ISSN: 2190-4790

LBzM Ludwigsburger Beiträge zur Medienpädagogik

Online-Magazin des Interdisziplinären Zentrums für Medienpädagogik und Medienforschung an der PH Ludwigsburg

Vom Self-Automator zum Zentauren: Theoriegeleitete KI-Steuerung als Gegenstrategie zur Entpädagogisierung

Monika Ueckert

Dieser Beitrag basiert auf der Masterarbeit *„KI-gestützte Konzeption von Lehrmaterialien nach medienpädagogischen Kriterien – Entwicklung und Evaluation eines Frameworks für die Generierung von H5P-Elementen für MINT-Lehrveranstaltungen“*, eingereicht und erfolgreich verteidigt an der FH Südwestfalen im April 2026. Die empirischen Primärdaten (Forschungstagebuch, Experteninterviews, Studierendenbefragung) wurden im Rahmen dieser Qualifikationsarbeit erhoben.

Zusammenfassung

Der Einsatz generativer Künstlicher Intelligenz (KI) verspricht enorme Effizienzgewinne in der Hochschullehre (z. B. durch Zeitersparnis bei der Foliensatz-Erstellung) birgt jedoch das Risiko einer schleichenden Entpädagogisierung, wenn didaktische Entscheidungen unreflektiert an Algorithmen delegiert werden. Dieser Beitrag stellt das theoriegeleitete *KMLE-Framework* (KI-gestützte Materialentwicklung – Lehre) vor, das diesem Risiko systematisch begegnet. In einer Design-Based-Research-Studie über zwei Makro-Zyklen (in den Lehrveranstaltungen Elektrotechnik und Physik) wurden fünf etablierte lerntheoretische und instruktionsdidaktische Ansätze (u. a. Cognitive Load Theory, ARCS-Modell) in maschinenlesbare Prompt-Restriktionen für die Generierung von H5P-Materialien übersetzt. Die Evaluation zeigt eine deutliche Verbesserung: Durch die theoriegeleitete Steuerung sank die Fehlerquote des KI-Outputs von initial 85 % auf 10–15 %, während die Studierenden eine hohe Akzeptanz und einen starken Transferwunsch (M = 4,72/5) zeigten. Der Beitrag demonstriert, wie eine bewusste Mensch-KI-Kollaboration im *Zentauren-Modus* einem didaktischen Deskilling entgegenwirkt und lerntheoretische sowie instruktionsdidaktische Prinzipien in algorithmische Steuerungsprinzipien transformiert.

Schlüsselbegriffe: ● *Medienpädagogik* ● *Generative KI* ● *Design-Based Research* ● *Cognitive Load Theory* ● *Entpädagogisierung* ● *Prompt Engineering* ● *Hochschullehre* ● *MINT-Didaktik*

1. Einleitung

Noch nie war es so einfach, digitale Lernmaterialien für die Hochschullehre zu erstellen und noch nie war die Frage nach ihrer didaktischen Qualität so dringend. Der KI-Monitor 2025

(vgl. Budde/Tobor 2025) dokumentiert eine rasante Dynamik: 97% der befragten Hochschulen entwickeln Unterstützungsangebote für Lehrende im Umgang mit generativer KI, 96% bieten bereits Workshops an (vgl. Budde/Tobor 2025, S. 13, 28). Auch das HFD-Arbeitspapier 87 belegt eine breite, zunehmend selbstverständliche KI-Nutzung in Studium und Lehre (vgl. Wannemacher et al. 2025, S. 6–8). Wichtig bleibt dabei die entscheidende Folgefrage: Wie wird sichergestellt, dass KI-gestützt entwickelte Materialien tatsächlich lernwirksam sind?

Erste Befunde deuten darauf hin, dass der Einsatz häufig explorativ und toolgetrieben erfolgt und Qualitätsfragen erst nachgelagert gestellt werden. Damit entsteht ein Spannungsfeld, das Thomas Damberger (2023) als Entpädagogisierung beschreibt. Gemeint ist die Reduktion bewusster didaktischer Entscheidungen zugunsten automatisierter KI-Vorschläge. In der Folge droht ein schleichender Kompetenzverlust bei Lehrenden, den Doris Weßels und Miriam Maibaum (2026) als Deskilling bezeichnen. Die kognitive Auseinandersetzung mit Lernzielen, Kompetenzstufungen und mediendidaktischen Prinzipien wird nicht mehr aktiv vollzogen, wenn sie vollständig an die KI abgegeben wurde.

Der vorliegende Beitrag knüpft genau an dieser Lücke an. Er stellt die Entwicklung und Evaluation des KMLE-Frameworks (KI-gestützte Materialentwicklung – Lehre) vor, die im Rahmen der zugrunde liegenden Masterarbeit erfolgten. Das Framework übersetzt fünf empirisch fundierte lerntheoretische Qualitätskriterien in maschinenlesbare KI-Restriktionen und macht KI so nicht als Entscheidungsautomatismus, sondern als Instrument didaktischer Professionalisierung nutzbar. Entwickelt, erprobt und empirisch validiert wurde es im Rahmen eines Design-Based-Research-Prozesses (DBR) über zwei Makro-Zyklen in Hochschullehrveranstaltungen (Elektrotechnik und Physik). Der Beitrag zeigt, dass eine theoretisch fundierte Steuerung der KI durch den Menschen, der sogenannte Zentauren-Modus (vgl. Kapitel 2.2), sowohl die fachliche und didaktische Qualität der generierten Materialien deutlich verbessert als auch dem Deskilling-Risiko strukturell entgegenwirkt.

Der Beitrag ist wie folgt aufgebaut: Nach dem theoretischen Rahmen (Kapitel 2) werden Forschungsdesign und Methodik erläutert (Kapitel 3), gefolgt von der Beschreibung des KMLE-Frameworks (Kapitel 4). In Kapitel 5 werden die empirischen Befunde präsentiert und in Kapitel 6 erfolgt deren theoretische Einordnung. Im siebten Kapitel werden die Erkenntnisse zusammengefasst und ein Forschungsausblick ausformuliert.

2. Theoretischer Rahmen

2.1 Entpädagogisierung und das Deskilling-Risiko

Die Debatte um KI in der Bildung wird häufig entlang einer Effizienz-Qualitäts-Achse geführt. Aktuelle Beiträge zeigen das Potenzial generativer KI zur Entlastung Lehrender bei der Aufgaben- und Materialdifferenzierung (vgl. Bahr et al. 2024) ebenso wie zur kreativen Personalisierung von Lernprozessen an Hochschulen (vgl. Martin/Pengel 2024). Die zum Teil erheblichen Zeitersparnisse bei der Materialentwicklung stehen dem von Damberger (2023) formulierten Entpädagogisierungsrisiko gegenüber. Dieses Risiko konkretisiert sich in einer Reihe grundlegender Entscheidungen, die bei unreflektierter KI-Nutzung nicht mehr bewusst getroffen, sondern stillschweigend an den Algorithmus delegiert werden: Welche Lernziele werden auf welcher Kompetenzstufe angestrebt? Welche kognitive Belastung ist angemessen? Wie wird

Motivation systematisch gestaltet? Die implizite Delegation an KI-Systeme kann als Substitution pädagogischer Handlungshoheit betrachtet werden und mittelfristig zu einem Erosionsprozess fachlicher Expertise führen.

Dieses Risiko ist nicht theoretisch abstrakt. Weßels und Maibaum (2026) unterscheiden zwischen Deskilling, dem Verlust bestehender Kompetenzen durch Delegation, und Newskilling bzw. Upskilling, dem Erwerb neuer Kompetenzen durch bewusste Auseinandersetzung mit KI-Möglichkeiten und -grenzen. Entscheidend ist also nicht die KI-Nutzung an sich, sondern die Art der Nutzung. Eine theoriegeleitete, reflexive Nutzung kann bspw. Newskilling anstoßen. Diese konzeptionelle Unterscheidung bildet den Ausgangspunkt für das im weiteren Verlauf entwickelte KMLE-Framework (Kapitel 4), dessen Newskilling-fördernde Wirkung anschließend empirisch geprüft wird (Kapitel 5).

2.2 Taxonomie der Mensch-KI-Kollaboration: Vom Self-Automator zum Zentauren

Für die Analyse unterschiedlicher Nutzungsmodi generativer KI bieten Randazzo et al. (2025, S. 13) eine aufschlussreiche Taxonomie. Sie unterscheiden drei Grundtypen der Mensch-KI-Kollaboration:

Self-Automator (Abdicated Co-Creation)
Der Mensch delegiert Inhalt und Entscheidungsstruktur an die KI. Die Outputs werden weitgehend übernommen, eine kritische Reflexion erfolgt kaum. Dieser Modus maximiert die kurzfristige Effizienz auf Kosten von Qualitätskontrolle und didaktischer Handlungshoheit.

Zentaur (Directed Co-Creation)
Der Mensch definiert Ziele, Restriktionen und Qualitätskriterien, die KI führt operativ aus. Der Mensch behält die strategische Kontrolle und Verantwortung. Der Begriff *Zentaur* ist bewusst metaphorisch-plakativ und stammt aus der internationalen Managementliteratur (vgl. Randazzo et al. 2025)[1]. Er wird hier trotz seiner populärwissenschaftlichen Prägung übernommen, um Anschlussfähigkeit an die aktuelle Human-AI-Collaboration-Diskussion herzustellen.

Cyborg (Fused Co-Creation)
Mensch und KI sind so eng verflochten, dass die Grenzen der jeweiligen Beiträge kaum noch trennbar sind. Dieser Modus setzt hohe Kompetenzen beim Umgang mit KI-basierten Tools und kritische Reflexionsfähigkeit voraus.

Für den Kontext der hochschulischen Lernmaterialentwicklung ist der Zentauren-Modus aus mediendidaktischer Perspektive das zielführende Leitbild. Er verbindet die Stärken der KI (Operationsgeschwindigkeit, Konsistenz, skalierbare Textgenerierung) mit menschlicher Expertise (Lernzielbestimmung, didaktisches Qualitätsurteil, fachliche Korrektheitsprüfung, Wissen über

[1] „Zentaur“ bezeichnet in Anlehnung an das mythologische Mischwesen die Mensch-Maschine-Teams des sogenannten „Centaur Chess“, das Kasparovs nach seiner Niederlage gegen Deep Blue (1997) etablierte. Der Mensch übernimmt die strategische Steuerung, die Maschine die taktische Ausführung (vgl. Kasparov/Greengard 2017). Die Managementliteratur überträgt die Metapher in Bezug auf die Mensch-KI-Kollaboration (vgl. Randazzo et al. 2025).

die Zielgruppe). Die Zentauren-Metapher ist dabei nicht nur Rollenbeschreibung, sondern epistemischer Modus (vgl. Randazzo et al. 2025; Geyer 2026). Der Mensch muss seine didaktischen Entscheidungen explizit formulieren, um sie der KI übermitteln zu können, ein Prozess, der selbst lernwirksam ist.

2.3 Lerntheoretische Qualitätskriterien als Steuerungsgrundlage

Das KMLE-Framework integriert fünf empirisch fundierte lerntheoretische Ansätze, die als Operationalisierungsgrundlage für KI-Prompts dienen. Sie werden hier in ihrer Funktion als Qualitätssicherungssystem skizziert.

Cognitive Load Theory (CLT)
John Sweller, Jeroen van Merriënboer und Fred Pass (2019) unterscheiden – mit Blick auf die Lehrenden – zwischen intrinsischer Belastung (Komplexität des Inhalts), extrinsischer Belastung (durch schlechtes Design verursacht) und lernförderlicher Belastung (Germane Load). Für die Materialgestaltung folgt daraus das Prinzip der Informationsportionierung: Informationsblöcke (Chunks) von maximal 150 Wörtern, konsistente Strukturschemata, keine parasitären Gestaltungselemente. Im KMLE-Framework wird CLT zu einem algorithmischen Steuerungsprinzip. Die Größenbegrenzung von Inhaltsbausteinen wird als Hard Constraint in den Prompt-Wizard (ein im Laufe der Forschung entwickelter HTML-Prompt-Generator) einprogrammiert und *vor jeder Generierung* als Randbedingung gesetzt.

ARCS-Modell
Das Motivationsdesign-Modell nach Keller (2010) strukturiert Lernmotivation entlang von vier Dimensionen: Attention (Aufmerksamkeit), Relevance (Bedeutsamkeit), Confidence (Selbstwirksamkeit) und Satisfaction (Befriedigung). Es liefert einen normativen Rahmen für die Gestaltung von Einstiegsseiten, Übungsformaten und Feedbackstrukturen und wird im Framework als motivationaler Qualitätsprüfpunkt operationalisiert.

Bloom's Taxonomie / Revidierte Taxonomie
Die hierarchische Kompetenzstufung (Erinnern, Verstehen, Anwenden, Analysieren, Evaluieren, Erschaffen) dient als Grundlage für die Differenzierung von Aufgabenformaten nach kognitiver Anforderung (vgl. Anderson/Krathwohl 2001). Das KMLE-Framework nutzt die Bloom-Stufen als Selektionskriterium. Jeder generierte H5P-Interaktionstyp wird einer Bloom-Stufe zugeordnet und die Verteilung der Stufen wird als Checkpoint (vgl. Tabelle 1) vor der Veröffentlichung geprüft.

Constructive Alignment
Das Kohärenzprinzip fordert eine schlüssige Abstimmung zwischen Lernzielen, Lehr-Lern-Aktivitäten und Prüfungsformen (vgl. Biggs/Tang/Kennedy 2022). Im KMLE-Framework manifestiert sich dies im Quellentreue-Prinzip. Die Ableitung sämtlicher Inhalte aus den von Lehrenden bereitgestellten Quellmaterialien ist zwingend erforderlich, um die curriculare Kohärenz im Sinne des Constructive Alignment (vgl. Biggs/Tang/Kennedy 2022) mit fachlicher

Korrektheit zu verbinden. Diese Bindung an kuratiertes Quellmaterial dient zugleich der Vermeidung von KI-Halluzinationen und inhaltlichen Verzerrungen (Bias). Das Framework funktioniert in dieser Hinsicht analog zu einem Expertensystem, dessen Wissensbasis bewusst auf geprüfte, fachlich verantwortete Quellen begrenzt wird, statt auf das ungefilterte Trainingswissen des Sprachmodells zurückzugreifen.

Multimedia Learning Principles
Mayers Cognitive Theory of Multimedia Learning (CTML) formuliert evidenzbasierte Gestaltungsprinzipien für multimediale Lernmaterialien, darunter das Redundanzprinzip (keine redundante Text-Bild-Kombination), das Segmentierungsprinzip (Aufteilung in lernbare Einheiten) und das Kohärenzprinzip (Reduktion auf das Lernrelevante) (vgl. Mayer 2021). Die genannten Prinzipien werden im Framework als Negativliste in den Prompt-Wizard, eine strukturierte Eingabemaske, eingetragen. Über diese wird die Übertragung der theoriegeleiteten Vorgaben in die KI-Generierung realisiert (eine ausführliche Darstellung findet sich in Kapitel 4.2). Der Prompt-Wizard teilt somit mit, welche Tätigkeiten die KI explizit *nicht* ausführen darf.

Die theoretische Innovation des KMLE-Frameworks liegt darin, diese fünf lerntheoretischen Ansätze aus abstrakten Qualitätsnormen in maschinenlesbare, operative Restriktionen zu überführen. Dies entspricht dem, was Thomas Knaus als Verschiebung vom konzeptionellen Technologie-Diskurs hin zur in Technik eingeschriebener Normativität beschreibt, mit dem entscheidenden Unterschied, dass hier nicht die Pädagogik der Technologie, sondern die Technologie der Pädagogik untergeordnet wird (vgl. Knaus 2024, S. 5).

3. Methodisches Vorgehen

3.1 Design-Based Research als methodischer Rahmen

Die Studie folgt dem Paradigma des Design-Based Research (DBR) nach Susan McKenney und Thomas Reeves (2019), das sich durch eine enge Verzahnung von Theorieentwicklung und Praxiserprobung in iterativen Zyklen auszeichnet. DBR ist insbesondere für komplexe Bildungsinterventionen geeignet, deren Wirksamkeit nicht unter Laborbedingungen, sondern im Echtbetrieb untersucht werden muss (vgl. McKenney/Reeves 2019, S. 7). Der zyklische Prozess (Plan → Act → Observe → Reflect) ermöglicht es, das Framework als Reaktion auf empirische Befunde iterativ weiterzuentwickeln.

Die Studie umfasst zwei Makro-Zyklen: einen Entwicklungs- und Validierungszyklus in der Lehrveranstaltung *Elektrotechnik* (WS 2024/25, SS 2025 und WS 2025/26) und einen Transfer- und Validierungszyklus in der Lehrveranstaltung *Ingenieurwissenschaftliche Grundlagen (Physik)* (WS 2025/26). Die unterschiedliche Dauer ergibt sich aus der jeweiligen Funktion im DBR-Prozess. Der Elektrotechnik-Zyklus diente der iterativen Entwicklung des Frameworks selbst über mehrere Plan-Act-Observe-Reflect-Schleifen. Die berichtete Evaluation bezieht sich auf das abschließende WS 2025/26. Der Physik-Zyklus prüfte dagegen nur die Übertragbarkeit des bereits entwickelten Frameworks, weshalb ein Semester als ausreichend erachtet wurde.

Beide Zyklen fanden an der Hochschule Hamm-Lippstadt statt. Die Grundgesamtheit umfasste Studierende mehrerer technischer Studiengänge (u. a. Energietechnik und Ressourcen-

optimierung, Umweltmonitoring und forensische Chemie) sowie, im Physik-Zyklus des Studiengangs Technisches Management und Marketing, eines wirtschaftswissenschaftlich geprägten Studiengangs mit ingenieurwissenschaftlichen Anteilen, der nicht als rein technischer Studiengang einzuordnen ist. Erkenntnistheoretisch folgt DBR einem pragmatischen Paradigma. Es zielt nicht auf universelle Gesetzmäßigkeiten, sondern auf transferierbare Design-Prinzipien, die in ähnlichen Kontexten Anwendung finden können (vgl. McKenney/Reeves 2019, S. 85). Die externe Validität ist damit analytisch-generalisierend, nicht statistisch-inferenziell zu verstehen.

3.2 Datenerhebung und Triangulation

Die Evaluation der Wirksamkeit des KMLE-Frameworks, operationalisiert über die Entwicklung der Fehlerquote des KI-Outputs sowie die studentische Akzeptanz, erfolgt im Rahmen des übergeordneten DBR-Prozesses über einen Mixed-Methods-Ansatz mit methodischer Datentriangulation (vgl. Kuckartz/Rädiker 2024, S. 37-42). Während DBR den forschungsmethodischen Gesamtrahmen der Studie bildet (iterative Entwicklung und Erprobung in Praxiszyklen), stellt der Mixed-Methods-Ansatz die konkrete Datenstrategie innerhalb dieses Rahmens dar. Dabei wurden die folgenden Erhebungsinstrumente miteinander kombiniert:

Forschungstagebuch
Über den gesamten Entwicklungsprozess wurde ein systematisches Forschungstagebuch (Tool: Obsidian) geführt, das Entwicklungsentscheidungen, Krisenpunkte, Iterationsschritte und emergente Erkenntnisse dokumentiert. Die qualitative Analyse erfolgte nach Kuckartz und Rädiker (2024) durch eine inhaltlich strukturierende Inhaltsanalyse.

Experteninterviews
Leitfadengestützte Interviews mit den beteiligten Lehrenden (Elektrotechnik und Physik) lieferten Daten zur wahrgenommenen fachlichen Korrektheit, didaktischen Passung und Praktikabilität des Workflows.

Studierendenbefragung
Eine standardisierte quantitative Befragung mit 5-stufiger Likert-Skala (1 = stimme gar nicht zu, 5 = stimme voll zu) erhob die studentische Akzeptanz entlang der ARCS-Dimensionen, der CLT-Dimensionen Verständlichkeit und Strukturierung sowie eines Transfer-Wunsch-Items. Deskriptive Statistik (Mittelwerte, Standardabweichungen, Top-2-Box-Werte) wurden berechnet. Aufgrund der Stichprobengröße (N = 45, verteilt über Elektrotechnik N = 17 und Physik N = 28) und des Fehlens einer Kontrollgruppe sind inferenzstatistische Generalisierungen nicht angemessen; die Befunde sind als validierende Indikatoren im DBR-Kontext zu interpretieren.

Als vierte Datendimension wurde die Fehlerquote KI-generierter H5P-Materialien im Längsschnitt dokumentiert: Als Fehler galten inhaltliche oder didaktische Abweichungen, die im Checkpoint-Verfahren identifiziert und im Forschungstagebuch nach einem 6-Punkte-Kategoriensystem erfasst wurden (vgl. Tabelle 1). Die Klassifikation erfolgte durch mich allein, ergänzt

durch die Einschätzungen des Elektrotechnik-Lehrenden in den wöchentlichen Jour-fixe-Sitzungen. Ein formales Inter-Rater-Reliability-Verfahren war im DBR-Kontext nicht realisierbar und stellt eine methodische Limitation dar (vgl. Kapitel 6.3). Die Fehlerquoten wurden je Zyklus als prozentualer Anteil fehlerbehafteter H5P-Elemente an allen im jeweiligen Checkpoint geprüften Elementen berechnet.

Tabelle 1: 6-Punkte-Checkpoint-System des KMLE-Frameworks (Operationalisierung der Fehlerquoten)
Anmerkung: Fehlerquote = prozentualer Anteil H5P-Elemente mit mindestens einem positiven Checkpoint-Befund an allen geprüften Elementen (eigene Erhebung; Forschungstagebuch 2025/2026).

#	Checkpoint	Theoretische Basis	Fehler-Kriterium
1	Chunk-Größe	CLT (Sweller/van Merriënboer/Pass 2019)	Textblock > 150 Wörter
2	Strukturschema	CLT/MLP (Mayer 2021)	Abweichung vom 5-Teile-Schema (Einleitung – Wissensvermittlung – Übung – Zusammenfassung – Abschluss)
3	ARCS-Elemente	ARCS-Modell (Keller 2010)	Einführungsseite fehlt oder enthält keine Relevanzherstellung
4	Bloom-Passung	Bloom (Anderson/Krathwohl 2001)	Interaktionstyp passt nicht zur gewählten Kompetenzstufe
5	Quellentreue	Constructive Alignment (Biggs/Tang/Kennedy 2022)	Inhalt nicht aus Lehrenden-Quellmaterial ableitbar (Halluzination)
6	Fachliche Korrektheit	Expertenurteil (Lehrende)	Sachlich inkorrekte Aussage, nachgewiesen durch Fachprofessor*in

4. Das KMLE-Framework: Konzept und Implementierung

4.1 Operationalisierungslogik: Theorie → Prompt → Checkpoint

Das Kernprinzip des KMLE-Frameworks ist eine Operationalisierungslogik auf drei Ebenen, die didaktische Theorie in technische Praxis übersetzt (vgl. Abb. 1). Das Framework ist dabei kein statisches Modell, sondern in den iterativen DBR-Prozess eingebettet. Die Operationalisierungslogik Theorie → Prompt → Checkpoint entspricht dem Mikro-Zyklus Act → Observe innerhalb des übergeordneten DBR-Makro-Zyklus (Plan → Act → Observe → Reflect) nach McKenney/Reeves (2019).

Der Kern des Frameworks ist eine Pyramide mit drei Ebenen:

Ebene 1: Theorie
Aus den fünf lerntheoretischen Qualitätskriterien werden konkrete Gestaltungsprinzipien abgeleitet, die als normative Grundlage dienen. Beispiel aus der Cognitive Load Theory: „Jeder Textblock eines H5P-Elements darf maximal 150 Wörter umfassen."

Ebene 2: Prompt
Die Gestaltungsprinzipien werden als Hard Constraints in das Prompt-System des Prompt-Wizards eingetragen. Hard Constraints sind nicht verhandelbar. Die KI darf diese Grenzen

nicht überschreiten, unabhängig von inhaltlichen Erwägungen. Sie unterscheiden sich von Soft Constraints (Empfehlungen, die die KI abwägen kann). Diese Unterscheidung ist zentral, da Large Language Models (LLMs) ohne klare Restriktionen dazu neigen, Vollständigkeit und Detailtiefe über Kognitionseignung zu priorisieren.

Ebene 3: Checkpoint
Nach jeder Generierungsrunde erfolgt eine manuelle Qualitätsprüfung anhand eines 6-Punkte-Checklisten-Schemas, das alle fünf Theoriedimensionen abdeckt. Erst wenn alle Checkpoints grünes Licht geben, wird das Material zur weiteren Bearbeitung freigegeben. Dieser Checkpoint ist der entscheidende Mechanismus gegen Entpädagogisierung: Er zwingt Lehrende zur expliziten didaktischen Urteilsbildung.

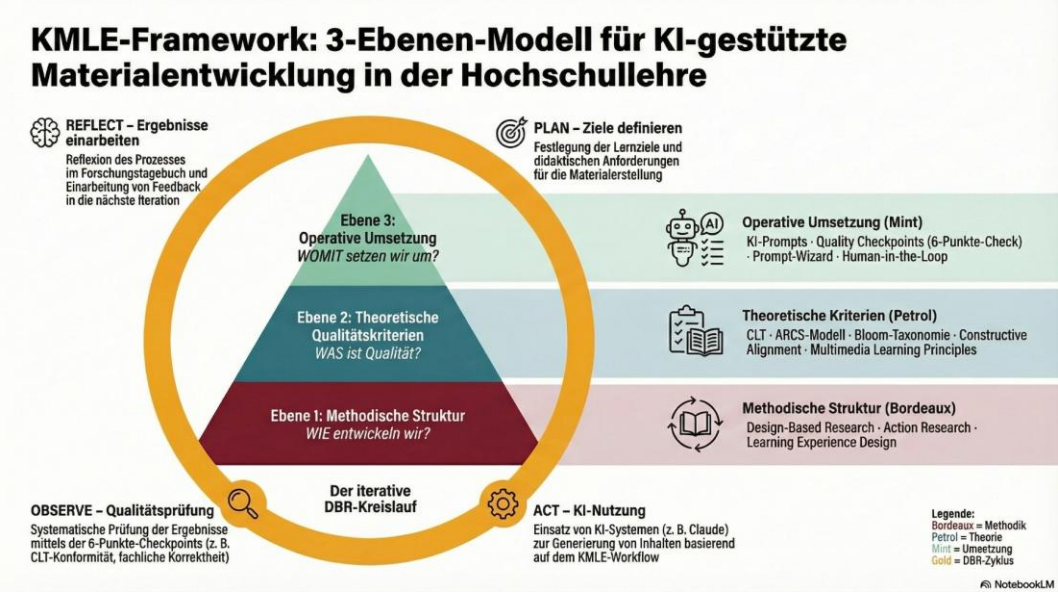


Abbildung 1: Illustration des KMLE-Frameworks (eigene Darstellung mit grafischer KI-Unterstützung, NotebookLM)

Die Logik Theorie → Prompt → Checkpoint konstituiert, was in der Arbeit als Zentauren-Modus bezeichnet wird: Nicht die KI entscheidet über didaktische Qualität, sondern der Mensch setzt die Parameter, und die KI operiert innerhalb dieser Parameter. Die pädagogische Handlungshoheit soll dieser Logik zufolge beim Lehrenden verbleiben – eine Prämisse, die dem Framework zugrunde liegt und die in Kapitel 6.4 aus postdigitaler Perspektive kritisch diskutiert wird.

4.2 Der Prompt-Wizard als technischer Enabler

Der Prompt-Wizard (HTML-basiert) ist die technische Implementierung des KMLE-Frameworks (Enabler). Er stellt eine strukturierte Benutzeroberfläche bereit, die von den Lehrenden eigenständig bedient wird. Sowohl der Elektrotechnik- als auch der Physik-Lehrende nutzen den Wizard selbst zur Erstellung der von ihnen verantworteten Materialien und trugen dabei

die theoretischen Qualitätskriterien vor der KI-Generierung selbst ein.[2] In der zum Zeitpunkt der Forschung aktuellen Version (v3.2) umfasst der Wizard folgende Konfigurationsbereiche:

(1) Quellentreue-Parameter: Das Wizard-Interface verlangt vor jeder Generierung die Einspeisung der Quellmaterialien der Lehrenden (Vorlesungsfolien, Skripte). Das Quellentreue-Prinzip besagt, dass die KI ausschließlich auf Basis dieser Materialien generiert und keine autonomen Ergänzungen aus dem eigenen Training vornimmt. Dieses Prinzip ist die zentrale Sicherung gegen fachliche Halluzinationen.

(2) CLT-Regelwerk: Automatische Einbindung der Hard Constraints (Wortlimitierung, Chunk-Größe, Strukturschema) in jeden Prompt.

(3) Bloom-Selektor: Lehrende wählen explizit aus, welche Bloom-Stufe(n) das zu generierende H5P-Element adressieren soll. Diese Wahl steuert den Interaktionstyp-Selector des Wizards.

(4) ARCS-Motivationsparameter: Gestaltungsanweisungen für die Einführungsseite jedes H5P-Elements (Relevanzherstellung, Alltagsanbindung, Kompetenztransparenz).

(5) Qualitäts-Checkpoint-Liste: Eine abzuarbeitende Checkliste, die nach jeder Generierung zu durchlaufen ist und alle fünf Theoriedimensionen abdeckt.

Technisch ist der Prompt-Wizard als variables, KI-system-unabhängiges Werkzeug konzipiert. Die Prompts sind für gängige Large Language Models (Claude, GPT-4 u. Ä.) optimiert und lassen sich als HTML-gesteuerter Prompt-Generator jederzeit an Fachtermini anpassen.

4.3 Die Krise als Wendepunkt: Vom explorativen zum theoriegeleiteten Prompten

In der explorativen Phase des ersten DBR-Zyklus (April–Oktober 2025), der auf erste kleinere Projekte im vorangegangenen Semester aufbaute, wurde KI zunächst im Self-Automator-Modus genutzt. Materialien wurden primär KI-generiert, Qualitätsprüfungen erfolgten unsystematisch. Die Folge war eine Fehlerquote von rund 85 %, wie beispielsweise Halluzinationen in Fachbegriffen, fehlerhafte Schaltkreisdarstellungen sowie didaktisch unstrukturierte Inhaltsfolgen.

Der Wendepunkt ereignete sich im November 2025 durch ein kritisches Feedback des betreuenden Fachprofessors (Critical Incident). Die negative Rückmeldung erzwang eine grundlegende Revision des Entwicklungsansatzes. Dieser Moment der Krise lässt sich aus einer DBR-Perspektive methodologisch als abduktiver Schluss im Sinne von Charles S. Peirce verstehen (vgl. McKenney/Reeves 2019, S. 92). Aus dem Widerspruch zwischen Erwartung (funktionierende KI-Outputs) und Beobachtung (massive Fehler) entstand die Hypothese, dass die KI theoriegeleitete Leitplanken statt freier Entfaltung benötigt.

[2] Ich selbst setzte den Wizard im Projektverlauf und darüber hinaus zunehmend nur noch zur Kontrolle bereits erstellter Materialien ein. Für die eigene Arbeitsweise mit der KI wurde er durch eine erweiterte „Globale Konvention“ sowie Entscheidungslogs ergänzt, die auf dem Wizard aufbauen, aber stärker auf die individuelle Arbeitsweise zugeschnitten sind (vgl. Kapitel 7).

Konkret operationalisiert wurde diese Hypothese durch eine konsequente Master-Prompt-Logik. Statt einzelner, situativ formulierter Prompts trat ein strukturiertes System expliziter, nicht verhandelbarer Vorgaben (Hard Constraints), aus dem sich die in Kapitel 4.1 beschriebene Operationalisierungslogik entwickelte. Die Krise war damit kein Scheitern, sondern, wie in der Tradition des Action Research (vgl. McKenney/Reeves 2019, S. 31), ein notwendiger Katalysator für die Theoriebildung.

5. Empirische Befunde

5.1 Qualitätsentwicklung: Fehlerquotenreduktion

Die markanteste Wirkung des KMLE-Frameworks zeigt sich in der Fehlerquotenentwicklung über die DBR-Zyklen. Dokumentiert im Forschungstagebuch und durch Expert*inneninterviews trianguliert, verlief die Qualitätsentwicklung wie folgt. Als Fehler galten inhaltliche oder didaktische Abweichungen gemäß dem 6-Punkte-Checkpoint-Schema (vgl. Tabelle 1; zur methodischen Einordnung der Einzel-Klassifikation vgl. Kapitel 3.2 und 6.3).

In der Self-Automator-Phase (Zyklus 1, Phase 1) lag die Fehlerquote bei rund 85 %. Diese Fehler verteilten sich auf die drei Kategorien fachlicher (Checkpoints 5–6), didaktischer (Checkpoints 3–4) und struktureller Fehler (Checkpoints 1–2).

Nach Einführung des KMLE-Frameworks in Phase 3 (Zentauren-Modus) sank die Fehlerquote auf 10–15 %. Dies entspricht einer Fehlerreduktion um ca. 70–75 Prozentpunkte. Die verbleibenden Fehler sind nach Einschätzung der interviewten Lehrenden primär Grenzfälle fachlicher Interpretation und keine strukturellen Defizite des Frameworks.

Der Physik-Transferzyklus bekräftigte diese Entwicklung. Trotz eines neuen fachlichen Kontextes, einer anderen Lehrform (traditionelle Vorlesung statt Flipped Classroom) und eines neuen H5P-Typs (Branching Scenario zusätzlich zum Interactive Book) wurden von der Physiklehrkraft in den bi-wöchentlichen Jour-fixe-Sitzungen ebenfalls nur vereinzelte Korrekturen dokumentiert.

5.2 Studierenden-Akzeptanz und ARCS-Validierung

Die Studierendenbefragung (N = 45) lieferte Daten zur Akzeptanz und zur Wirksamkeit der theoretischen Dimensionen des Frameworks. Die unterschiedliche Teilnahmequote erklärt sich durch die Erhebungsbedingungen. Im Physik-Zyklus fand die Befragung während einer regulären Vorlesungsstunde und beim Erstkontakt mit H5P statt, was Teilnahme und Rückmeldebereitschaft erhöhte; im Elektrotechnik-Zyklus war H5P nur ein ergänzendes Nebenfach-Angebot, entsprechend geringer war die freiwillige Beteiligung. Tabelle 2 fasst die zentralen Ergebnisse zusammen.

Die unterschiedlichen strukturellen Rahmenbedingungen beider Zyklen (Pflicht- bzw. Wahlpflichtcharakter, Erst- versus wiederholter Kontakt mit H5P-Materialien) schränken die Vergleichbarkeit der beiden Teilstichproben ein. Die gemeinsame deskriptive Auswertung in Tabelle 2 ist daher als indikative Gesamttendenz zu verstehen, nicht als Beleg statistischer Äquivalenz zwischen den Gruppen (vgl. zu den methodischen Limitationen Kapitel 6.3).

Tabelle 2: Zentrale Ergebnisse der Studierendenbefragung (5-Punkt-Likert-Skala, 1 = niedrig, 5 = hoch)

Dimension	Elektrotechnik	Physik	Δ
Gesamtakzeptanz	M = 4,34 (SD = 0,77)	M = 4,13 (SD = 0,97)	−0,21
ARCS: Relevance	M = 4,54 (SD = 0,52)	M = 3,92 (SD = 0,93)	−0,62
ARCS: Confidence	M = 4,43 (SD = 0,65)	M = 3,90*	−0,53
ARCS: Satisfaction	M = 4,47 (SD = 0,52)	nicht erfasst	–
ARCS: Attention	M = 3,60 (SD = 1,18)	nicht erfasst	–
CLT: Strukturierung	M = 4,40 (SD = 0,83)	M = 4,08 (SD = 0,95)	−0,32
Transfer-Wunsch	M = 4,59 (SD = 0,62)	M = 4,72 (SD = 0,72)	+0,13

Anmerkung: *Physik-Confidence-Wert aggregiert aus zwei H5P-Formaten (Interactive Book M = 4,18, SD = 1,05; Branching Scenario M = 3,62, SD = 1,02). ARCS Satisfaction und Attention wurden im Physik-Zyklus nicht erhoben, da der auf konzeptionelle Übertragbarkeit angelegte Transferzyklus nur zwei H5P-Elemente umfasste und damit außerhalb des Rahmens einer vollständigen ARCS-Erhebung lag.*

Die hohen Gesamtakzeptanzwerte (ET: M = 4,34; Physik: M = 4,13) belegen eine breite Akzeptanz der Materialien. Besonders aussagekräftig ist der CLT-Strukturierungswert (M = 4,40 ET, M = 4,08 Physik), der die Wirksamkeit des Chunking-Prinzips und des konsistenten 5-Teile-Schemas bestätigt.

Die ARCS-Dimension Attention erreichte in Elektrotechnik den niedrigsten Wert (M = 3,60, SD = 1,18), eine designkonforme Konsequenz des Quellentreue-Prinzips. Nach der Krise im November 2025 wurde bewusst auf nicht in den Lehrenden-Materialien verankerte Alltagsbeispiele verzichtet, fachliche Korrektheit also vor motivationaler Kreativität priorisiert. Dieser Trade-off ließe sich künftig durch gezielte Attention-Strategien innerhalb der fachlichen Quellen adressieren.

Der herausragende Befund ist der Transfer-Wunsch (M = 4,72 Physik, M = 4,59 ET). Er ist didaktisch bedeutsam, weil er nicht nur Zufriedenheit, sondern wahrgenommenen Mehrwert ausdrückt: Studierende, die sich vergleichbare Lerninhalte auch in anderen Modulen wünschen, haben einen Nutzen erlebt, der über das Einzelmodul hinausweist. Die geringe Differenz zwischen den Kontexten (Δ = 0,13) liefert einen ersten, deskriptiven Hinweis darauf, dass die Designprinzipien auch in einem neuen fachlichen Kontext ähnlich wahrgenommen werden; angesichts der Stichprobengröße und des Fehlens eines Signifikanztests ist dies als vorsichtiger Anhaltspunkt, nicht als Beleg statistisch abgesicherter Transferrobustheit zu werten (vgl. Kapitel 6.3).

5.3 Transferierbarkeit des KMLE-Frameworks

Die Transfervalidierung erfolgte entlang zweier Dimensionen: des *Intra-Fach-Transfers* (Elektrotechnik-Grundlagen auf weitere Studiengänge: Biomedizintechnik, Umweltmonitoring und forensische Chemie, Energietechnik und Ressourcenoptimierung, Intelligent Systems Design) und des *Inter-Fach-Transfers* (Elektrotechnik → Physik).

Beim Intra-Fach-Transfer konnten 80–90 % der H5P-Inhalte ohne substantielle Überarbeitung auf weitere Studiengänge übertragen werden. Der verbleibende Anteil studiengangsspezifischer Anpassungen (10–20 %) betrifft marginale Inhaltsbereiche (spezifische Themengebiete einzelner Studiengänge), nicht die Framework-Struktur.

Der Inter-Fach-Transfer in die Physik umfasste vier simultane Transformationen: (1) Fachdomäne (Elektrotechnik → Physik), (2) Lehrformat (Flipped Classroom → traditionelle Vorlesung), (3) Zielgruppe (technikaffine Ingenieurstudiengänge → breites Spektrum inkl. Wirtschaftsstudierender), (4) Inhaltstyp (Interactive Book → Interactive Book + Branching Scenario). Trotz dieser Mehrfachtransformation blieben die Framework-Kernkomponenten – CLT-Regelwerk, ARCS-Parameter, Bloom-Selektor, Checkpoint-Schema – vollständig stabil. Anpassungsbedarf bestand lediglich auf der operativen Ebene (ca. 40 % der Prompts): Fachterminologie und Beispielrahmen wurden aktualisiert; die methodische Tiefenstruktur blieb unverändert.

Dieses Befundmuster belegt die *konzeptionelle Invarianz* des KMLE-Frameworks. Die theoretischen Qualitätskriterien (CLT, ARCS, Bloom) sind domänenübergreifend gültig; lediglich ihre inhaltliche Befüllung ist kontextspezifisch. Dies entspricht der DBR-Prämisse, dass robuste Design-Prinzipien ihren Kern in der Übertragung behalten, während ihre Oberfläche sich anpasst (vgl. McKenney/Reeves 2019, S. 94).

6. Diskussion: Theoretische Einordnung

6.1 Cognitive Load Theory als algorithmisches Steuerungsprinzip

Ein zentraler Beitrag dieser Studie liegt in der Transformation der Cognitive Load Theory (CLT) von einer Gestaltungsempfehlung zu einem algorithmischen Steuerungsprinzip. Die Grundidee, pädagogisches Handeln informationstheoretisch zu formalisieren, ist dabei nicht gänzlich neu. Bereits die Kybernetische Pädagogik der 1960er/70er Jahre verfolgte mit Lerntechnologien und Teaching Machines das Ziel einer *Objektivierung pädagogischer Tätigkeiten* auf Basis der Informations- und Codierungstheorie (vgl. Frank/Meder 1971, S. 30, 86), mit anderen theoretischen Bezügen und einem deutlich geringeren Automatisierungsgrad als im hier vorgestellten Framework. Der eigenständige Beitrag der vorliegenden Studie liegt entsprechend weniger in der Grundidee der algorithmischen Formalisierung selbst als in ihrer konkreten, KI-gestützten Umsetzung im Kontext generativer Sprachmodelle und in ihrer empirischen Erprobung im Hochschulalltag.

Die CLT wurde von Sweller, van Merriënboer und Pass (2019) als Theorie der menschlichen Informationsverarbeitung entwickelt, die ihre Aussagen auf die Lernenden beziehen, nicht auf das Produktionssystem des Lernmaterials. Das KMLE-Framework vollzieht nun eine Umkehrung dieser Perspektive: CLT-Prinzipien werden nicht mehr nur nachträglich als Qualitätsprüfmaßstab angelegt, sondern als *a priori*-Restriktionen in das KI-Generierungssystem eingespeist. Die KI-Outputs müssen CLT-konform sein, bevor sie überhaupt entstehen (vgl. Abbildung 2).

Diese Übertragung ist aus theoretischer Perspektive als nicht trivial zu erachten. Die Prämisse, dass CLT-Prinzipien in der Form „If [Bedingung] then [Restriktion]“ formalisierbar sind, bildet die Grundlage für die weitere Analyse. Die vorliegende Formalisierung ist im Kontext von Large Language Models möglich, da diese Systeme auf natürlichen Sprachbefehlen basieren. Das Chunk-Prinzip, das eine maximale Textblocklänge von 150 Wörtern vorsieht, stellt eine formalisierbare Restriktion dar, die in einem LLM-Prompt implementiert werden kann und empirisch nachweislich zu CLT-konformen Ergebnissen führt.

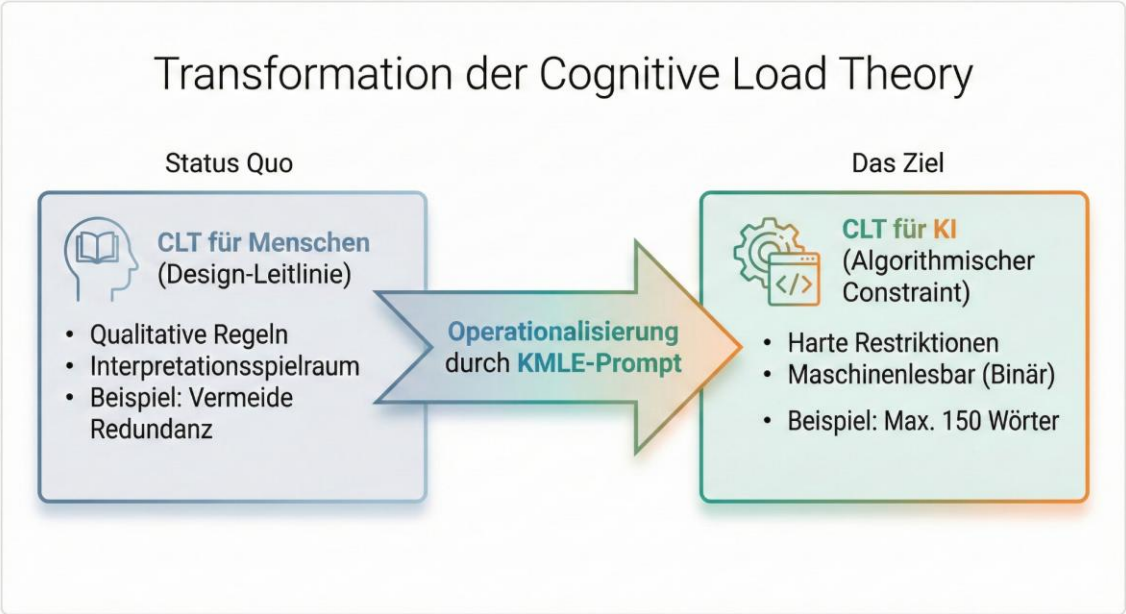


Abbildung 2: Transformation der Cognitive Load Theory (eigene Darstellung mit grafischer KI-Unterstützung, KI: NanoBanana)

6.2 Das Entpädagogisierungs-Paradox und seine Auflösung

Die Befunde bestätigen empirisch die eingangs (Kapitel 1) formulierte These: Die Nutzung von KI zur Materialerstellung kann sowohl Entpädagogisierung verstärken (Self-Automator-Modus) als auch diese aktiv bekämpfen (Zentauren-Modus). Entscheidend ist dabei nicht die KI selbst, sondern die Struktur des menschlichen Inputs. Ein Zusammenhang, der sich im vorliegenden DBR-Prozess wiederholt beobachten ließ.

Der Zentauren-Modus bedingt eine explizite Darlegung der didaktischen Entscheidungen der Lehrenden, etwa zur angestrebten Bloom-Stufe, zur angemessenen Chunk-Größe oder zur adressierten ARCS-Dimension auf der Einführungsseite. Dieser Explizierungszwang wirkt paradox – einerseits bürokratischer als naives Prompten, andererseits pädagogisch produktiver, da er didaktische Reflexion institutionalisiert. Die KI fungiert damit nicht als Entscheidungsersatz, sondern als Spiegel der eigenen Didaktik.

Dies korrespondiert mit Weßels und Maibaums (2026, S. 31) Konzept des Upskillings. Durch die strukturierte Auseinandersetzung mit lerntheoretischen Kategorien im Rahmen des KMLE-Workflows können Lehrende ihr eigenes didaktisches Repertoire explizit machen und erweitern, auch dann, wenn sie diese Theorien vorher nicht systematisch kannten.

6.3 Limitationen und kritische Reflexion

Die Befunde sind im Licht methodischer Limitationen zu interpretieren. Die Stichprobengröße (N = 45) erlaubt keine inferenzstatistischen Generalisierungen. Das Fehlen einer Kontrollgruppe macht kausale Aussagen über die alleinige Wirkung des KMLE-Frameworks methodisch problematisch. Verbesserungen könnten teilweise auch auf den Lerneffekt der Lehrenden selbst, auf die H5P-Plattformnutzung oder auf semesterspezifische Faktoren zurückzuführen sein. Zudem schränken die unterschiedlichen strukturellen Rahmenbedingungen

der beiden Erhebungskontexte (Pflicht- bzw. Wahlpflichtcharakter, Erst- versus wiederholter Kontakt mit H5P-Materialien) die Vergleichbarkeit der beiden Teilstichproben ein; die in Kapitel 5.2 berichteten Gesamtwerte sind entsprechend als deskriptive Tendenzaussage und nicht als Beleg einer statistisch abgesicherten Gruppenäquivalenz zu interpretieren.

Die ökologische Validität der Studie (Echtlehr-Kontext, iterative Erprobung) ist als Stärke zu werten, limitiert aber die Kontrollmöglichkeiten. Zudem ist das KMLE-Framework an spezifische H5P-Elementtypen und an die Nutzung von Large Language Models gebunden. Seine Übertragbarkeit auf andere Content-Formate und zukünftige KI-Architekturen ist ein offenes Forschungsfeld.

Schließlich ist die Rolle der Forschenden als zugleich Entwicklerin und Evaluatorin des Frameworks im Sinne der Aktionsforschung reflexiv zu markieren. Der DBR-Ansatz toleriert diesen doppelten Standpunkt explizit, verlangt jedoch eine kontinuierliche Reflexion der eigenen Perspektiven, die im Forschungstagebuch dokumentiert wurde.

6.4 Souveränität versus verteilte Kompetenz: Eine offene Anschlussfrage

Die im vorangegangenen Abschnitt benannte Doppelrolle der Forschenden als Entwicklerin und Evaluatorin wirft eine grundsätzlichere, über methodische Limitation hinausweisende Frage auf. Die bisherige Argumentation geht davon aus, dass didaktische Handlungshoheit beim Menschen liegen sollte und ihre Erosion durch Deskilling problematisch ist, eine Annahme, die nicht selbstverständlich ist. Michael Kerres und Sarah von Querfurth stellen in ihrer postdigitalen Bildungstheorie die Kategorie der Souveränität selbst zur Disposition. Das Festhalten am souveränen, technikbeherrschenden Subjekt könnte eine nostalgische Reaktion auf eine Konstellation sein, in der Mensch und KI ohnehin untrennbar verwoben sind (vgl. Kerres/Querfurth 2026). Kompetenz wäre demnach keine Eigenschaft der einzelnen Lehrperson mehr, sondern die emergente Leistung eines Geflechts aus Lehrenden, KI-System, Quellmaterial und Prompting-Praxis. Die Verteilung von Kompetenz im Sinne von Rammerts verteilten Aktion und hybrider Intelligenz (vgl. 2016, S. 105–168, 227–242) sowie Dellermanns Hybrid Intelligence (vgl. 2019, S. 637–643).

Diese Perspektive lässt sich nicht ohne Weiteres mit dem zugrunde liegenden Zentauren-Modell vereinbaren, das eine klare Grenze zwischen menschlichem und maschinellem Beitrag voraussetzt: Der Mensch behält die strategische Kontrolle, die KI führt operativ aus (vgl. Randazzo et al. 2025). Genau diese Grenze problematisiert die postdigitale Sicht. Entsteht Kompetenz im Zusammenwirken mehrerer Netzwerk-Knoten, wird die Frage ‚Wer entscheidet – Mensch oder KI?' analytisch weniger trennscharf, als das Zentauren-Modell nahelegt (vgl. Kerres/Querfurth 2026).

Das KMLE-Framework liefert hierzu einen Gegenbefund, der die Diskussion empirisch anreichert, ohne sie aufzulösen. Die Fehlerquotenreduktion (Kapitel 5.1) wurde gerade durch explizite Trennung von menschlicher Entscheidung und maschineller Ausführung erzielt, eine Stärkung, keine Auflösung der Grenze. Das steht im Spannungsverhältnis zum Metabefund von Vaccaro et al. (2024), wonach Mensch-KI-Kollaborationen Einzelleistungen häufig nicht übertreffen. Diese Diskrepanz könnte mit der Struktur der Kollaboration zusammenhängen. Ein klar konturierter Zentauren-Modus mit Hard Constraints dürfte andere Koordinations-

kosten erzeugen als das eher unstrukturierte Zusammenwirken, das Vaccaro et al. (2024) primär untersucht haben dürften, eine Erklärung, die sich mit dem vorliegenden Design allerdings nicht kausal prüfen lässt.

Eine abschließende Klärung, ob Souveränität oder verteilte Kompetenz die tragfähigere Leitkategorie der Mediendidaktik ist, kann dieser Beitrag nicht leisten. Bemerkenswert bleibt jedoch: Die hier vorgelegte Qualitätssteigerung durch explizite Grenzziehung zwischen menschlicher und maschineller Zuständigkeit spricht zumindest für die H5P-Materialentwicklung dafür, diese Trennung nicht vorschnell aufzugeben. Ob sich das auf stärker emergente Mensch-KI-Interaktionen übertragen lässt, bleibt eine offene Forschungsfrage.

7. Fazit und Ausblick

Dieser Beitrag hat gezeigt, dass generative KI in der Lernmaterialentwicklung kein Instrument der Entpädagogisierung sein muss, wenn sie theoriegeleitet eingesetzt wird. Das KMLE-Framework bietet einen praktisch erprobten und empirisch gestützten Ansatz, um fünf etablierte lerntheoretische und instruktionsdidaktische Qualitätskriterien in maschinenlesbare KI-Restriktionen zu übersetzen und damit einen strukturierten Zentauren-Modus der Zusammenarbeit zwischen Mensch und KI zu ermöglichen.

Die zentralen Befunde lassen sich in drei Thesen verdichten:

These 1: Qualitätseffekt
Die Reduktion der fachlichen und didaktischen Fehlerquote von KI-generiertem Lernmaterial von 85% auf 10–15% spricht dafür, dass theoriegeleitete KI-Steuerung im hier untersuchten Setting einen deutlichen Qualitätsgewinn ermöglicht. Der Unterschied zwischen dem Self-Automator- und dem Zentauren-Modus zeigt sich dabei nicht nur graduell, sondern als praktisch relevanter Wechsel der Arbeitslogik.

These 2: Anti-Deskilling-Effekt
Der Explizierungszwang des KMLE-Workflows wirkt einem didaktischen Deskilling entgegen. Lehrende müssen Lernziele, Bloom-Stufen, CLT-Parameter und ARCS-Dimensionen vor jeder Generierung aktiv bestimmen und damit ihre didaktische Urteilskraft ausdrücklich machen. KI fungiert in diesem Kontext als Verstärker und Spiegel vorhandener Expertise, nicht als deren Ersatz.

These 3: Transferprinzip
Die methodische Tiefenstruktur des Frameworks folgt einer stabilen Operationalisierungslogik, Theorie, Prompt und Checkpoint greifen ineinander. Diese Struktur erscheint grundsätzlich auf andere Kontexte übertragbar, bedarf aber in jeder Domäne einer sorgfältigen inhaltlichen Anpassung. Gerade darin liegt ihre Stärke, nicht in der Übertragbarkeit fertiger Inhalte, sondern in der Übertragbarkeit eines kontrollierten Vorgehens.

Für zukünftige Forschung ergeben sich mehrere Ansatzpunkte: Erstens wären Kontrollgruppenstudien mit größeren Stichproben sinnvoll, um die Kausalaussagen weiter abzusichern.

Zweitens sollte in Langzeitstudien untersucht werden, ob der Anti-Deskilling-Effekt auch nach Routinisierung des Workflows bestehen bleibt. Drittens wäre zu prüfen, ob sich das Framework auf andere Content-Formate, etwa Video-Scripting oder Game-based Learning, sowie auf weitere Hochschulkontexte, etwa die Geistes- und Sozialwissenschaften, übertragen lässt. Viertens wäre in einem Prä-Post-Design zu untersuchen, ob und wie sich die Arbeit mit einem solchen Framework auf die didaktische Professionalität der beteiligten Lehrenden auswirkt. Fünftens hat sich die operative Umsetzung des Prompt-Wizards in der Praxis bereits in Richtung eines versionierten Entscheidungslogs weiterentwickelt, das Konventionen, Korrekturprioritäten und offene Entscheidungen fachbereichsübergreifend dokumentiert; eine systematische Auswertung dieser Praxis steht noch aus und ist Gegenstand künftiger Untersuchungen.

Im medienpädagogischen/-didaktischen Diskurs trägt diese Studie damit zu einer zentralen Frage bei: Wie lässt sich KI so einsetzen, dass sie pädagogische Kompetenz stärkt, statt sie zu erodieren? Die vorliegenden Befunde sprechen dafür, dass dies vor allem dort gelingt, wo KI nicht als Ersatz didaktischer Reflexion, sondern als deren strukturiertes Werkzeug verstanden wird. Gerade die Theorie erweist sich dabei nicht als bürokratische Zusatzlast, sondern als Bedingung dafür, dass KI überhaupt in pädagogisch sinnvoller Weise eingesetzt werden kann.

Zugleich macht der anschließende theoretische Diskurs deutlich, dass diese Antwort nicht vorschnell verallgemeinert werden sollte. Die im Beitrag beobachtete Stärkung durch explizite Grenzziehung zwischen menschlicher und maschineller Zuständigkeit ist eine überzeugende Lösung für den hier untersuchten Anwendungsfall der Lernmaterialentwicklung. Ob sich diese Form der Zusammenarbeit auch auf stärker emergente, weniger klar strukturierte Mensch-KI-Konstellationen übertragen lässt, bleibt eine offene Frage. In diesem Sinne ist der hier entwickelte Ansatz nicht als abschließende Antwort, sondern als empirisch begründeter Beitrag zu einer weiter offenen Debatte über Souveränität, verteilte Kompetenz und die zukünftige Rolle didaktischer Professionalität zu verstehen.

Anmerkung:
Der vorliegende Beitrag entstand unter unterstützender Nutzung generativer KI-Systeme (Claude, Perplexity, Gemini/Google), insbesondere zur sprachlichen Verdichtung, Umformulierung und redaktionellen Überarbeitung bereits eigenständig verfasster Textentwürfe. Eine wörtliche Übernahme KI-generierter Formulierungen in zentrale inhaltliche oder theoretische Aussagen fand nicht statt. Alle KI-unterstützten Passagen wurden von der Autorin inhaltlich geprüft und verantwortet. Die fachliche und theoretische Konzeption, die Auswahl und Bewertung der Inhalte sowie die finale Verantwortung für den gesamten Text lagen bei der Autorin. Die hier gewählte Form der Offenlegung folgt etablierten Transparenzkonventionen, auch wenn diese das iterative, ko-konstruktive Verhältnis zwischen Autorin und KI-Systemen (vgl. Dellermann et al. 2019; Randazzo et al. 2025) nur unzureichend abbilden. Es handelt sich um ein Spannungsfeld, das die Wissenschaftskommunikation derzeit insgesamt noch nicht gelöst hat. Stand der KI-Unterstützung: zuletzt August 2026.

Literatur

Anderson, Lorin W./Krathwohl, David R. (Eds.) (2001): A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives: complete edition. New York, Munich: Longman.

Bahr, Tobias/Manzocco, Mario/Schuster, Dennis/Wacker, Albrecht (2024): Textgenerierende KI zur Erleichterung der Arbeit von Lehrpersonen? In: Ludwigsburger Beiträge zur Medienpädagogik, S. 1–16. doi.org/10.21240/lbzm/24/07.

Biggs, John B./Tang, Catherine S./Kennedy, Gregor (2022): Teaching for quality learning at university. Maidenhead: Open University Press McGraw Hill.

Budde, Jannica/Tobor, Jens (2025): KI Monitor 2025. Hochschulen gestalten den KI-Alltag. Berlin: Zenodo.

Damberger, Thomas (2023): Digitalisierung zwischen Entpädagogisierung und ihrer Überwindung. In: academia.edu, S. 1–12.

Dellermann, Dominik/Ebel, Philipp/Söllner, Matthias/Leimeister, Jan M. (2019): Hybrid Intelligence. In: Business & Information Systems Engineering, 61, pp. 637–643.

Frank, Helmar/Meder, Brigitte S. (1971): Einführung in die kybernetische Pädagogik. München: Deutscher Taschenbuch-Verlag.

Geyer, Barbara (2026): Bin ich eine Betrügerin, weil ich mit KI schreibe? In: KI in Lehre und Weiterbildung. barbarageyer.substack.com/p/bin-ich-eine-betrugerin-weil-ich.

Kasparov, Garri/Greengard, Mig (2017): Deep thinking. Where machine intelligence ends and human creativity begins. New York: PublicAffairs.

Keller, John M. (2010): Motivational Design for Learning and Performance. The ARCS Model Approach. Boston, MA: Springer Science+Business Media LLC.

Kerres, Michael/Querfurth, Sarah von (2026): Natürliche und künstliche Intelligenz: „Verteilte Kompetenz“ in einer postdigitalen Bildungstheorie. In: Elvany, Nele/Becker, Michael/Heppt, Birgit/Alscher, Pascal (Hrsg.): Bildung im Wandel – Zwischen technologischer Innovation und gesellschaftlicher Herausforderung. Münster: Waxmann.

Knaus, Thomas (2024): Künstliche Intelligenz und Pädagogik – ein Plädoyer für eine Perspektiverweiterung. In: Ludwigsburger Beiträge zur Medienpädagogik, S. 1–34.

Kuckartz, Udo/Rädiker, Stefan (62024): Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und künstlicher Intelligenz. Weinheim: Juventa.

Martin, Anne/Pengel, Norbert (2024): Die kreative Nutzung von KI zur Personalisierung des Lernens an Hochschulen. In: Ludwigsburger Beiträge zur Medienpädagogik, S. 1–11. doi.org/10.21240/lbzm/24/03.

Mayer, Richard E. (2021): Multimedia Learning. Cambridge: Cambridge University Press.

McKenney, Susan/Reeves, Thomas C. (2019): Conducting educational design research. London, New York: Routledge Taylor & Francis Group.

Rammert, Werner (2016): Technik – Handeln – Wissen. Wiesbaden: Springer Fachmedien.

Randazzo, Steven/Lifshitz-Assaf, Hila/Kellogg, Katherine/Dell'Acqua, Fabrizio/Mollick, Ethan R./Candelon, François/Lakhani, Karim R. (2025): Cyborgs, Centaurs and Self-Automators: The Three Modes of Human-GenAI Knowledge Work and Their Implications for Skilling and the Future of Expertise. Working Paper No. 26-036.

Sweller, John/van Merriënboer, Jeroen J. G./Paas, Fred (2019): Cognitive Architecture and Instructional Design: 20 Years Later. In: Educational Psychology Review, 2, pp. 261–292.

Vaccaro, Michelle/Almaatouq, Abdullah/Malone, Thomas (2024): When combinations of humans and AI are useful: A systematic review and meta-analysis. In: Nature human behaviour, 12, pp. 2293–2303.

Wannemacher, Klaus/Bosse, Elke/Lübcke, Maren/Kaemena, Alena. (2025): Wie KI Studium und Lehre verändert. In: Hochschulforum Digitalisierung, Arbeitspapier Nr. 87, S. 1–74.

Weßels, Doris/Maibaum, Miriam (2026): Vom Deskilling zum Newskilling. Der blinde Fleck emergenter Kompetenzverschiebungen durch generative KI. In: Forschung & Lehre, 2, S. 30–32.

Informationen zur Autorin



Monika Ueckert (M. A. Medienpädagogik [FH SWF], B. A. Bildungswissenschaften [FernUniversität Hagen]) ist als wissenschaftliche Mitarbeiterin für digitale Lehre an der Hochschule Hamm-Lippstadt tätig.

Durch ihre vorherige Tätigkeit im administrativen Hochschulmanagement (Prüfungsamt) vereint sie in ihrer Arbeit eine ganzheitliche institutionelle Perspektive mit modernem Lehr-Lern-Design. In ihrer Forschung beschäftigt sie sich intensiv mit der Frage, wie digitale Bildungsprozesse gefördert und qualitativ gesichert werden können. Ihre aktuellen Forschungsschwerpunkte liegen an der Schnittstelle von mediendidaktischer Theoriebildung, der theoriegeleiteten Nutzung generativer KI in der Hochschullehre (Educational Technology) und der MINT-Didaktik.

Monika.Ueckert@hshl.de; Monika.Ueckert@icloud.com

www.linkedin.com/in/monika-ueckert-hshl

Zitationshinweis:
Ueckert, Monika (2026): Vom Self-Automator zum Zentauren: Theoriegeleitete KI-Steuerung als Gegenstrategie zur Entpädagogisierung. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 26, S. 1–18. doi.org/10.21240/lbzm/26/19.