenden Person inklusive ihrer Lebens-, Lern- und Bildungswege. Eine Achtung des Kontextes, innerhalb dessen Prüfungen abgenommen werden, erinnert daran, dass im Dreieck der Prüfung immer zugleich Wissen, Prüfling und Prüfender geprüft werden. Das trifft selbstredend auch auf solche Prüfungen zu, die in erster Linie die Reproduktion von Wissen und seine Relationierung (gemeinhin: Anforderungsbereiche 1 und 2) zum Gegenstand haben, wenngleich im geringeren Maße. Ein Prüfling wie der oben zitierte 18-jährige Schüler, der zwar ohne den Gebrauch gängiger Fachbegriffe, aber mit eigenen Worten bspw. ethische Prinzipien zu präsentieren und kritisieren weiß, wäre nach objektiven Maßstäben weniger gut zu bewerten als ein Mitschüler, der mit Hilfe von ChatGPT Grundbegriffe auswendig gelernt hat, diese aber nur mit Hilfe auswendig gelernter Phrasen zu erklären vermag. In einer fairen *Prüfung* hingegen würde das Denkenkönnen des Prüflings als wertvoller und wünschenswerter erachtet, weil es auf das Durchdringen des Stoffs abhebt, ohne die Notwendigkeit präziser Fachsprache in Abrede zu stellen.

Faire pädagogische Prüfungen sind Situationen der Dauerhermeneutik. Schriftliche und mündliche Äußerungen von Prüflingen werden im Sinne wohlwollender Interpretation von Prüfenden interpretiert und beeinflussen so den Prüfungsverlauf und die abschließende Beurteilung. Nehmen wir als Prüfende unsere Aufgabe ernst, eröffnen wir in ihnen jenen Prüflingen Chancen, die aufgrund von Migrationsgeschichten, Disability, Bildungsaufstieg usf. nicht der normierten Vorstellung einer Schülerin oder eines Studenten entsprechen (vgl. Buck, 2025b, S. 62 ff.). Vielmehr wären sie deswegen *genetisch*, weil sie den Bildungsprozess Einzelner als Leistung einer entwickelten Haltung zur Welt als Ergebnis von Denkprozessen berücksichtigen. Prüfungen können somit zentrale Orte sein, an denen Bildungsinstitutionen zur Chancengerechtigkeit beitragen können, indem sie den inneren Maßstab der Bildung im Sinne Humboldts freizulegen versuchen. Möglicherweise schmilzt in dieser Erkenntnis das Versprechen der Zeitersparnis durch KI-Gebrauch. Hermeneutik ist situativ und sehr zeitintensiv, aber idealiter eben auch eines: fair.

Faire Prüfungen beinhalten somit einen dezidiert demokratischen Aspekt. Sie imaginieren keine Gesellschaft, die normierte Ausdrucksformen hegemonial und über das Prüfungswesen des Bildungssystems legitimiert reproduzieren – in radikaler Lesart eine Form des *Social Engineering* (vgl. Buck, 2015). Stattdessen honorieren sie die Eigentümlichkeit und Originalität jeder und jedes Einzelnen für das Ziel einer gebildeten Gesellschaft bzw. des *Menschengeschlechtes*, wie es Humboldt imaginierte.

Auch Werner Loch führt ein politisches Argument an, wenn er die Zunahme von Prüfungen in modernen Gesellschaften als Ausweis von deren Demokratisierung verstanden wissen möchte:

»Wenn Bildung jedem zugänglich und zur entscheidenden Voraussetzung für den beruflichen und sozialen Aufstieg wird, gewinnt die Prüfung dieser Bildung bzw. der darauf beruhenden Tüchtigkeit zwangsläufig eine erhöhte Bedeutung als bestätigende und auslesende Funktion und verstärkt sich entsprechend die Tendenz, sie in Institutionen zu befestigen. In Kulturen und Epochen, wo demgegenüber die soziale Herkunft über den Beruf und den gesellschaftlichen Rang, den der Nachwuchs einnehmen wird, entscheidet, muß die Prüfung an Bedeutung verlieren.« (Loch, 1963, S. 192)

Dieses Argument ist freilich nur dann statthaft, wenn in pädagogischen Prüfungen tatsächlich Genanntes geprüft wird und sich nicht eine Selektion nach habitueller Ähnlichkeit vollzieht – wir also im Vollzug der Prüfung keiner *Meritokratiellusion* aufsitzen, wie es Heike Solga (2005) pointiert. Abermals liegt die pädagogische Aufgabe im Einbezug des individuellen Kontextes der Person.

Diese Mahnung wird eindringlich vor dem Hintergrund der Diskriminierungsforschung der vergangenen Jahrzehnte und möglicherweise umso notwendiger, je mehr LLMs sprachliche Ausdrucksfähigkeit von Prüflingen nicht kompensieren, sondern verschleiern und somit die Prüfung individueller Leistung angesichts ungleicher Startchancen, Lebenslagen und Gelinensbedingungen erschwert oder gar verunmöglicht.<sup>25</sup> Die Verschleierung

---

25 Es ist nicht so, dass es nicht (zirkulär argumentierende) Gegenvorschläge für mehr Technisierung gäbe, etwa dahingehend, dass schriftliche Prüfungen

von habituellen Unterschieden im sprachlichen Ausdruck durch die Verwendung KI-generierter Texte bleibt eine Illusion von sozialer Angleichung und die willfährige Affirmation dessen, was Theodor W. Adorno (2003) als Halb- bildung auswies: Die »[f]risch-fröhliche Verbreitung von Bildung unter den herrschenden Bedingungen ist unmittelbar eins mit ihrer Vernichtung« (ebd., S. 110), wohingegen *genetisches Zeigen* und *genetisches Prüfen* (lediglich) einen Versuch des Widerstandes darstellen. Sie verweigern sich eines »auf den Konformismus vereidigt[en]« (ebd., S. 116) Geistes der Halbbildung.

Nicht nur ihre reproduktiven, auch die angedeuteten sozialen Effekte gegenseitigen Misstrauens durch die Verbreitung generativer KI bedürfen im Zuge einer Neubesinnung auf pädagogische Prüfungen einer gründlichen Erforschung; Gleiches gilt für leibliche Effekte jenseits psychologisierter Phänomene wie Prüfungsangst. Pädagogische Prüfungen – ohne die Möglichkeit der Kritik ihrer exklusiven Effekte zu negieren – lassen sich hingegen unter Einbezug der genannten Argumente einer handlungstheoretischen Theoretisierung überführen, die das Pädagogische im Blick behält und gegenüber Überformungsversuchen jedweder Art zu verteidigen versteht.

Eine auf *Originalität* abhebende Praxis der genetischen Prüfung, die den Anspruch der *Fairness* anstelle eines solchen der Objektivität zu verwirklichen sucht, markiert eine doppelte Differenz zu KI und der ihr zugrundeliegenden digital-ökonomischen Rationalität. Der Gegenvorschlag bezieht theoretische und historische Bestände der Erziehungswissenschaft ein, die es aus sachhaltigen Gründen zu pflegen gilt. Eine solche Wiedergewinnung der Prüfung als spezifisch pädagogische Handlung jedenfalls scheint überfällig, in der Praxis wie in der Theorie.

---

zukünftig *innerhalb* von LLM-Umgebungen zu verfassen sind, um von der KI eine Bewertung eigenständiger Leistung vornehmen zu lassen. Diesen Hinweis verdanke ich Thomas Koinzer.

## Literaturverzeichnis

- Addey, C.; Sellar, S.; Rutkowski, D.; Gorur, R.; Waldow, F.; Takayama, K.; Grek, S. (2025). International Large-Scale Assessments (ILSAs): what have they done for education? In: *Discourse. Studies In the Cultural Politics of Education* 46, S. 1–14.  
<https://doi.org/10.1080/01596306.2025.2578289>.
- Adorno, T. W. (2003). Theorie der Halbbildung. In: ders.: *Gesammelte Schriften, Bd. 8. Soziologische Schriften I*. Frankfurt/Main: Suhrkamp.
- Andresen, S; Casale, R.; Gabriel, T.; Horlacher, R.; Larcher Klee, S.; Oelkers, J. (Hrsg.) (2009). *Handwörterbuch Erziehungswissenschaft*. Weinheim: Beltz.
- BBC (2025). *Watch: Pope Leo to US students: Don't let AI do your homework*  
<https://web.archive.org/web/20251202084838/>  
<https://www.bbc.com/news/videos/c99704g0x4ro> (Memento vom 02.12.2025).
- Bedorf, T. & Hermann, S. (Hrsg.) (2016). *Das soziale Band. Geschichte und Gegenwart eines sozialtheoretischen Grundbegriffs*. Frankfurt/Main: Campus.
- Bellmann, J. & Weiß, M. (2009). Risiken und Nebenwirkungen Neuer Steuerung im Schulsystem. Theoretische Konzeptualisierung und Erklärungsmodelle. In: *Zeitschrift für Pädagogik* 55(2), S. 286–308.
- Benner, D./Oelkers, J. (Hrsg.) (2010). *Historisches Wörterbuch der Pädagogik*. Studienausgabe. Weinheim: Beltz.
- Berkemeyer, N. (2020a). Über die Schwierigkeit, das Leistungsprinzip im Schulsystem gerechtigkeits-theoretisch zu begründen. In: *Zeitschrift für Erziehungswissenschaft* 23, S. 447–464. <https://doi.org/10.1007/s11618-020-00932-2>.
- Berkemeyer, N. (2020b). Weniger Dissens als man meinen könnte. Ergänzungen zu Elmar Tenorths Metakritik. In: *Zeitschrift für Erziehungswissenschaft* 23, S. 451—454. <https://doi.org/10.1007/s11618-020-00934-0>.
- Bichler, T. & Brauner, K. (2007). Prüfung. In: Lenzen, D. & Rost, F. (Hrsg.): *Pädagogische Grundbegriffe*. Bd. 2: Jugend—Zeugnis. 8. Aufl. Reinbek: Rowohlt. S. 1285–1289.
- Bilstein, J. (1990). Kitsch in der Pädagogik. In: *Neue Sammlung* 30, S. 419–435.

- BLLV [=Bayerischer Lehrer- und Lehrerinnenverband] (2025). *KI-Nutzung an Schulen erfordert anderen Leistungsbegriff*. <https://web.archive.org/web/20251202085226/https://www.bllv.de/vollstaendiger-artikel/news/ki-nutzung-schulen-erfordert-anderen-leistungsbegriff-6713> (Memento vom 02.12.2025).
- Böhm, W. & Seichter, S. (2025). *Wörterbuch der Pädagogik*. 19. Aufl. Paderborn: Schöningh/UTB.
- Bourdieu, P. & Passeron, J.-C. (1971). *Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Klett.
- Buck, M. F. (2012). *Einführung in die biographische Erziehungstheorie Werner Lochs*. Norderstedt: BOD.
- Buck, M. F. (2015). Erziehung, Kompetenzorientierung und Social Engineering. In: Krause, S. & Breinbauer, I. M. (Hrsg.): *Im Raum der Gründe. Einsätze theoretischer Erziehungswissenschaft IV*. Würzburg: Königshausen & Neumann. S. 247–265.
- Buck, M. F. (2023). *Ökonomisierung der Bildung. Eine Einführung*. Weinheim: Beltz. <https://doi.org/10.25656/01:29138>.
- Buck, M. F. (2025a). Pädagogische Verantwortung: Generative KI, Irreversibilität und Genetisches Zeigen. In: *Vierteljahrsschrift für wissenschaftliche Pädagogik* 101(2), S. 179–196. <https://doi.org/10.30965/25890581-10102003>.
- Buck, M. F. (2025b). *Digitalisierung der Bildung. Eine Einführung*. Stuttgart: Kohlhammer.
- Buck, M. F. (2026). Fungibilität, Transaktionalität, Binarität. Zur Konfluenz von ökonomisierenden und digitalisierenden Transformationen der Bildung. In: *Bildung und Erziehung* 79(1), S. 55–72.
- Burghardt, D.; Krinninger, D.; Seichter, S. (Hrsg.) (2015). *Pädagogischer Takt. Theorie – Empirie – Kultur*. Paderborn: Schöningh.
- Brinkmann, M. (2009). Fit für PISA? Bildungsstandards und performative Effekte im Testregime. Vorschläge zur theoretischen und pädagogischen Differenzierung von Bildungsforschung und Aufgabenkultur. In: Bilstein, J. & Ecarius, J. (Hrsg.): *Standardisierung – Kanonisierung. Erziehungswissenschaftliche Reflexionen*. Wiesbaden: Springer VS. S. 97–116. [https://doi.org/10.1007/978-3-531-91726-9\\_5](https://doi.org/10.1007/978-3-531-91726-9_5).
- Cornelsen (2024). *Wie Künstliche Intelligenz das Lernen und Prüfen verändert*. <https://web.archive.org/web/20251202085550/https://www.cornelsen.de/magazin/beitraege/alternative-pruefungsformate-ki> (Memento vom 02.12.2025).

- DLF (2023). *Uni in Tschechien schafft wegen KI Bachelorarbeiten ab*. <https://web.archive.org/web/20250504053502/https://www.deutschlandfunk.de/wegen-ki-erste-uni-in-tschechien-schafft-bachelorarbeiten-ab-dlf-e693e4e5-100.html> (Memento vom 04.05.2025).
- Decuyperre, M. (2021). Decuyperre, Mathias (2021): »The Topologies of Data Practices: A Methodological Introduction«. In: *Journal of New Approaches in Educational Research* 10(1), S. 1–17. <https://doi.org/10.7821/naer.2021.1.650>.
- Drerup, J. (2016). (Re-)Konstruktion praxisinhärenter Normen. Zur Eigenstruktur pädagogischer Rechtfertigungsverhältnisse. In: *Zeitschrift für Pädagogik* 62(4), S. 531–550. <https://doi.org/10.25656/01:16830>.
- Engelhard, K. von (2013). Das Gutachten zur Reifeprüfung. Befunde aus dem Archiv eines West-Berliner Gymnasiums (1890-1970). In: *Zeitschrift für Pädagogik* 69(2), S. 171–185. <https://doi.org/10.3262/ZP2302171>.
- Feldmann, M.; Rieger-Ladich, M.; Voß, C.; Wortmann, K. (Hrsg.) (2024). *Schlüsselbegriffe der Allgemeinen Erziehungswissenschaft. Pädagogisches Vokabular in Bewegung*. 2. Aufl. Weinheim/Basel: Beltz Juventa.
- Fend, H. (2008). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen*. 2. Aufl. Wiesbaden: VS. <https://doi.org/10.1007/978-3-531-91788-7>.
- Fink, E. (1970). *Erziehungswissenschaft und Lebenslehre*. Freiburg: Rombach.
- Förschler, A. (2021). Der wachsende politische Einfluss privater (EdTech-) Akteure im Kontext digitaler Bildungsbeobachtung und -steuerung. Bemühungen um ein ‚dateninfrastrukturfreundliches Ökosystem‘. In: *Zeitschrift für Pädagogik* 67(3), S. 323–337. <https://doi.org/10.3262/ZP2103323>.
- Fuchs, T. (2011). *Bildung und Biographie. Eine Reformulierung der bildungstheoretisch orientierten Biographieforschung*. Bielefeld: transcript.
- Fuhr, T. (2022). Erwachsenenbildung erzieht? Zur Notwendigkeit einer erwachsenenpädagogischen Bereichsethik. In: *Debatte. Beiträge zur Erwachsenenbildung* 5(2), S. 115–128. <https://doi.org/10.3224/debatte.v5i2.03>.
- Gelhard, A. (2018). *Kritik der Kompetenz*. 3. Aufl. Zürich: diaphanes.
- Giesecke, H. (2015). *Pädagogik als Beruf. Grundformen pädagogischen Handelns*. 12. Aufl. Weinheim/Basel: Beltz.
- Gröger, J.; Behrens, F.; Gailhofer, P.; Hilbert, I. (2025). *Umweltauswirkungen künstlicher Intelligenz*. Hamburg: Greenpeace. <https://web.archive.org/web/20251004190532/https://www.greenpeace.de/publikationen/20250514-greenpeace-studie-umweltauswirkungen-ki-pdf> (Memento vom 04.10.2025).

- Grözinger, G. & Baillet, F. (2015): Gibt es auch beim Abitur eine Noteninflation? Zur Entwicklung der Abiturnoten als Hochschulzugangsberechtigung – eine Darstellung und Analyse aus soziologischer Perspektive. In: *Bildung und Erziehung* 68(4), S. 473–494. <https://doi.org/10.7788/bue-2015-0407>.
- Habermas, J. (1994). *Die Moderne – ein unvollendetes Projekt. Philosophisch-politische Aufsätze 1977–1992*. Leipzig: Reclam.
- Hartmann, M. (2002). *Der Mythos von den Leistungseliten. Spitzenkarrieren und soziale Herkunft in Wirtschaft, Politik, Justiz und Wissenschaft*. Frankfurt/Main: Campus.
- Haverkamp, H. (2024). *Wie sich Prüfungsaufgaben jetzt verändern müssen*. In: Deutsches Schulportal, <https://web.archive.org/web/20251202084424/https://deutsches-schulportal.de/expertenstimmen/ki-und-schule-wie-sich-pruefungsaufgaben-jetzt-veraendern-muessen/> (Memento vom 02.12.2025).
- Heidegger, M. (2002). Was heisst Denken? In: ders.: *Gesamtausgabe, I. Abteilung: Veröffentlichte Schriften 1910–1976, Bd. 8*. Hrsg. v. P.-L. Coriando. Frankfurt/Main: Klostermann.
- Helsper, W. (2016). Antinomien und Paradoxien im professionellen Handeln. In: Dick, M.; Marotzki, W.; Miege, Harald (Hrsg.): *Handbuch Professionsentwicklung*. Bad Heilbrunn: Klinkhardt/UTB. S. 50–61.
- Helsper, W.; Krüger, H.-H.; Lüdemann, J. (Hrsg.) (2019). *Exklusive Bildung und neue Ungleichheit. Ergebnisse der DFG-Forschergruppe „Mechanismen der Elitebildung im deutschen Bildungssystem“*. Zeitschrift für Pädagogik, 65. Beiheft.
- Herbart, J. F. (1982). Die ersten Vorlesungen über Pädagogik (1802). In: ders.: *Pädagogische Schriften, Bd. 1. Kleinere pädagogische Schriften*. Hrsg. v. W. Asmus. 2. Aufl. Stuttgart: Klett-Cotta. S. 121–131.
- Herbart, J. F. (2022). Allgemeine Pädagogik aus dem Zweck der Erziehung abgeleitet. In: Ders.: *Joh. Friedr. Hartbart's sämtliche Werke, Bd. 2*. Hrsg. v. K. Kehrbach. Berlin: De Gruyter. S. 1–211. <https://doi.org/10.1515/9783112686843>.
- Höhne, T. (2015). *Ökonomisierung und Bildung. Zu den Formen ökonomischer Rationalisierung im Feld der Bildung*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-08974-0>.
- Höhne, T. (2020). Ökonomisierung im schulischen Feld. In: Bollweg, P./Buchna, J./Coelen, T./Otto, H.-U. (Hrsg.): *Handbuch Ganztagsbildung*. Wiesbaden: Springer VS. S. 1189–1204. [https://doi.org/10.1007/978-3-658-23230-6\\_88](https://doi.org/10.1007/978-3-658-23230-6_88).
- Höhne, T. & Striebing, C. (Hrsg.) (2020). *Stiftungen im Schulwesen*. Weinheim/Basel: Beltz.

- Humboldt, W. von (1980a). Theorie der Bildung des Menschen. In: ders.: *Werke in fünf Bänden I. Schriften zur Anthropologie und Geschichte*. Hrsg. v. A. Flitner und K. Giel. 3. Aufl. Stuttgart: Klett Cotta. S. 234–240.
- Humboldt, W. von (1980b). Über den Geist der Menschheit [1797]. In: ders.: *Werke in fünf Bänden I. Schriften zur Anthropologie und Geschichte*. Hrsg. v. A. Flitner und K. Giel. 3. Aufl. Stuttgart: Klett Cotta. S. 506–518.
- Humboldt, W. von (1980c). Ideen zu einem Versuch, die Gränzen der Wirksamkeit des Staates zu bestimmen [1792]. In: ders.: *Werke in fünf Bänden I. Schriften zur Anthropologie und Geschichte*. Hrsg. v. A. Flitner und K. Giel. 3. Aufl. Stuttgart: Klett Cotta. S. 56–233.
- Humboldt, W. von (1988). Wie weit darf sich die Sorgfalt des Staats um das Wohl seiner Bürger erstrecken? In: ders.: *Wilhelm von Humboldts gesammelte Werke. Zweiter Band*. Nachdruck der Ausgabe von 1841. Berlin/New York: De Gruyter.
- Illich, Ivan (1971). *Deschooling Society*. San Francisco: Harper & Row.
- Ingenkamp, K. (1989). *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte*. 8. Aufl. Weinheim/Basel: Beltz.
- Karcher, Martin (2021). Das Ende der Prüfung als Anfang des ‹Data Realism›? Erkundungen zur Digitalisierung pädagogischen Wissens. In: *MedienPädagogik 45*, S. 80–97.  
<https://doi.org/10.21240/mpaed/45/2021.12.19.X>.
- Karcher, M.; Voß, C.; Höhne, T. (2020). Wolkige Verheißungen. Die Schul-Cloud als Mittel der Technologisierung von Schule und Lernen. In: *Zeitschrift für Pädagogik 66(3)*, S. 324–340.  
<https://doi.org/10.25656/01:25797>.
- Kämper-van den Boogaart, M.; Reh, S.; Schindler, C.; Scholz, J. (Hrsg.) (2023). *Abitur und Abituraufsätze zwischen 1882 und 1972. Prüfungspraktiken, professionelle Debatten und Aufsatztexte*. Bad Heilbrunn: Klinkhardt. <http://doi.org/10.35468/6052>.
- Klafki, W. (1964). Didaktische Analyse als Kern der Unterrichtsvorbereitung. In: Roth, H. & Blumenthal, A. (Hrsg.): *Grundlegende Aufsätze aus der Zeitschrift Die Deutsche Schule*. Hannover u.a.: Schroedel. S. 5–34.
- KMK (2016). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz*. Berlin: KMK. [https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie\\_2017\\_mit Weiterbildung.pdf](https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit>Weiterbildung.pdf) (Abruf: 20.01.2026).
- KMK (2024). *Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen*. [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2024/2024\\_10\\_10-Handlungsempfehlung-KI.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf) (Abruf: 02.12.2025).

- Kraft, V. (2021). *Erziehung – Beratung – Psychotherapie. Eine Einladung zu Unterscheidungen*. Stuttgart: Kohlhammer.
- Kranzberg, M. (1986): Technology and History: "Kranzberg's laws". In: *Technology and Culture* 27(3), S. 544–560. <https://www.jstor.org/stable/3105385>.
- Kron, F. W.; Jürgens, E.; Standop, J. (Hrsg.) (2013). *Grundwissen Pädagogik*. 8. Aufl. Stuttgart: UTB/Reinhardt.
- Krüger, H.-H.; Rabe-Kleberg, U.; Kramer, R.-T.; Budde, J. (Hrsg.) (2011). *Bildungsungleichheit revisited. Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule*. Wiesbaden: VS. <https://doi.org/10.1007/978-3-531-93403-7>.
- Kvale, S. (1972). *Prüfung und Herrschaft. Hochschulprüfungen zwischen Ritual und Rationalisierung*. Weinheim/Basel: Beltz.
- Khrono (2026a). *Kunstig intelligens. Medisin-studenter reagerer på bruk av KI i sensur*. <https://www.khrono.no/medisinstudenter-reagerer-pa-bruk-av-ki-i-sensur/1035921> (Abruf: 10.02.2026).
- Khrono (2026b). *Storbritannia. Hver femte student innrømmet KI-fusk*. <https://www.khrono.no/hver-femte-student-innrømmet-ki-fusk/992446> (Abruf: 10.02.2026).
- Langela-Bickenbach, A.; Dreier, R.; Wampfler, P.; Albrecht, C (Hrsg.) (2024). *Wege zu einer zeitgemäßen Prüfungskultur. Grundlagen und Praxisbeispiele*. Weinheim/Basel: Beltz.
- Lenzen, D. & Rost, F. (Hrsg.) (2007). *Pädagogische Grundbegriffe. Bd. 2: Jugend—Zeugnis*. 8. Aufl. Reinbek: Rowohlt.
- Loch, W. (1963). Historische Vorbetrachtungen zu einer Pädagogik der Prüfung. In: *Bildung und Erziehung* 16(3/4), S. 180–193. <https://doi.org/10.7788/bue-1963-jg21>.
- Loch, W. (1996). Forschungen zur Anthropologie des Kindes. In: Bartmann, T. & Ulonska, H. (Hrsg.): *Kinder in der Grundschule. Anthropologische Grundlagenforschung*. Bad Heilbrunn: Klinkhardt. S. 147–179.
- Mikhail, T. (2016). *Pädagogisch handeln. Theorie für die Praxis*. Paderborn: Schöningh.
- Montessori, M. (1996). *Die Entdeckung des Kindes*. Hrsg. und eingeleitet v. Paul Oswald und Günter Schulz-Benesch. 12. Aufl. Freiburg i.Br.: Herder.
- Münziger, W./Klafki, W. (Hrsg.) (1995). *Schlüsselprobleme im Unterricht. Thematische Dimensionen einer zukunftsorientierten Allgemeinbildung*. Weinheim: Juventa.
- Muth, J. (1962). *Pädagogischer Takt. Monographie einer aktuellen Form erzieherischen und didaktischen Handelns*. Heidelberg: Quelle & Meyer.

- Nerowski, C. (2017). Leistung als Kriterium von Bildungsgerechtigkeit. In: *Zeitschrift für Erziehungswissenschaft* 21, S. 441–464. <https://doi.org/10.1007/s11618-017-0775-x>.
- Nerowski, C. (2020). Zur Relevanz sprachlicher Präzision für die Rede über Schule, Leistung und Gerechtigkeit. Eine Antwort auf Nils Berkemeyer. In: *Zeitschrift für Erziehungswissenschaft* 23, S. 455–459. <https://doi.org/10.1007/s11618-020-00935-z>.
- Nohl, A.-M. (2025). Bürger erziehen? Erwachsene als Adressaten von Orientierungszumutungen. In: *APuZ* 75(52), S. 4–9.
- ORF [=Österreichischer Rundfunk] (2024). *Gegen KI-Tricks. FH-Bereich ersetzt Bachelorarbeit*. <https://web.archive.org/web/20241004062111/https://science.orf.at/stories/3223859/> (Memento vom 04.10.2024)
- Petzelt, A. (2018). *Grundzüge systematischer Pädagogik*. Neu herausgegeben von T. Mikhail und J. Ruhloff. Freiburg: Lambertus.
- Pötzsch, H. & Buck, M. F. (2023). Digitalisierung, Daten, Eigentum: Bildung im digitalen Kapitalismus. In: Pfeiffer, J.; Güney, S.; Hille, L.; Porak, L.; Theine, H (Hrsg.): *Eigentum, Medien, Öffentlichkeit. Verhandlungen des Netzwerks Kritische Kommunikationswissenschaft*. Frankfurt: Westend. S. 385–406. <https://doi.org/10.53291/OFDA5558>.
- Prahl, H.-W. (1977a). *Prüfungsangst. Symptome, Formen, Ursachen*. München: Nymphenburger.
- Prahl, H.-W. (1977b). Die Gesellschaft und ihre Prüfungen. In: *BWP* 4/77, S. 1–5.
- Prahl, H.-W. (Hrsg.) (1980a). *Prüfungssysteme und Prüfungsreformen an den Hochschulen in der BRD (einschl. West-Berlin). Band I: Thesen – Konzepte – Ergebnisse – Forderungen – Bibliographie*. Hamburg: AHD.
- Prahl, H.-W. (Hrsg.) (1980b). *Prüfungssysteme und Prüfungsreformen an den Hochschulen in der BRD (einschl. West-Berlin). Band II: Gutachten*. Hamburg: AHD.
- Prahl, H.-W. (1983). Prüfungen. In: Huber, L. (Hrsg.): *Enzyklopädie Erziehungswissenschaft, Bd. 10. Ausbildung und Sozialisation in der Hochschule*. Stuttgart: Klett-Cotta. S. 438–450.
- Prange, K. (2012). *Die Zeigestruktur der Erziehung*. 2. Aufl.
- Prange, K./Strobel-Eisele, G. (2014). *Die Formen des pädagogischen Handelns*. 2. Aufl. Stuttgart: Kohlhammer.
- Precht, R. D. (2013). *Anna, die Schule und der liebe Gott. Der Verrat des Bildungssystems an unseren Kindern*. München: Goldmann.
- Radtke, F.-O. (1997). Schulautonomie und Sozialstaat. Wofür ist die Bildungspolitik (noch) verantwortlich? In: *Die Deutsche Schule* 89(3), S. 278–291. <https://doi.org/10.25656/01:31012>.

- Radtke, F.-O. (2003). Die Erziehungswissenschaft der OECD – Aussichten auf die neue Performanz-Kultur. In: *Erziehungswissenschaft* 14(27), S. 109–137.
- Radtke, F.-O. & Weiß, M. (Hrsg.) (2000). *Schulautonomie, Wohlfahrtsstaat und Chancengleichheit. Ein Studienbuch*. Opladen: Leske + Budrich.  
<https://doi.org/10.1007/978-3-322-95122-9>.
- Reble, A. (2004). *Geschichte der Pädagogik*. 21. Aufl. Stuttgart: Klett-Cotta.
- Reddit [u/my-name-is-aki] (2025). *KI macht das Leben als Schüler zur Hölle*.  
<https://redd.it/1kgf3b5>.
- Reichenbach, R. (2003). Pädagogischer Kitsch. In: *Zeitschrift für Pädagogik* 49(6), S. 775–789. <https://doi.org/10.25656/01:3902>.
- Reichenbach, R. (2013). *Für die Schule lernen wir. Plädoyer für eine gewöhnliche Institution*. Seelze: Klett Kallmeyer.
- Reichenbach, R. (2016). Über Neomanie und die posttheoretische Phase in der Erziehungswissenschaft. In: Blömeke, S.; Caruso, M.; Reh, S.; Salaschek, U.; Stiller, J. (Hrsg.): *Traditionen und Zukünfte. Beiträge zum 24. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft*. Opladen: Budrich. S. 17–28.
- Reichenbach, R.; Oser, F. (Hrsg.) (2002). *Die Psychologisierung der Pädagogik. Übel, Notwendigkeit, Fehldiagnose*. Weinheim/München: Juventa.
- Ricken, N. & Reh, S. (2017). Prüfungen – systematische Perspektiven der Geschichte einer pädagogischen Praxis. Einführung in den Themen- teil. In: *Zeitschrift für Pädagogik* 63(3), S. 247–258.  
<https://doi.org/10.25656/01:18538>.
- Rora, C.; Peters, N.-E.; Borchert, J.; Krönig, F. K. (Hrsg.) (2025). *Das Ästhetische im Zeichen künstlicher Verstehenssimulation*. Berlin: Logos.
- Rucker, T. (2024). Allgemeine Didaktik als Theorie eines erziehenden Unterrichts mit Bildungsanspruch. In: Ammann, K.; Emmenegger, F.; Gerstgrasser, R.; Ibrahim O.; Rucker, T.; Welti, G. (Hrsg.): *Bildsamkeit – Komplexität – Werteorientierung. Beiträge zur erziehungswissenschaftlichen Grundlagenforschung*. Bad Heilbrunn: Klinkhardt. S. 222–244.  
<https://doi.org/10.25656/01:30888>.
- Ruhloff, J. (2008). *Prüfungswandel und Prüfungswahn. Qualitätskontroll-Prüfungen, emanzipatorische Prüfungen, tauglichkeitskritische Prüfungen*. Oldenburger Universitätsreden 178. Oldenburg: BIS.  
<https://diglib.bis.uni-oldenburg.de/pub/unire-den/2008/ur178/pdf/ur178.pdf> (Abruf: 02.12.2025).
- Sandfuchs, U.; Melzer, W.; Dühlmeier, B.; Rausch, A. (Hrsg) (2012). *Handbuch Erziehung*. Bad Heilbrunn: Klinkhardt/UTB.
- Schröder, S. (2026). *Prüfung als pädagogisches Problem*. Weinheim: Beltz Juventa.

- Schweiger, G. & Zichy, M. (Hrsg.) (2023). *Zwischenmenschliche Beziehungen im Zeitalter des Digitalen. Ethische und interdisziplinäre Perspektiven*. Heidelberg: Metzler. <https://doi.org/10.1007/978-3-662-66931-0>.
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. In: *Zeitschrift für Pädagogik* 60(6), S. 850–866. <https://doi.org/10.25656/01:14686>.
- Solga, Heike (2025). Meritokratie – die moderne Legitimation ungleicher Bildungschancen. In: Berger, P. & Kahlert, H. (Hrsg.): *Institutionalisierte Ungleichheiten. Wie das Bildungswesen Chancen blockiert*. 3. Aufl. Weinheim: Beltz. S. 19–38. <https://hdl.handle.net/10419/307005>.
- taz [=Die Tageszeitung] (2026). Künstliche Intelligenz an Hochschulen. Universität radikal umdenken. <https://taz.de/Kuenstliche-Intelligenz-an-Hochschulen/!6149714/> (Abruf: 30.01.2026).
- Tenorth, H.-E. (2020). Über die Schwierigkeiten der Pädagogik, über Leistung und Gerechtigkeit im Schulsystem zu reden. Eine Metakritik zu Berkemeyers Nerowski-Kritik. In: *Zeitschrift für Erziehungswissenschaft* 23, S. 439–449. <https://doi.org/10.1007/s11618-020-00933-1>.
- Tenorth, H.-E. (2024). Bildung – Schulzwang – Recht auf Bildung. Historische Traditionen und systematische Optionen. In: *Die Deutsche Schule* 116(1), S. 14–30. <https://doi.org/10.25656/01:29065>.
- Tenorth, H.-E. & Tippelt, R. (Hrsg.) (2007). *Beltz Lexikon Pädagogik*. Weinheim/Basel: Beltz.
- Universität zu Köln (2025). *ChatGPT*. [https://web.archive.org/web/20251205153613/https://verwaltung.uni-koeln.de/stabsstelle02.1/content/faq/data/chatgpt/index\\_ger.html](https://web.archive.org/web/20251205153613/https://verwaltung.uni-koeln.de/stabsstelle02.1/content/faq/data/chatgpt/index_ger.html) (Memento vom 05.12.2025).
- Vulture (2025). *Even God Is Worried About ChatGPT*. <https://www.vulture.com/article/pope-leo-ai-homework.html> (Abruf: 02.12.2025)
- Walgenbach, K. (2023). Digitaler Ableismus im Feld der Bildung. In: *Medien-Pädagogik*, Jahrbuch 20. S. 1–26. <https://doi.org/10.21240/mpaed/jb20/2023.09.01.X>.
- Willatt, C. & Ossa, C. (2026). Effects of Using Generative Artificial Intelligence (GenAI) to Feedback Academic Writing. A Case Study in Chilean Teacher Education. In: Abreu, A.; Vidal de Carvalho, J.; Castanho, R. A. (Hrsg.): *Perspectives and Trends in Education and Technology, Volume 3. Selected Papers from ICITED25*. Cham: Springer Switzerland. S. 471–483. [https://doi.org/10.1007/978-3-032-09074-4\\_47](https://doi.org/10.1007/978-3-032-09074-4_47).
- Wittgenstein, L. (2025). *Philosophische Untersuchungen*. 12. Aufl. Berlin: Suhrkamp.

Zenke, T. (2022). Von der „Entschulung der Gesellschaft“ zur „Entschulung der Schule“: Ein Versuch über das Verschieben „radikaler Grenzpfiler“. In: *Pädagogische Rundschau* 76(1), S. 61–76.  
<https://doi.org/10.3726/PR012022.0005>.