

Künstliche Intelligenz als Herausforderung und Chance für die Weiterentwicklung von Prüfungsformaten

Thomas Hanstein

Inhalt

1	KI-Detektoren: Beispiel und Symptom des (hoch-)schulischen Kontrollparadigmas	2
2	Kompetenzermöglichung statt Wissensvermittlung	3
2.1	Handlungsorientierung als berufspädagogische Grundhaltung	4
2.2	Taxonomische Ordnung als pädagogische Musteränderung	5
2.3	Eigenverantwortliches Lernen als naturgemäßes Lernen von Menschen	7
3	Das SAMR-Modell als Spiegelung von KI-Einsatz	9
4	Fazit & Ausblick: Prüfungen auf dem Prüfstand	10
	Literatur	13

Zusammenfassung

Jede:r gute Handwerker:in wird nicht an seinem Werkzeug gemessen, sondern am Ergebnis seiner Arbeit, dem Produkt, und auch der Identifikation mit diesem. Der Rückschluss auf die verwendeten Werkzeuge erlaubt noch keine Aussage über die Güte und Funktionalität des Ergebnisses. Vielmehr gehört es zur handwerklichen Kompetenz, die richtigen Werkzeuge für den jeweiligen Einsatz zu wählen. Bereits daran wird deutlich, dass die Kenntnis und Einschätzung der Tools zur Expertise selbstverständlich dazugehören. Beim Thema KI indes herrscht schulisch und hochschulisch einige Aufregung, da man den Einsatz von KI bei schulischen und wissenschaftlichen Leistungen nicht vollumfänglich überprüfen könne. Der Beitrag plädiert dafür, sich vom Kontrollparadigma nach und nach zu lösen und sich vielmehr auf das zu konzentrieren, was schon immer guten Unterricht und gute Lehre ausmacht: ein angemessener Transfer auf die jeweilige Handlungssituation. Was dabei überprüfbar ist, ist nicht nur die fachliche Angemessenheit und die situative Originalität eines Lösungsansatzes, sondern – bei angemessenen Prüfungsformaten – auch, inwieweit der Prüfling sein Thema durchdrungen hat. Die

T. Hanstein (✉)

DIPLOMA Hochschule & Institut für Lehrerbildung und Berufsbildungsforschung,
Friedrichshafen, Deutschland

E-Mail: thomas.hanstein@diploma.de

Analogie zum handwerklichen Schaffensprozess zeigt, dass dieser immer auch mit einem hohen Maß an fachlicher und persönlicher Identifikation verbunden ist. Der Einsatz von Werkzeugen versteht sich von selbst – alles andere wäre wohl ein kulturgeschichtlicher Rückschritt –, bestimmt allerdings nicht den Fokus. Anstatt daher Energie in die Überprüfbarkeit des Einsatzes von KI zu legen, plädiert der vorliegende Beitrag für Prüfungsformate, die nicht darauf abzielen, Wissensbestände zu reproduzieren (oder zu halluzinieren), sondern stattdessen an einer Identifikation im Schaffens- wie im Lernprozess ansetzen. Denn diese didaktische Perspektive führt wie von selbst dazu, Prozesswissen zu reorganisieren, zu transferieren und damit eigenständig und personalkompetent zu reflektieren.

Schlüsselwörter

Künstliche Intelligenz · Prüfungen · Klausuren · Prüfungsgespräch · Fragetechnik · Trialog · Transfer

1 KI-Detektoren: Beispiel und Symptom des (hoch-)schulischen Kontrollparadigmas

Einleitend soll ein reales Beispiel aus der beruflichen Erfahrungswelt des Autors in die Phänomenologie des Themas einführen: Im Kontext von Schul- und Unterrichtsbesuchen blieb mir ein besonders eindrückliches Beispiel von Kontrollparadigma haften: Ein Schulleiter hatte seinen Lehrkräften unnötige Mehrarbeit gemacht, indem sämtliche Texte von Schüler:innen zunächst durch KI-Detektoren gejagt werden mussten, bevor sie durch die Brille der menschlichen Lehrkraft gelesen werden durften. Im Gespräch zitierte der Kollege den Technologiekritiker John Culkin: „Erst gestalten wir unsere Werkzeuge, dann gestalten sie uns.“ Ich verstand diese Sorge. Sie ist nicht neu, sondern eine Grundangst, seitdem Menschen Werkzeuge kreieren und von ihren unbändigen Energien und ungeahnten Möglichkeiten – zugleich – fasziniert wie erschrocken sind. Das Paradoxon in diesem Beispiel war für mich ebenso erschreckend wie offenkundig: KI-Technologie gegen KI-Tools einsetzen zu wollen, ohne technisch prüfen zu können, mit welcher Zuverlässigkeit diese KI-Detektoren arbeiten; auch, welche sich eher anbieten und welche nicht. Das (sich selbst suggerierte) Gefühl, wieder Kontrolle bekommen zu können, war leitend, die Angst zu reduzieren. Doch änderte dieser Versuch der Wiedergewinnung von Kontrolle nichts am eigentlichen Phänomen. Der Schulleiter hatte lediglich die KI-Player auf dem Spielfeld der Macht und Ohnmacht verschoben. Ich nutzte im Gespräch einen pädagogischen Kniff, um seine Wahrnehmung zu verschieben. Und brachte den ausgebildeten Fremdsprachenlehrer dazu, über seine schulischen Erfahrungen mit Englisch- und Französischprüfungen zu sprechen. In beiden Fächern war ich mittelmäßig begabt, allerdings hatte ich schon als Schüler ein Gespür für Syntax und Semantik, für sprachliche Logik. Es war mir immer möglich, durch das Erkennen entsprechender Schlüsselwörter in Texten passende Antworten zu konstruieren – ohne weder den Sinnzusammenhang noch die Frage richtig verstanden zu haben. Für

dieses Phänomen hatte der Anglist auch umgehend eine Bezeichnung parat: Simulationskompetenz. Und stimmte mir zu, dass es eine übliche Strategie sprachlich Begabter sei. Als solchen verstand ich mich jedoch nicht, was ich ihm damit verdeutlichte, dass diese Taktik in Englisch zwar gut, in Französisch noch teilweise, in Latein jedoch nicht mehr funktioniert habe. Mit diesem sprachlichen Vergleich waren wir an einem entscheidenden Punkt angekommen: Das Latein besitzt eine komplexere Grammatik, die sich auch in komplizierteren Satzkonstruktionen (und ohne zeichenhafte Unterstützung wie das Altgriechisch, das mir persönlich deshalb als Schüler viel eingängiger war) wiederfindet. Hier benötigt man, neben einer Intuition für wahrscheinlich passende Zusammenhänge ein ausgeprägtes Grundverständnis für Grammatik sowie alle möglichen grammatikalischen Besonderheiten. Mit purem Sprachgefühl und Wahrscheinlichkeit ist kein Latinum zu machen, eher ist man mit seinem Latein früh am Ende. Das Beispiel zeigt: Der Versuch, Verstehen zu simulieren und dazu nach einer Wahrscheinlichkeit passende Antworten zu parieren, ist nichts neues. Er kommt mit dem KI-Einsatz nur auf ein neues Level. Und er fordert von Lehrenden zunächst eine Reflexion über ihre Vorstellung von Lehren und Lernen – und ihr Rollenverständnis – ein.

2 Kompetenzermöglichung statt Wissensvermittlung

In Krisen- und Transformationszeiten gibt es bekanntlich auch die Tendenz, auf Altes, Gewohntes zurückzufallen. Das war in und nach der COVID-19-Pandemie erkennbar, als eine starke Fixierung auf das Verteilen von Arbeitsblättern und auf die Kontrolle der Anwesenheit erfolgt ist. Und das ist auch aktuell wieder zu beobachten: Regelmäßig sind Posts von Bildungsexpert:innen im Umlauf, in denen davon ausgegangen wird, dass man dem KI-Einsatz nur mit viel Wissen beikommen könne. Das Argument ist nachvollziehbar: Um wahr von falsch zu differenzieren, brauche es Wissen. Das ist grundsätzlich nicht von der Hand zu weisen, vor allem aber braucht es – neben Grundwissen über das entsprechende Feld – Kompetenzen. In den Transformationsprozessen, die bildungstheoretisch und -praktisch durch KI ausgelöst werden, muss daher zuallererst an altbewährte Prinzipien der Berufspädagogik erinnert werden: an das Prinzip der Handlungsorientierung, das Lernverständnis nach Taxonomiestufen und an spiralförmige Curricula.

Die *Kernthese* dieses Beitrags lautet: In der *Gestaltung des „Outcomes“* eines Bildungsprozesses wird das bis dahin angewandte *Grundverständnis von Lehren und Lernen* offenbar. Der KI-Einsatz kann als Hemmschuh für eigenständige Prüfungsergebnisse gedeutet werden, aber auch als Impetus für die Weiterentwicklung von Prüfungsformaten. Dabei muss das Rad nicht neu erfunden (oder von KI halluziniert) werden, sondern die Anpassung darf sich auf der Basis vorliegender Konzepte, Modelle und Praktiken vollziehen. Ob und inwiefern Bildungseinrichtungen und Lehrende dafür offen sind, zeigt sich nicht nur an der Anpassung von Lern-, sondern letztlich in der Weiterentwicklung von Prüfungsformaten. Der „Modus probandi“ wird damit zum Indikator des pädagogischen Umgangs mit KI.

2.1 Handlungsorientierung als berufspädagogische Grundhaltung

Das pädagogische Prinzip der Handlungsorientierung, welches von einer Erweiterung vorliegender Kompetenzen durch (die Bewährung von) Handlungssituationen ausgeht, muss als ein epochaler Paradigmenwechsel in der Berufspädagogik der letzten Jahrzehnte verstanden werden. Mit diesem wurde das Verständnis von Lehren und Lernen quasi auf den Kopf gestellt. Franz Weinert definierte Kompetenzen als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen [die willentliche Steuerung von Handlungen und Handlungsabsichten] und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2001, S. 27–28).

Für die Konzeption des handlungsorientierten Unterrichts und handlungsorientierter Lehre ist eine komplexe, problemindizierende berufliche Lernsituation – herkömmlich Aufgabenstellung genannt – leitend. Diese steht im Zentrum und am Anfang des Unterrichts. Der Aufbau der Unterrichtsstunde und die Lerninhalte richten sich an dieser Lernsituation aus. Damit ist *handlungsorientierte Lehre immer kompetenzorientiert*: Kompetenzen ergeben sich somit aus dem Wissen, dem Können und dem Handeln. Sie bezeichnen Fähigkeiten und Fertigkeiten, um berufliche Herausforderungen nicht nur fachlich angemessen zu lösen. Kompetenz bedeutet insofern die Fähigkeit, handelnd mit Wissen umgehen zu können – und zwar situativ und ohne, genau diese Herausforderung schon einmal so erlebt (und gelöst) zu haben. Kompetenzen können nur im Handeln (prozesshaft lernend) erworben werden. Kompetenzen zeigen sich im konkreten handelnden Tun. Damit ist *kompetenzorientierte Lehre im Grundsatz induktiv*: Die Praxis steht am Anfang, nicht (noch so relevante) Theorien und Modelle. Darauf baut sich in einer ganzheitlichen – nicht nur kognitiven – Aktivierung die Ableitung der Herausforderungen aus der Problemsituation auf: Was ist nötig, um die vorliegende Situation in einer hohen Qualität lösen zu können? Welche Theorien können dafür herangezogen werden? Gibt es handlungsleitende Modelle, die zur vorliegenden Anwendung überprüft werden können? Damit ist *handlungsorientierte Lehre immer experimentell und forschend*: Faktenwissen allein reicht nicht aus, es muss in einem ausgewogenen Verhältnis mit Begründungswissen, Verfahrenswissen und mit Einsatzwissen angereichert werden. Damit ist *handlungsorientierte Lehre immer prozesshaft*. Hier zeigt sich letztlich, ob der Paradigmenwechsel vollzogen ist: Lehrende und Lernende denken nicht (alleinig) vom Ergebnis her, sondern dieses steht immer unter einer gewissen Vorläufigkeit. Diese Charakteristika haben Auswirkungen auf die Leistungsbewertung: Diese kann nicht mehr vom Inhalt erfolgen, sondern muss den Grad der Anwendung als Richtschnur nehmen. Feedbackschleifen im Prozess gehören zum handlungsorientierten Unterricht selbstverständlich dazu. Sie sind ein äußerst wichtiger Teil der Rückkopplung und damit erst *ermöglichter Lehr-Lern-Resonanz*.

Die *These* dieses Rekurses lautet: Handlungs- und Kompetenzorientierung darf nicht vor der Phase der Prüfungsvorbereitung enden, sondern muss auch die *Grundlage für Prüfungsformate* bilden und diese als vollständige Handlungen prägen.

2.2 Taxonomische Ordnung als pädagogische Musteränderung

Die Logik der Taxonomieeinstufung unterstützt den Paradigmenwechsel: vom Wissen zum Können, von Inhalten zu Kompetenzen, von Theorien zur Praxis. Bereits vor mehreren Jahrzehnten, 1965, wurde mit „Taxonomy of Educational Objectives“ (Bloom et al., 1965) eine erste operationalisierte Systematik zur Analyse von Lernzieldimensionen vorgelegt. Ergänzt wurden die kognitiven Lernzieltaxonomien durch die psychomotorische Taxonomie nach Harrow (1972). Bis heute ist der Ansatz der Taxonomie vielfältig fortgeschrieben worden – aber noch nicht konsequent bis hinein in Prüfungsformate.

Die Taxonomie der Lernziele intendiert Werte des Schwierigkeitsgrades zu berücksichtigen. Insofern hat die Taxonomielogik keinen auswählenden, sondern einen analytischen Charakter. Dazu werden Lernziele hinsichtlich der Reihenfolge von Anforderungen kategorisiert. Dies geschieht vor dem Hintergrund, dass Lernziele – der Handlungsorientierung folgend – nicht einfach nur theoretisch definiert werden können, sondern die Kompetenzen der Lernenden berücksichtigen sollen. Dazu werden drei Dimensionen analysiert, die sich in den im Folgenden dargestellten Wissensarten in der entsprechenden Ausprägung wiederfinden: Der *Grad der Komplexität* (kognitive Dimension), der *Grad der Verinnerlichung* (Verhaltensdimension) sowie der *Grad der psychomotorischen Koordinierbarkeit* (psychomotorische Dimension). Durch dieses Vorgehen können Lernziele operationalisiert und attribuiert werden. Die Zuschreibung erfolgt durch die Nutzung von Operatoren, die das jeweilige Lernziel beschreib- und damit bewertbar machen. Anderson und Krathwohl (2001) entwickelten das hierarchische Ausgangsmodell weiter. Während sich die Taxonomie von Bloom et al. auf Attribute stützt, kann mit diesem Taxonomiekonzept ein direkter *Bezug zur Wissensdimension* hergeleitet werden. Die Wissensdimension stellt dabei den Grad der kognitiven Anforderung an Lernende dar (Tab. 1):

Passend zu den Wissensarten entwickelten Anderson und Krathwohl Tätigkeitswörter, die in der *Schulpädagogik Operatoren* genannt werden oder sich davon ableiten lassen: Das *Erinnern* ruft Faktenwissen ab, listet dieses auf, und verweist auf eine Reproduktionsleistung. Im *Verstehen* wird die Bedeutung erfasst, nur so kann sie erklärt, dargestellt und zusammengefasst werden. Das *Anwenden* setzt auf den Transfer des Gelernten auf eine konkrete Situation. Unter *Analysieren* wird das Differenzieren der Informationen sowie das Erkennen der zugrunde liegenden Zusammenhänge verstanden. Eine ebenso fundierte Kompetenz setzt *Beurteilen* voraus, denn hierzu müssen Kriterien angewandt werden, um zu begründeten Entscheidungen zu gelangen. Und im *Erschaffen* schließlich werden Elemente neuartig so strukturiert, dass etwas Neues entsteht.

Tab. 1 Wissensarten nach dem Taxonomieansatz. (Eigene Darstellung, nach Anderson und Krathwohl 2001)

Wissensart	Beschreibung
<i>Faktenwissen</i>	Lernende verfügen über Faktenwissen und können dieses je nach kognitivem Vermögen entweder auswendig wiedergeben/reproduzieren (erinnern). Eine Anwendung von Wissen und das Generieren neuer Fakten stehen am Anfang.
<i>Konzeptionelles Wissen</i>	Hier ist bereits eine kumulative Vernetzung erkennbar. Lernende verfügen über die Fähigkeit, Fakten in einen Sinnzusammenhang zu stellen, Kategorien zu bilden, Ordnungsprinzipien abzuleiten. Auf der Ebene der kognitiven Prozesse kann zwischen der Wiedergabe (Erinnern/Reproduzieren) und der Erzeugung eigener Schemata (Reorganisieren) differenziert werden.
<i>Prozedurales Wissen</i>	Hier ist implizites, verinnerlichtes Wissen erkennbar, welches nicht zwingend verbalisiert werden muss. Es geht um die handlungsorientierte Beherrschung von Fertigkeiten, Handlungsweisen sowie um das Prozessverständnis. Auf der Ebene der kognitiven Prozesse können demnach einfache Fertigkeiten anhand einer Lernsituation umgesetzt werden.
<i>Metakognitives Wissen</i>	Hier zeigt sich die Ebene der Reflexion dessen, was gewusst und beherrscht wird. Eigene Lernstrategien sind entwickelt und Problemlösungsprozesse möglich. Auf der Ebene der kognitiven Prozesse kann dies bedeuten, dass die Lernstrategie soweit etabliert ist, dass Prüfungswissen kurzfristig auswendig gelernt, aber auf die vorliegende Situation mit einer hohen Präzision angewandt wird. Auf der Ebene der Erzeugung können adäquate Handlungsschritte geplant, umgesetzt und abschließend evaluiert werden.

Konsequenzen der Taxonomielogik für KI-resiliente Prüfungsformate

Aus diesen Verben lassen sich Operatoren ableiten, die im Erwartungshorizont mit einem passenden Punktwert verbunden werden. Damit ergibt sich die erste Verschiebung durch den Einsatz von KI: Was Rechensysteme und Parameter – insofern korrekt – wiedergeben können, sollte nicht mehr den bisherigen Stellenwert haben. KI-Systeme können in einer schnelleren Zeit Inhalte abrufen. Um die Qualität dieser aber bewerten zu können, bedarf es Kompetenzen in allen weiteren Wissens Ebenen. Im Überprüfen, Bewerten, Analysieren, Verknüpfen zeigen sich Leistungen, die zwar mehr und mehr auch von KI abgenommen werden, die aber eines voraussetzen: schlaue wie zugleich offene Fragen. Der Anteil der Reorganisations- und Transferfragen in Prüfungskolloquien sollte dabei prozentual wesentlich höher bepunktet werden, als dies bisher der Fall war.

Diese Systematik lässt sich an Unterrichtsbeispielen verschiedener Fachschaften veranschaulichen (Hanstein, 2024, S. 30–34). Dabei werden die handlungsorientierten Lernsituationen als berufliche Handlungssituationen gefasst, die Lernende aktivieren, alle Taxonomiestufen bzw. Wissensarten zu aktivieren. Dieser Zugang stellt zugleich eine situationsorientierte Elementarisierung dar (Hanstein, 2025, S. 88–90). Wo eine derart handlungsorientiert-elementarisierte Vorbereitung im Unterricht und in Seminaren stattfindet, ist der Schritt zur Prüfung eine leistbare Hürde.

Die *These* lautet hier: Prüfungsformate sollten sich konsequent an der Taxonomielogik orientieren und ihren Schwerpunkt von der Reproduktion zum handlungsorientierten Transfer verschieben. Die Bedeutung der Fragetechnik bekommt dabei eine neuartige Bedeutung.

2.3 Eigenverantwortliches Lernen als naturgemäßes Lernen von Menschen

Die Vorstellung, Wissen vermitteln zu können, sitzt tief in der Bildungstradition des Abendlandes. Die Erinnerung daran, wie wir als Heranwachsende gelernt haben, macht Lernen als selbstgesteuerten Aneignungsprozess erkennbar. In den Generationen bis in die 60er-Jahre des letzten Jahrhunderts hinein wurde Heranwachsenden relativ genau vorgegeben, was sie zu tun hatten. Sie wurden in der Regel auch nicht dazu gefragt, was sie davon hielten. Das pädagogische Muster hieß: Erwachsene haben die größere Erfahrung, daher wissen sie es besser. Orientierst du dich daran, wirst du es gut machen. Die Crux besteht in der Dynamik der Veränderungen der letzten zwei Generationen. Sodass das, was gestern erfolgsversprechend war, keine Blaupause für morgen sein kann – oder deutlicher: eine für den wahrscheinlichen Misserfolg. Weil die Herausforderungen von morgen nicht mit den Rezepten von gestern zu bewältigen sind. Das bedeutet pädagogisch: Durch das Nachahmen wird man nur insoweit kompetent, als dass dieselbe (oder zumindest eine doch sehr affine) Situation nochmal auf uns einwirkt – und genau so hat das mechanistische Industriezeitalter funktioniert. Man muss sogar so weit gehen, zu sagen: Reproduzieren schafft als kopierte Handlung Kompetenzen für genau wiederkehrende Situationen. In aller Regel aber nicht mehr. Heranwachsende wachsen aber insbesondere an neuartigen Herausforderungen, am Finden eigener Lösungsansätze für Probleme, auf die es noch gar keine vorgegebene Antwort gibt. Kinder freuen sich natürlich, wenn sie etwas genauso gut können wie ihre Eltern oder Geschwister. Mehr noch freuen sie sich aber, wenn sie etwas Kniffliges selbstständig gelöst haben. Die dabei entstehenden Gefühle und Botenstoffe sind lernförderliche Wundermittel.

Heinz Klippert hat dafür in der Berufspädagogik den Ansatz des „*Eigenverantwortlichen Arbeitens und Lernens*“ – kurz: EVA – entwickelt (Klippert, 2001). Die Prinzipien sind einfach erklärt, doch in der Umsetzung ein (innerer) Prozess bei Lehrenden: Die Lernenden sind aktiv Handelnde, die Lehrenden rücken aus dem Zentrum, nehmen dazu ihre Dominanz zurück, und verändern so im Prozess ihre Rolle – von der vermittelnden Fachperson zum: zur Lernbegleiter:in. Für ganzheitliche Lernprozesse werden alle Wahrnehmungskanäle genutzt, Kommunikation und Kooperation zwischen den Lernenden werden durch ein angemessenes Methodenrepertoire gefördert. Offene Aufgabenstellungen fördern die Selbstständigkeit, Kreativität und Problemlösekompetenz der Lernenden. Eine so erfahrene Eigenverantwortlichkeit wirkt sich als Bestärkung und Attraktor aus.

Die auf diesem Klassiker beruhende *These* lautet: Das EVA-Prinzip lässt sich als Ganzes auf Prüfungen übertragen. Etwaige Unsicherheit durch (zu) offene Fragestel-

lungen kann durch ein angemessenes Scaffolding – Hilfe durch thematische Leitplanken, wenn der:die Lehrende seinen bzw. ihren Weg sucht und situativ Unterstützung braucht – verhindert werden. Empfohlen wird eine interaktionale Form des Scaffolding, je nach der bereits mitgebrachten Erfahrung an Eigenverantwortlichkeit der Lernenden bzw. zu Prüfenden. Das setzt bei Lehrenden und Prüfenden eine hohe Kompetenz und eine große Sicherheit in der eigenen Fachlichkeit voraus. Denn die Offenheit im Prozess spiegelt sich in einer ebenso notwendigen Offenheit im Scaffolding: Der:die Prüfende kommt in die Rolle des Coachs, indem er oder sie nicht anbietet, was auf seinem bzw. ihrem vorbereiteten Fragezettel steht, sondern indem er oder sie *passgenaue Leitfragen* bietet, die den Prüfling in seinem kreativen Prozess weiterbringen.

Fallbeispiel: Kompetenzfeststellungsprozess im Masterprogramm

Das Mastereinstiegsprogramm der DIPLOMA Hochschule sieht unter bestimmten Voraussetzungen einen Kompetenzfeststellungsprozess im Fachbereich Berufspädagogik vor. In diesem reflektieren die Studierenden zum einen ihre Rollen, die sie während der schulpraktischen Phase ihres Studiums als Lehrende einnehmen. Zum anderen werden sie durch konkrete schulische Herausforderungen in die Situation gebracht, ohne viel Vorwissen Lösungsvorschläge zu erarbeiten. Zunächst sollen die Studierenden assoziativ vorgehen, danach tauschen sie ihre Assoziationen aus. Das Brainstorming hat das Ziel, die wesentlichen Herausforderungen, welche die Problemsituation mit sich bringen, zu erkennen. In einem weiteren Schritt werden Ziele und Unterziele definiert. Erst jetzt macht es Sinn, nach geeigneten Modellen zu recherchieren, um praxisorientiert und theoriegeleitet die Aufgabe zu lösen. Für diese Erarbeitung bekommen die Studierenden Zeit zur eigenständigen Verfügung. Ob sie in dieser Zeit künstliche Intelligenz nutzen oder nicht, stellt sich als Frage nicht, wenn verstanden worden ist, dass die Prüfungsleistung nicht im Googeln und Präsentieren fertiger Lösungen besteht. Stattdessen läuft das Kolloquium auf ein PrüfungsGespäch hinaus, das nicht im inhaltlichen Vorstellen, sondern in einem wirklichen Gespräch zwischen Prüfenden, je einem Prüflingstandem und künstlicher Intelligenz darstellt – in einem *resonanten Trialog, der spiralförmig angelegt* ist. Der Prozess ist im Vorfeld derart transparent, dass die zu Prüfenden sich als Team vorher kennen und in die Prüfung auch einstimmen und diese trainieren können. Ebenso bekannt sind die inhaltlichen Themenbereiche. Einzig die je konkrete berufliche Handlungssituation wird mit Beginn der Prüfung erst bekanntgegeben.

Diese Prüfungsform findet online und sogar von daheim aus statt. Durch eine doppelte Kamera und eine entsprechende Technik ist sichergestellt, dass während des online stattfindenden Gesprächs kein KI-Prompt angelegt werden kann. Da je zwei Prüflinge wechselseitig und inhaltlich aufeinander aufbauend, aber immer mit dem Prüfer im Gespräch sind, hätte dazu auch keiner von beiden hinreichend Zeit. Die Dokumentation der Prüfungstrialoge indizieren eine hohe Prozesshaftigkeit und Dramaturgie der Kolloquien. Jedes Prüfungsgespräch gibt sich dabei als einmalig zu erkennen, ebenso die Fragen der Prüfenden. Diese kommen dabei in eine neuartig aktive Rolle, da sie *situativ* Fragen und Impulse einstreuen – und zwar sowohl offen in die Runde, als auch (über die direkte Chat-Funktion) direkt an die Prüflinge. Da sie

sich in ihren Nachfragen an der von den Prüflingen dargebotenen Durchdringungstiefe und deren Transfer auf die Praxis orientieren, sind sowohl die prüfungsrechtliche Vergleichbarkeit – und zwar als je einzelne Bewertungen – wie gleichzeitig die adäquate Passung des Anforderungsniveaus aller Tandems gewährleistet. Die größte Herausforderung besteht insofern nicht aufseiten der Prüflinge, sondern in der didaktischen Vorbereitung, Schulung und Weiterbildung der Prüfenden. Denn der Paradigmenwechsel liegt in der prozessoffenen, in die Reproduktion, den Transfer und die Reflexion zielende Fragestellung. Das ist nicht weniger als eine *Habituserweiterung bei Dozierenden*, die sich genauso, dem Wort gemäß verstehen und diese Haltung auf Prüfungen übertragen. Ihre menschliche Kompetenz als Prüfende zeigt sich daran, dass sie ihre fachliche Befähigung und ihre pädagogische Erfahrung dazu nutzen, „Antworten“ von Systemen künstlicher Intelligenz durch die Prüflinge kritisch hinterfragen zu lassen. Mit einem derart kompetenten Transfer bleibt das Gegenüber in jedem Fall der zu prüfende *Mensch*. Neben der pädagogischen Weiterentwicklung der Prüfenden besteht die Hauptherausforderung dieser Transformationsaufgabe im Faktor der in einem Prüfungsgespräch zur Verfügung stehenden *Zeit*. Dieser Aspekt ist bei der *Planung trialogischer Prüfungsformate* ernst zu nehmen, wobei die meiste Zeit – konsequent zur oben dargestellten taxonomischen Ordnung – für Gespräche auf Reproduktions- und TransfERNiveau eingeplant werden sollte.

3 Das SAMR-Modell als Spiegelung von KI-Einsatz

Die Digitalisierungsthese von Jürgen Handke (Handke, 2020) lassen sich leicht auf das Thema KI-Nutzung übertragen: Digitalisierung und Digitalität beginnen nicht mit Hardware und Programmen, sondern mit einer *Veränderung im Mindset*. Den Grad der digitalen Durchdringung veranschaulicht das SAMR-Modell (Substitution/Augmentation/Modifikation/Redefinition) nach Ruben Puentedura. Der Einsatz von Lerntechnologie wird hier im Bild von Strand und Meer visualisiert: Am Strand steht – in die Ferne blickend – der, der sich die Füße nicht nass machen will. Er bleibt fragend zurück. Das beintiefte Wasser versinnbildlicht in dieser Metapher die Ersetzung (Substitution). Hier wird Technik als direkter Ersatz bisheriger Lernmittel gesehen – eine funktionale Änderung findet nicht statt. Der Klassiker ist der Austausch eines Arbeitsblattes aus Papier durch ein PDF-Dokument, üblicherweise auf der Substitutionsebene (ohne Möglichkeit, diese zu beschreiben). Sobald Lernmittel derart ersetzt werden, dass eine funktionale Verbesserung beobachtet werden kann, spricht Puentedura von der Erweiterung (Augmentation). Was einfach klingt, kennen wir in seiner Dysfunktionalität zu hauf: Papiere, die digital übermittelt wurden, dann aber zur Unterschrift wieder als Ausdruck auf dem Postweg zurückkommen. Derart analog-digitale Doppelstrukturen erschweren Prozesse und können der Motivation für digitalisierte Weiterentwicklungen nicht sonderlich dienlich sein. Die Phasen der Substitution und Augmentation bezeichnet Puentedura als Ebene des *Enhancement*, eine Umgestaltung von Lerntechnologie hat diese Ebene (noch) nicht zum Ziel und nicht in der Wirkung. Diese Ebene der *Transformation* setzt allerdings erst mit den Phasen Änderung (Modification) und Neubelebung (Redefinition) ein. In der dritten Phase bewirken

Tab. 2 SAMR-Modell von Ruben Puentedura. (Eigene Darstellung, nach Schulentwicklung, NRW, 2021)

Keine Technik	Ersetzung	Erweiterung	Änderung	Neubelebung
Strand	seichtes Wasser	Schwimmbereich	tieferes Wasser	Hochsee
	Substitution	Augmentation	Modification	Redefinition
	Enhancement		Transformation	

SAMR Substitution/Augmentation/Modification/Redefinition

die durch den Technikeinsatz erfolgten Umgestaltungen wichtige Veränderungen von Abläufen. Die vierte Phase ist vom Erzeugen neuartiger Zusammenhänge geprägt. Dadurch werden Inhalte und Lehr- und Lernprozesse weitestgehend bis vollständig neugestaltet (Tab. 2).

Der Unterschied liegt auf der Hand: Künstliche Intelligenz – ohne Einführung und Vorbereitung – konfrontiert ganz unmittelbar auf der vierten Stufe. Transformation und Redefinition ereignen sich ungeplant, ungewollt und permanent in beschleunigten Rhythmen und verdichteten Strukturen. Künstliche Intelligenz verlangt, anders als Digitalisierung ohne Vorbereitung wie im Kontext der COVID-19-Pandemie, keinen „Sprung ins kalte Wasser“, sondern verlangt uns das Eintauchen und Schwimmenkönnen in tiefen, kalten Wassern ab – und zwar ohne thematisches Einschwimmen. Insofern braucht es eine Kompetenz des Freischwimmens, die sich ebenfalls erst noch redefinieren muss. Und zwar bei Lehrenden/Prüfenden wie Lernenden. Warm wird einem beim Schwimmen von selbst. Was zur Umkehrung klassischer Lehr- und Lernarrangements beobachtet und analysiert worden ist (Hanstein und Naun, 2024), lässt sich auf Prüfungssituationen übertragen: dass die Expert:innen des Lernens vor uns sitzen und wir Digital Natives nichts von Medendidaktik zu erzählen brauchen. Gerade, weil sie einen technischen Vorsprung haben, sollte dieser auch für Lehr- und Lernprozesse genutzt werden – alles andere wäre lernpsychologisch gesehen eine Bremse. Die Frage ist insofern: Was ist dienlich zur Erledigung einer (Prüfungs)Aufgabe? Was kann KI – und was KI besser kann als Menschen, sollte sie abnehmen dürfen. Und umgekehrt: *Worin liegt die eigentliche menschliche Kompetenz* – und an genau diesen (fachspezifischen) Antworten kann – und muss künftig – die Konzeption von Prüfungen ansetzen. Und was zu mündlichen Prüfungsformaten ausgeführt worden ist, lässt sich analog auf schriftliche Klausuren übertragen.

4 **Fazit & Ausblick: Prüfungen auf dem Prüfstand**

Bereits mit dem bis heute selbstverständlich verwendeten, allerdings jede Pädagogik engführenden Begriff der „Lernstandskontrollen“ ist alles gesagt: Die Einschränkung des „Outputs“ auf den Inhalt, die Messbarkeit von Wissen und das vermutlich tiefsitzendste Muster in der pädagogischen Zunft, das Kontrollparadigma. Im aktuellen und fortwährenden Transformationsprozess, den künstliche Intelligenz und Digitalisierung ausgelöst oder zumindest beschleunigt und qualitativ verdichtet

Tab. 3 Prüfungen auf dem Prüfstand. (Eigene Darstellung)

Kriterien KI-robuster Prüfungsformate nach Hanstein	• sozialenergetisch
	• handlungsorientiert
	• transferschaffend
	• resonanzfördernd
	• beziehungsdidaktisch
	• kompetenzerweiternd

haben, verliert damit der sicherste Garant der schulischen Sozialisation seine Bedeutung. Die Erkenntnis, vom Prinzip der Kontrolle ablassen zu müssen, weil schlichtweg nicht zu kontrollieren ist, was nicht kontrollierbar ist, verursacht Unbehagen, Unsicherheit und Angst. Da ist der „KI-Teufel“ schnell an die Wand gemalt. Wie wäre es mit der Vision, dass *prägende, innovative, pädagogische Ansätze* der letzten Jahre und Jahrzehnte – wie beispielhaft: Bildung als Beziehung, Lehr- und Lernprozesse als Resonanzgeschehen sowie als Potenzialentfaltung – *sich nicht auch, sondern gerade beim (immer vorläufigen) Abschluss eines Lernprozesses – einer Prüfungsleistung – zeigen* müssen? Denn ein noch so handlungsorientiertes Projekt konterkariert sich selbst, wenn es auf eine Klausur oder ein Kolloquium vorbereiten soll, in dem Wissen abgefragt wird.

Die abschließende These lautet daher:
Prüfungsformate, die sich an dieser Vision orientieren, erfüllen die *sechs Kriterien*, die in Tab. 3 dargestellt sind.

Am oben beschriebenen Beispiel des Kompetenzfeststellungsprozesses kann dies verdeutlicht werden:

Die Prüflinge werden in Zweier- oder Dreierteams (sozialenergetisch) mit einer praktischen Lern- und Handlungssituation konfrontiert (handlungsorientiert). Diese verkörpert eine Herausforderung, die nicht allein durch fachliches Wissen, sondern nur durch gelungene Verbindung fachlicher und überfachlicher Kompetenzen (transferschaffend) sowie durch den intensiven Austausch miteinander (resonanzfördernd) gelöst werden kann. Durch das anschließende Kolloquium führt der:die Prüfende mit offenen Transferfragen (beziehungsdidaktisch). Die größte Herausforderung für traditionelle Prüfende besteht darin, spiralförmig und in iterativen Schleifen von der Oberfläche in den Grund vorzudringen (kompetenzerweiternd). Damit ist auch die *Hauptherausforderung der kommenden Jahre* benannt: Sie besteht nicht – als ersten Schritt – im Schaffen neuer Prüfungsformate, sondern – zunächst – in der Musteränderung bei Lehrenden, in ihrer Rolle als Prüfende. Hierzu braucht es Zeit und Fehlerfreundlichkeit, Reflexion und Feedback sowie adäquate Fortbildungsangebote.

Die Frage nach passenden Prüfungsformaten sollte erst der zweite Schritt sein. Eine große Zahl an Methoden, die für den Unterricht und die Lehre eingesetzt werden können (Hanstein und Lanig, 2020) und die genannten Kriterien erfüllen, sind auch als Prüfungsformate – in Präsenz, online oder hybrid – geeignet. Bewährt haben sich beispielhaft *die Pro-und-Kontra-Debatte* und das *Expert:innengespräch*. Die Feedback- und Bewertungskriterien lassen sich – im besten Fall im Vorfeld der Prüfungsphase, zusammen mit den Lernenden – von den genannten Kriterien ab-

leiten und orientieren sich an einer gewünschten taxonomischen Durchdringungstiefe. Je mehr diese Form zum selbstverständlichen Prozess mit Lernenden wird, je angstfreier gehen sie in eine Prüfung. Dass ihre Selbstreflexion und Selbsteinschätzung – als Team und als Person – Teil des Prozesses ist, gibt zusätzliche Sicherheit. Nicht anders verhält es sich in der Prüfungssituation. Allein das Wissen darum gibt Gelassenheit.

Die Eigenständigkeit einer Prüfungsleistung bemisst sich nach diesem Ansatz nicht danach, ob ein Inhalt reproduziert werden kann, sondern wie kompetent Problemlösung durch Reorganisation und Transfer möglich wird: Der Kontext einer Lern- und beruflichen Handlungssituation ist vorgegeben. Die Lernenden bringen zusätzlich ihre eigenen Erfahrungen und Kompetenzen ein, reflektieren sie auf die jeweilige Lernsituation, erkennen Aspekte, die noch erschlossen oder recherchiert werden müssen, und verdichten dieses Konglomerat aus Perspektiven und Hintergründen als *vollständige Handlung*. Auf weitergehende Fragen der Lehrenden können sie situativ und angemessen reagieren. Sie reflektieren dabei zusätzlich die Rolle und Funktion der ausgewählten KI-Anwendungen, ebenso ethische Fragen im Zusammenhang mit dieser Nutzung künstlicher Intelligenz. Die obengenannten Kriterien zeigen sich damit nicht nur in den vorbereitenden Teams, sondern auch im nachgelagerten Prüfungsgespräch. Und dieses gestaltet sich konsequenterweise mit jedem Prüfungsteam anders.

Das heißt praktisch: Prüfungsformen im Kontext künstlicher Intelligenz benötigen mehr Zeit auf der menschlichen Ebene, mehr Beziehung, mehr Kommunikation: Dialoge und Trialoge. So kann auch die *Prüfung zu einem Resonanzraum* (Rosa, 2019) werden, in dem sich etwas – je – Einzigartiges ereignet. Und zwar zwischen Menschen, unter Nutzung von Werkzeugen aus künstlicher Intelligenz – im Übrigen ganz praktisch auch für das Protokoll –, die eingeladen, ermutigt und inspiriert (Hüther, 2018) sind ihre Potenziale auch in der Prüfung nicht klassisch „unter Beweis“ zu stellen, sondern im Beziehungsgeschehen des Prüfungsgesprächs weiter zu entfalten.

Nach einem Kompetenzfeststellungsprozess (s. oben) hörte sich das beispielhaft so an: „Es war bemerkenswert für mich. Sie haben keine Theorien mit uns gepaukt, auch keine vorgegeben. Und trotzdem konnte jeder die Aufgaben lösen.“ (O-Ton eines Prüflings) Nicht trotzdem, sondern weil in diesem beispielhaften Format konsequent induktiv vorgegangen wird, funktioniert Lernen als selbstgesteuerte Aneignung. Und evoziert in den Studierenden nicht nur eine entsprechend große Selbstwirksamkeit, sondern auch eine Musteränderung bei angehenden Lehrkräften – weg vom traditionell vorrangigen „Was“, hin zum „Wie“ (Hanstein, 2021). Mein Credo auf dem Weg lautet: „Vertraut dem Prozess!“ Ebenso gilt dies für den Prozess einer Prüfung nach den genannten Kriterien: „Vertraut dem Prüfungsprozess!“ Auf diese Weise kann eine leidenschaftliche Bildung (Burow, 2024) unter adäquatem Einsatz von künstlicher Intelligenz evoziert werden, Pädagogik den Vermittlungshabitus hinter sich lassen und sich weiterentwickeln zu vielfältigsten methodischen (*Prüfungs-*)*Formen der Ermöglichung* (in Anlehnung an Arnold, 2012).

Ein (vermeintliches) Paradoxon gilt es dazu allerdings zu überwinden: Obwohl neue Technologien Anwendung finden, wächst nicht die (tradierte) mechanistische

Vorstellung von Lehren und Lernen weiter. Sondern, gerade weil künstliche Intelligenz eine große und ungeahnte Prozessoffenheit entfaltet, wird – und muss – sich mechanistisches Lernen alter Schule als antiquiert herausstellen. KI-Einsatz darf damit als Treiber modernen Lehrens und Lernens verstanden, mit professioneller Gelassenheit angenommen und mit einem offenen konstruktiv-kritischen Geist begleitet werden – wobei diese wissenschaftliche Grundkompetenz nicht nur den: die Prüfende im Blick behält, sondern mindestens so intensiv die KI.

Erklärung zu konkurrierenden Interessen T. Hanstein hat keine Interessenkonflikte zu erklären, die für den Inhalt dieses Manuskripts relevant sind.

Literatur

- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing. A revision of Bloom's taxonomy of educational objectives*. Addison-Wesley.
- Arnold, R. (2012). *Wie man lehrt, ohne zu belehren. 29 Regeln für eine kluge Lehre*. Carl Auer.
- Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1965). *Taxonomy of educational objectives*. Longman.
- Buraw, O.-A. (2024). *Mit KI zu leidenschaftlicher Bildung*. Beltz.
- Handke, J. (2020). *Handbuch Hochschullehre digital. Leitfaden für eine moderne und mediengerechte Lehre* (3. Aufl.). Tectum.
- Hanstein, Th. (2021). Die „5W“ der Unterrichtsplanung. <https://youtu.be/edt7lxZZTFU?si=kNW8LbCobtULhXHw>. Zugegriffen am 11.03.2026.
- Hanstein, Th. (2024). Planung, Durchführung und Reflexion von Unterricht. Studienheft. Diploma (unv.).
- Hanstein, Th. (2025). Elementarisierung – Renaissance eines Klassikers als Prinzip in veränderten Lernarchitekturen. In T. Hanstein & A. Lanig (Hrsg.), *Berufspädagogik und Bildungsmanagement* (Bd. 1, S. 73–94). Springer VS.
- Hanstein, Th., & Lanig, A. (2020). *Digital lehren. Das Homeschooling-Methodenbuch*. Tectum.
- Hanstein, Th., & Naun, J. (2024). KI-Einsatz an Schulen. Umkehrung klassischer Lehr-Lern-Arrangements. In *Zeitschrift für Sozialmanagement* 2/2024: 75–81.
- Harrow, A. J. (1972). *A taxonomy of the psychomotor domain: A guide for developing behavioral objectives*. Longman.
- Hüther, G. (2018). *Etwas vom Hirn, bitte*. Vandenhoeck & Ruprecht.
- Klippert, H. (2001). *Eigenverantwortliches Arbeiten und Lernen*. Beltz.
- Rosa, H. (2019). *Resonanz. Eine Soziologie der Weltbeziehung*. Suhrkamp.
- Schulentwicklung, NRW. (2021). https://www.schulentwicklung.nrw.de/cms/upload/Faecher_Seiten/Sport/digi/M4_SAMR_Didakt-Mehrwert-digitaler-Medien_2021-01-18.pdf. Zugegriffen am 13.03.2026.
- Weinert, F. (2001). *Leistungsmessungen in Schulen*. Beltz.

Prof. Dr. Thomas Hanstein ist Studiengangsleiter des Masterstudiums Berufspädagogik (M.A.) der DIPLOMA Hochschule am Studienzentrum Friedrichshafen, wissenschaftlicher Leiter am Institut für Lehrerbildung und Berufsbildungsforschung (ILBF) sowie Mitglied im Schulleitungsteam der Bernd-Blindow-Schulgruppe.