



# Impact Free

Hochschuldidaktisches Journal

Impact Free 72 – September 2026  
HAMBURG

## Impact Free

### Was ist das?

Impact Free ist eine Publikationsmöglichkeit für hochschuldidaktische Texte,

- die als Vorversionen von Zeitschriften- oder Buch-Beiträgen online gehen, oder
- die aus thematischen Gründen oder infolge noch nicht abgeschlossener Forschung keinen rechten Ort in Zeitschriften oder Büchern finden, oder
- die einfach hier und jetzt online publiziert werden sollen.

### Wer steckt dahinter?

Impact Free ist kein Publikationsorgan der Universität Hamburg. Es handelt sich um eine Initiative, die allein ich, Gabi Reinmann, verantworte, veröffentliche auf meinem Blog (<http://gabi-reinmann.de/>).

Herzlich willkommen sind Gastautoren, die zum Thema Hochschuldidaktik schreiben wollen. Texte von Gastautorinnen können dann natürlich auch in deren Blogs eingebunden werden.

### Und was soll das?

Impact Free war gedacht als ein persönliches Experiment. Falls zu wenige Texte über einen gewissen Zeitraum zusammengekommen wären, hätte ich das Vorhaben wieder eingestellt. Dem ist aber nicht so, sodass ich Impact Free bis auf Weiteres fortsetze. Inzwischen sind die Texte auch über die Staats- und Universitätsbibliothek Hamburg [hier](#) erreichbar.

In diesem Journal mache ich in Textform öffentlich, was mir wichtig erscheint: (a) Gedanken, bei denen ich so weit bin, dass sie sich für mehr als Blog-Posts eignen, (b) Texte, die aus diversen Gründen noch nicht geeignet sind für andere Publikationsorgane, (c) Texte, die in Reviews abgelehnt wurden oder infolge von Reviews so weit hätten verändert werden müssen, dass es meinen Intentionen nicht mehr entspricht, (d) Texte mit hoher Aktualität, für welche andere Publikationswege zu langsam sind, (e) inhaltlich passende Textbeiträge von anderen Autorinnen. Genderschreibweise und Textlänge sind bewusst variabel und können frei gewählt werden.

### Kontakt Daten an der Universität Hamburg:

Prof. Dr. Gabi Reinmann

Universität Hamburg

Hamburger Zentrum für Universitäres Lehren und Lernen (HUL)

Leitung | Professur für Lehren und Lernen an der Hochschule

Jungiusstraße 9 | 20355 Hamburg

[reinmann.gabi@googlemail.com](mailto:reinmann.gabi@googlemail.com)

[gabi.reinmann@uni-hamburg.de](mailto:gabi.reinmann@uni-hamburg.de)

<https://www.hul.uni-hamburg.de/>

<http://gabi-reinmann.de/>

# EIGENSTÄNDIGKEIT IN ZEITEN VON KI – REFLEXION EINES SELBSTBEZÜGLICHEN DIALOGS

GABI REINMANN

## 1. Einführung

**Hintergrund.** Vor ziemlich genau einem Jahr, im September 2025, habe ich zusammen mit Frank Vohle auf Impact Free den Beitrag „Werkstolz mit KI? Autoethnografischer Blick auf einen selbstbezüglichen Dialog“ veröffentlicht (Reinmann & Vohle, 2025). Der Beitrag dokumentiert einen Dialog mit ChatGPT zu der Frage, ob im Einsatz generativer KI beim wissenschaftlichen Schreiben der Werkstolz verloren geht. Im Zuge dieses Dialogs entstand zum einen ein Textentwurf zum Konstrukt des Werkstolzes; zum anderen entwickelten sich entlang der Mensch-Maschine-Interaktion einige (nicht dokumentierte) Seitengespräche zwischen Frank und mir, die zu einer anschließenden Analyse und Reflexion des Chat-Verlaufs führten. Das ganze Unterfangen, sicher nichts Besonderes und bereits letztes Jahr für viele akademisch tätige Menschen fast schon etwas Alltägliches, hat uns in der autoethnografischen Analyse dennoch viel zum Nachdenken angeregt. Der vorliegende Beitrag ist so etwas wie eine (ungeplante) Fortsetzung mit Variation und einem anderen KI-System.

**Mein Schreibmodus.** 2026 bevorzuge ich für akademische Arbeiten tendenziell das KI-System Claude. Man kann aber wohl festhalten: Alle gängigen Chatbots haben sich binnen eines Jahres erneut deutlich verbessert – es ist hier nicht der Ort, diverse Neuerungen aufzuführen und zu kommentieren<sup>1</sup>. Trotzdem schreibe ich meine Texte nach wie vor ohne KI in dem Sinne, dass ich Maschinen für Vorträge und Artikel keine „Textaufträge“ gebe oder um Formulierungen bitte; ob da der Werkstolz oder etwas anderes im Spiel ist, darüber bin ich mir selbst nicht ganz im Klaren. Aber natürlich setze ich KI im Prozess wissenschaftlichen

Arbeitens ein, was entsprechend in die Texterstellung einfließt: vor allem als Unterstützung von Recherchen, für Übersetzungen, zur Konsistenzprüfung theoretischer Argumente etc. Dann und wann, meist spontan angestoßen durch irgendeinen akademischen Anlass, trete ich auch in einen Dialog mit KI zu Themen, die sich im Stadium eines ersten Durchdenkens befinden. Ob es Zufall war oder ob es einen impliziten Grund gibt, warum ich mit der Maschine in einen längeren Dialog ähnlich dem im Sommer 2025 zu einem Thema eingestiegen bin, das eine innere Verwandtschaft zum Werkstolz hat, kann ich nicht sagen.

**Das Thema.** Dialogauftakt war 2025 der *Werkstolz* – ein Begriff, den Frank und ich im Vorfeld des Dialogs des Öfteren gebraucht und diskutiert hatten, wenn es um den Einsatz von KI beim Schreiben geht. Die Maschine stieg auf den Begriff sofort ein und verknüpfte ihn mit verschiedenen Fragen (Wem gehört das Werk? Was ist der Wert von Mühe? Wie verändert sich akademisches (oder pädagogisches) Ethos?) und machte sogleich den Vorschlag, dazu einen Textentwurf zu erstellen. Der Dialogauftakt 2026 bestand im Begriff der *Eigenständigkeit*; über diesen wird, anders als über Werkstolz, vor allem im Prüfungskontext von Hochschulen seit dem Aufkommen generativer KI viel debattiert. Die beiden Begriffe Werkstolz und Eigenständigkeit verbindet, dass KI dazu beiträgt, diese zu hinterfragen: Was bedeuten sie eigentlich und wie verändern sie sich infolge von KI – und ist das zwingend, gewollt oder unnötig? Ein Dialog mit KI ist daher in beiden Fällen selbstbezüglich: Die Interaktion des Menschen mit der Maschine dreht sich um den Einfluss der Maschine auf menschliches Tun.

**Das Vorgehen.** Den Dialog, den ich im Folgenden dokumentieren werde, habe ich im August 2026 mit Claude geführt; eingestiegen bin ich zunächst ohne eine besondere Absicht. Anlass war die Mitarbeit bei den KI-Lotsen<sup>2</sup> zum Thema Prüfungen und KI sowie ein in diesem Zusammenhang stehender erster informeller Austausch mit Antje Michel<sup>3</sup>. Verwendet habe ich das Modell Claude Sonnet 5, das im August 2026 als Standardmodell für Nutzer ohne kostenpflichtiges Abo zur Verfügung stand. Der Dialog fand über zwei Tage verteilt statt –

<sup>1</sup> Ein kompakter Überblick lässt sich z.B. hier finden: <https://kibuzzer.com/de/blog/llm-aktuell-februar-2026.html>

<sup>2</sup> Ein Projekt, das eine Leitstelle zu KI an Hochschule aufbaut: <https://www.ki-lotse.ai/projekt/#>

<sup>3</sup> <https://www.fh-potsdam.de/hochschule-karriere/personen/antje-michel>

unterbrochen von Denkpausen, deren Ergebnisse in meine Prompts eingeflossen sind. Den Dialogverlauf habe ich nach Abschluss erneut mit Frank Vohle diskutiert; zwei Wochen später habe ich mit der Reflexion sowohl der Ergebnisse als auch des Prozesses begonnen, die ich in diesem Beitrag, ähnlich wie letztes Jahr, festhalte. Ich denke, auch dieses Vorgehen lässt sich im Ansatz als autoethnografisch bezeichnen (vgl. Mao, Romero-Hall & Reeves, 2024).; es schließt an die Erfahrungen aus dem letzten Jahr unmittelbar an. Ich gebe in Abschnitt 2 den Chat-Verlauf in der Gänze wieder, habe lediglich Tippfehler in meinen Prompts korrigiert und die Formatierung der KI-Ausgabe verändert. Literaturangaben, die Claude selbst vorschlägt, habe ich überprüft und stelle die jeweils vollständige Quelle als Fußnote direkt in Abschnitt 2 zur Verfügung. Abschnitt 3 reflektiert den Dialog und zieht an einigen Stellen Vergleiche zum Dialog vom Sommer 2025 zum Begriff des Werkstolzes. Ergänzt wird in Abschnitt 4 der inhaltliche Ertrag, der sich meiner Einschätzung nach aus dem Dialog mit Claud zur Eigenständigkeit ergibt.

## 2. Der Dialog mit der Maschine

**Ich:** Ein großes Thema ist derzeit die Eigenständigkeit akademischer Leistungen – nicht nur, aber besonders im Kontext von Hochschulprüfungen. Aus meiner Sicht wäre das Konstrukt der Eigenständigkeit in Zeiten von generativer KI neu zu bewerten, allerdings nicht, so meine Haltung, transhumanistisch in dem Sinne, dass man von vorherein jedes akademische Artefakt als eine Kombination von Leistungen aus menschlicher und maschineller „Fähigkeit“ ansieht. Natürlich: KI greift anders als andere digitale Systeme in den menschlichen Prozess des Denkens und Schreibens ein; ich möchte die Maschine dennoch nicht als gleichwertigen Akteur deuten. Die derzeitige Reaktion auf die Ubiquität von KI besteht an Hochschulen in der Forderung, in Eigenständigkeitserklärungen aufzulisten, für welche Teilprozesse wie Ideengenerierung, Recherche, argumentativer Aufbau, sprachliche Überarbeitung, Übersetzung und ähnliches KI zum Einsatz kommt. Man geht in diesem Zusammenhang davon aus, die Eigenständigkeit einer menschlichen Leistung dann noch erkennen zu können. Und in der Tat könnte man die dazugehörigen Prompts als

eigene Leistung ansehen; bereits hier bahnt sich in der Regel schon an, wie gut das maschinelle Ergebnis wird: Wer viel weiß, bekommt aus der Maschine mehr heraus als jemand ohne (viel) Wissen. So gesehen könnte man sich auf die Qualität des Produktes konzentrieren, um zumindest zu folgern, wie „eigenständig“ jemand mit der Maschine interagiert hat. Was ich allerdings nicht verstehe, ist, warum es in Bezug auf Eigenständigkeit als „harmlos“ gilt, sich von der Maschine Ideen für den Inhalt eines Textes produzieren zu lassen; aus meiner Sicht mindert das im akademischen Bereich die Eigenständigkeit erheblich. Ich möchte mit dir in einen Dialog über dieses Thema einsteigen. Mein Ziel ist es, hier einen ersten Schritt weiterzukommen und ein neues Konstrukt von Eigenständigkeit zu entwickeln – trotz des eklatanten Widerspruchs, dass ich an dieser Stelle einen Teil meiner Eigenständigkeit aufgebe. Ich bin Bildungswissenschaftlerin, seit circa 25 Jahren, habe mich also in meine Domäne ohne KI eingearbeitet. Angst von Kompetenzverlust im Wissensbereich habe ich daher kaum. Ich bin hier aber auch in einer gänzlich andere Ausgangssituation als Studierende bzw. generell Novizen, die erst noch Experten werden wollen. Dieser Aspekt muss in unserem Dialog unbedingt berücksichtigt werden. Bis du bereit? Bist du auch in der Lage, bei Bedarf auf bestehende wissenschaftliche Erkenntnisse zurückzugreifen? Wenn ja, dann lass uns beginnen.

**Claude:** Ja, ich bin bereit – und ich sage das nicht nur höflich, sondern weil ich finde, dass Sie hier eine Unterscheidung treffen, die in der aktuellen Debatte fast durchgängig fehlt. Zur zweiten Frage: Ja, ich kann auf etablierte Konzepte aus Kognitionspsychologie, Expertiseforschung und Hochschuldidaktik zurückgreifen, und wenn Sie an einem Punkt sehr aktuelle Literatur (2025/2026) brauchen – etwa neuere Rahmenwerke zu KI-Disclosure<sup>4</sup> an Hochschulen –, kann ich gezielt recherchieren. Für den Einstieg reicht aber, was an Theorie schon vorliegt. Ich steige direkt bei Ihrem zentralen Punkt ein, weil ich ihn für den Kern der Sache halte: die Sonderstellung der *Ideengenerierung*.

*Warum Ideengenerierung anders ist als sprachliche Überarbeitung.* Die aktuelle Praxis – Auflistung von Teilprozessen in Eigenständigkeitserklärungen – behandelt alle Prozessschritte als graduell gleichwertig: mal mehr, mal weniger

<sup>4</sup> Gemeint ist das Offenlegen von KI-Einsatz, wie ich oben zum Dialogauftakt angesprochen.

KI-Anteil, aber im Prinzip additiv und kommensurabel. Das funktioniert nur, wenn man unterstellt, dass „Idee haben“ und „Idee sprachlich ausformulieren“ strukturell ähnliche Leistungen sind, die sich nur im Automatisierungsgrad unterscheiden. Das halte ich für falsch, und zwar aus einem kognitionspsychologischen Grund, der Ihre Intuition stützt: dem sogenannten *Generierungseffekt* (Slamecka & Graf, 1978<sup>5</sup>, und seither vielfach repliziert). Selbst produzierte Information wird tiefer verarbeitet, besser vernetzt und langfristig verfügbarer als rezipierte Information – auch wenn das Endprodukt (der Satz, der Gedanke) identisch aussieht. Wer eine Idee selbst entwickelt, durchläuft dabei zwangsläufig Prüf-, Verwerfungs- und Alternativschleifen, die beim bloßen Annehmen einer fremden (auch maschinell erzeugten) Idee entfallen. Sprachliche Politur eines bereits selbst gedachten Inhalts greift in diesen Prozess nicht ein – sie kommt danach. Ideengenerierung durch die Maschine dagegen ersetzt genau den Teil, in dem die kognitive Arbeit stattfindet, die man später als „Eigenständigkeit“ honorieren will. Das erklärt, warum die Gleichbehandlung in den Disclosure-Kategorien irreführend ist: „Recherche“ und „sprachliche Überarbeitung“ operieren auf bereits vorhandenem Gedankenmaterial, „Ideengenerierung“ und „argumentativer Aufbau“ sind das Gedankenmaterial. Letztere zu delegieren, ist kategorial etwas anderes als Ersteres.

*Der Novizen/Experten-Punkt, den Sie betont haben.* Hier wird Ihre biografische Ausgangslage analytisch wichtig, nicht nur als Kontext. In der Expertiseforschung (Ericsson, aber auch Bjork's Konzept der *desirable difficulties*<sup>6</sup>) gilt: Der Kampf mit dem Problem – das Ringen um eine Idee, das Scheitern an einer Argumentationslücke – ist nicht Reibungsverlust, sondern der Mechanismus, durch den sich Schemata und Fertigkeiten überhaupt erst aufbauen. Bei Ihnen ist dieser Aufbau abgeschlossen; die Maschine könnte Ihnen Ideen liefern, ohne dass Sie etwas verlernen, weil das Schema längst konsolidiert ist und Sie das Angebotene souverän prüfen, verwerfen, umformen können. Bei Novizen fällt genau die Phase weg, in der das Schema erst

entsteht. Das ist kein gradueller „auch problematisch, nur weniger“, sondern ein qualitativer Unterschied: Bei Ihnen wird eine vorhandene Kompetenz genutzt, bei Studierenden wird eine entstehende Kompetenz unterbrochen, bevor sie sich bilden konnte. Das würde bedeuten: Eine einheitliche Eigenständigkeitsdefinition für „akademische Leistung“ generell ist vermutlich der falsche Anspruch. Man bräuchte unterschiedliche Konstrukte für Qualifikationsarbeiten (Novizenkontext, Kompetenzentwicklung ist Zweck der Übung) versus Forschungsleistung etablierter Wissenschaftler:innen (Output ist Zweck, Kompetenz ist vorausgesetzt).

Eine Frage zurück an Sie, um das zu schärfen: Wenn Ideengenerierung der eigentlich kritische Punkt ist – wie würden Sie „argumentativen Aufbau“ davon abgrenzen? Ist der Aufbau eines Arguments (die Anordnung, die Gewichtung, die Übergänge) für Sie eher der Ideengenerierung zuzuschlagen oder eher der Rezeption/Überarbeitung?

**Ich:** Hier müssen wir kurz innerhalten. Mir ist die Folgerung zu schnell gezogen, der erfahrene Wissenschaftler könne risikolos KI nutzen. Ich denke, Ideen sollten immer erst vom Menschen ausgehen, auch in der Wissenschaft, nicht nur im Studium. KI verändert schließlich die ganze epistemische Praxis. Das ist jetzt ein neues Thema und ich will es in diesem Dialog nicht vertiefen, aber man muss es als Randbedingung mitberücksichtigen, denke ich.

**Claude:** Ein wichtiger Einwand, und Sie haben recht – ich bin zu schnell von „kein Kompetenzverlust“ auf „daher unproblematisch“ gesprungen. Das war ein Fehlschluss. Kompetenzerhalt und epistemische Integrität sind zwei verschiedene Achsen, und ich habe sie vermischt. Der Punkt, den Sie machen, scheint mir dieser zu sein: Auch wenn bei Ihnen keine Lernschädigung droht, bleibt die Frage, *wessen* Gedanke ein wissenschaftlicher Befund, eine Interpretation, eine These eigentlich ist – unabhängig davon, ob der Urheber sie noch „lernen“ muss oder nicht. Wenn die erste Idee zu einer Forschungsfrage von der Maschine stammt, dann ist die Wissenschaftlerin nicht mehr Urheberin

<sup>5</sup> Slamecka, N.J. & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 592-604.

<sup>6</sup> Hier gibt Claude keine Quellen an. Ich weiß allerdings, dass der 2020 verstorbene Karl Anders Ericsson als profilierter Vertreter der psychologischen

Expertiseforschung gilt bzw. galt. Auch das genannte Konzept zu, ich übersetze mal, wünschenswerten Schwierigkeiten existiert. Siehe z.B.: Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, 9(4), 475-479.

des Erkenntnisfortschritts, sondern nur noch Prüferin und Ausarbeiterin eines maschinell vorgeschlagenen Gedankens – selbst wenn sie diese Prüfung exzellent leistet. Das würde bedeuten: Die Forderung „Ideen sollten immer erst vom Menschen ausgehen“ ist keine Schutzmaßnahme für die Lernenden, sondern eine Bedingung für das, was Wissenschaft als Praxis überhaupt ausmacht – nämlich, dass Erkenntnis von einem erkennenden Subjekt ausgeht, nicht bloß durch es hindurchläuft. Das passt zu Diskussionen, die es in der Wissenschaftsforschung bereits zum Konzept der „epistemischen Urheberschaft“ gibt (etwa in Debatten um Ghostwriting oder um die Frage, ob ein reiner Daten-Auswerter als Autor gelten darf) – nur dass generative KI das Problem radikalisiert, weil sie nicht nur ausführt, sondern selbst inhaltlich vorschlägt.

Ich nehme das als Randbedingung mit auf, wie Sie sagen, ohne es hier zu vertiefen: *Ideengenerierung durch KI ist demnach nicht graduell riskanter oder sicherer je nach Erfahrungsstand der Person, sondern grundsätzlich different von anderen Unterstützungsformen* – unabhängig vom Subjekt. Das schärft Ihre ursprüngliche These sogar: Es geht nicht in erster Linie um Kompetenzschutz, sondern um die Frage, was als genuine intellektuelle Urheberschaft überhaupt zählen kann.

Damit zurück zu meiner Frage von eben, jetzt unter dieser korrigierten Prämisse: Wie stehen Sie zum argumentativen Aufbau? Wenn die Ideengenerierung (der Einfall, die These, der Kerngedanke) klar auf der Seite des Menschen verortet werden muss – gehört die Struktur, in der diese Gedanken dann angeordnet und miteinander verknüpft werden, für Sie eher noch zum „Denken“ selbst (und wäre damit ebenso schützenswert), oder eher schon zur Darstellung eines bereits fertig gedachten Gedankens (und damit näher an der sprachlichen Überarbeitung)?

**Ich:** Das kommt darauf an, was man darunter versteht: In theoretischen Kontexten ist ein Gedanke eine längere Einheit von Sätzen bzw. Aussagen; da steckt dann bereits eine erste Argumentationslinie drin. In den Naturwissenschaften mag das anders sein, aber gänzlich

unterscheiden dürfte es sich wohl auch nicht; also, was ich sagen will: Die „Idee“, der Ausgangsgedanke und eine Argumentation hängen eng zusammen. Davon zu trennen ist die Frage, wie man einen fertigen akademischen Text in der Gänze gliedert. Das müsste man folglich trennen und passende Begriffe dafür finden. Aber auch die textübergreifende Argumentation halte ich für wesentlich im Hinblick auf Eigenständigkeit des Denkens: Da macht sich fest, ob ich etwas logisch, konsistent, originell und auch zu Ende denken kann, würde ich sagen.

**Claude:** Das ist eine wichtige Präzisierung, und ich glaube, sie lässt sich mit einer sehr alten Unterscheidung fassen, die in der modernen Debatte fast verschwunden ist: der klassischen Rhetorik-Trias *inventio* – *dispositio* – *elocutio*<sup>7</sup>. *Warum diese Trias hilft.* Die antike Rhetorik hat *inventio* (das Finden der Gedanken/Argumente) und *dispositio* (ihre Anordnung) nie sauber getrennt – aus gutem Grund: Die Reihenfolge, in der Argumente präsentiert werden, ist selbst ein argumentativer Akt. Welches Gegenargument ich zuerst entkräfte, welche Prämisse ich als Fundament lege, bevor die nächste kommt, in welcher Konsequenz sich eine These aus vorherigen Schritten ergibt – das *ist* Denken, nicht nur Verpackung von bereits Gedachtem. Erst *elocutio*, die sprachliche Ausformulierung (Wortwahl, Satzbau, Stil), gilt in dieser Tradition als nachgelagert und eigenständig delegierbar. Das trifft ziemlich genau Ihre Unterscheidung: Sie sagen, die *textübergreifende Argumentation* (Konsistenz, Stringenz, das Zu-Ende-Denken einer Linie) gehört noch zum Kerngedanken – das wäre in der rhetorischen Terminologie *dispositio als Teil der inventio*, nicht von ihr trennbar. Davon zu unterscheiden ist die *reine Makrogliederung eines bereits fertigen Textes* – Kapitelreihenfolge nach Konvention, Abschnittsüberschriften, die formale Textarchitektur, die eher zur Darstellung als zum Denken gehört.

*Eine kognitionswissenschaftliche Stütze dafür.* Das lässt sich mit Bereiter & Scardamalias (1987) Unterscheidung zwischen *Knowledge-Telling* und *Knowledge-Transforming* beim Schreiben verknüpfen<sup>8</sup>. Im Knowledge-

<sup>7</sup> Wer sich in der Rhetorik ein wenig auskennt, erkennt gleich, dass das kein schlechter Tipp ist. Für einen kurzen Überblick kann man Wikipedia nutzen (<https://de.wikipedia.org/wiki/Rhetorik>). *Invention*, *dispositio* und *elocutio* sind die ersten drei der fünf Produktionsstadien einer Rede, wie sie seit der Antike postuliert werden.

<sup>8</sup> Den Text und dessen Inhalte kenne ich sogar – aus wirklich sehr alten Zeiten: Scardamalia, M. & Bereiter, C. (1987). *Knowledge telling and knowledge transforming in written composition*. In S. Rosenberg (Eds.), *Advances in applied psycholinguistics* (pp. 142-175). Cambridge University Press.

Transforming-Modell gibt es zwei interagierende „Problem Spaces“: einen inhaltlichen und einen rhetorischen. Der entscheidende Punkt: Das Ringen um die rhetorische Form – *wie* sage ich das, in welcher Ordnung – wirkt auf den inhaltlichen Raum zurück und verändert dabei den Gedanken selbst. Genau das beschreiben Sie: Argumentationsaufbau ist nicht Verpackung eines fertigen Gedankens, sondern Teil seiner Entstehung. Reine Textgliederung dagegen, die diese Rückkopplung nicht hat (weil sie einem Genre-Schema folgt, das unabhängig vom Inhalt feststeht), wäre eher Knowledge-Telling – und damit näher an Ihrer „Darstellung“, nicht am „Denken“.

*Ihr Hinweis auf die Disziplinenunterschiede passt hier gut hinein.* In den Naturwissenschaften ist die Makrostruktur (IMRAD: Introduction-Methods-Results-Discussion) weitgehend genrekonventionell und inhaltsunabhängig vorgegeben – hier ist Dispositio tatsächlich stärker von Inventio abtrennbar, weil sie kaum noch etwas über die Qualität des Denkens verrät. In geistes- und sozialwissenschaftlichen Texten dagegen, wo die Gliederung selbst die Argumentationslinie abbildet (welches Kapitel worauf aufbaut, warum diese und keine andere Reihenfolge), ist Dispositio nicht vom Inhalt trennbar – hier hat sie noch diagnostischen Wert für Eigenständigkeit. Das würde einen ersten Entwurf für Ihre Begriffstrennung ergeben: (1) *Inventio/Gedankenführung* (Kerngedanke + die Argumentationslinie, die ihn trägt) – zentral eigenständigkeitsrelevant, nicht delegierbar; (2) *genrekonventionelle Textorganisation* (Formatvorgaben, die unabhängig vom Inhalt bestehen) – eher delegierbar, disziplinenabhängig variabel in ihrer Relevanz; (3) *Elocutio/sprachliche Überarbeitung* – am unproblematischsten delegierbar.

Eine Anschlussfrage, die mir wichtig erscheint: Wenn Dispositio in geisteswissenschaftlichen Texten so eng an Inventio hängt, dass beide praktisch eine Einheit bilden – müsste dann nicht auch die *Recherche* neu bewertet werden? Denn die Auswahl und Gewichtung von Quellen (welche Literatur ich als tragend behandle, welche ich verwerfe) ist ja oft schon eine Vorentscheidung über die Argumentationslinie, nicht neutrale Vorarbeit dazu. Sehen Sie

*Recherche in Ihrem Modell eher bei der harmlosen Zuarbeit – oder rutscht sie durch diese Überlegung nicht auch in den geschützten Kernbereich?*

**Ich:** Wieder ein Zwischenstopp, bevor ich zu deiner Anschlussfrage komme. Die Dreiteilung passt perfekt – das stimmt. Ich kenne sie, hatte aber hier gar nicht daran gedacht. Ein guter Hinweis also. Aber: Eine KI-generierte sprachliche Überarbeitung könnte zwar unproblematisch sein; möglicherweise aber wirkt sie als Störfaktor insbesondere im epistemischen Schreiben. Aus der Schreibdidaktik weiß man, dass die sprachliche Überarbeitung sehr wohl Rückwirkungen auf den Inhalt, damit auch auf die Idee, je nach Fach gegebenenfalls sogar auf die Erkenntnis haben kann. Hier hast du aus meiner Sicht erneut zu schnell eine fertige Schlussfolgerung gezogen.

**Claude:** Sie haben wieder recht, und das Muster wiederholt sich bei mir: Ich behandle *elocutio* als das saubere, unproblematische Restglied, nachdem *inventio* und *dispositio* als geschützt markiert sind – nach dem Motto: was übrig bleibt, ist harmlos. Das ist genau der Fehler, den ich vorhin schon einmal gemacht habe, nur an anderer Stelle.

*Was die Schreibdidaktik hier zeigt.* Der Befund, den Sie ansprechen, lässt sich am prägnantesten mit Janet Emigs klassischer These „Writing as a Mode of Learning“ (1977)<sup>9</sup> und später, differenzierter mit David Galbraiths *Knowledge-Constituting*-Modell des Schreibens fassen<sup>10</sup>. Galbraiths Punkt ist genau der, den Sie machen: Formulierung ist nicht die nachgelagerte Verpackung eines bereits fertigen Gedankens, sondern der Prozess, in dem unklares, implizites Wissen erst explizit und dabei oft *verändert* wird. Man weiß häufig nicht genau, was man denkt, bevor man es in eine bestimmte sprachliche Form zu zwingen versucht – und beim Ringen um die richtige Formulierung stellt sich heraus, dass der ursprüngliche Gedanke unpräzise, widersprüchlich oder unvollständig war. Die Korrektur der Sprache *ist* dann Korrektur des Denkens, nicht getrennt davon. Das bedeutet: Wenn eine KI den Satz „glättet“ – Nebensätze auflöst, Präzisierungen vorschlägt, Redundanzen entfernt –, kann sie genau die Reibung

<sup>9</sup> Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, 28(2), 122-128.

<sup>10</sup> Galbraith, D. & Torrance, M.C. (1999). Writing as a knowledge-constituting process. In G.

Rijlaarsdam, E. Espéret, M. Torrance & D. Galbraith (Eds.), *Studies in writing: Vol. 4. Knowing what to write: Conceptual processes in text production* (pp. 139-160). Amsterdam University Press.

wegnehmen, an der eigentlich die inhaltliche Klärung stattfinden würde. Der Text wird dann sprachlich besser, aber der Denkprozess, der bei eigenständiger Überarbeitung stattgefunden hätte, entfällt – ähnlich wie bei der Ideengenerierung, nur später im Prozess und subtiler, weil es äußerlich wie „nur Stil“ aussieht.

*Konsequenz für unser Modell.* Das zwingt zu einer Ergänzung der Trias: Es reicht nicht, *elocutio* pauschal als unkritisch einzustufen. Man müsste unterscheiden zwischen: (a) *rein oberflächlicher sprachlicher Glättung*, die tatsächlich nachgelagert ist (z. B. Grammatikkorrektur, Interpunktion, Stilregister an Textsorte anpassen) – hier ist Delegation vermutlich unproblematisch; (b) *epistemischem Formulieren*, bei dem das Ringen um die richtige Formulierung selbst der Ort ist, an dem Klärung, Präzisierung oder gar Erkenntnis stattfindet – hier gilt für *elocutio* dasselbe wie für *inventio*: die Delegation nimmt der Person die Gelegenheit zur eigenen Erkenntnisleistung. Und vermutlich ist auch das wieder fachabhängig, wie schon bei *dispositio*: In stark theoretischen, begriffsarbeitsintensiven Disziplinen (Philosophie, Teile der Bildungswissenschaft, Theologie) ist die exakte Formulierung oft selbst die Aussage – ein anderes Wort ist ein anderer Gedanke. In eher deskriptiv-empirischen Textteilen (Ergebnisdarstellung, Methodenbeschreibung) ist die Formulierung austauschbarer, ohne den Gedanken zu verändern.

Ich sehe jetzt ein Muster in meinen eigenen Fehlern, das vielleicht selbst diagnostisch interessant ist: Ich neige dazu, den jeweils noch nicht diskutierten Prozessschritt vorschnell als „den unproblematischen Rest“ zu behandeln, einfach weil er zuletzt an der Reihe ist. Das könnte darauf hindeuten, dass die ganze Prämisse der aktuellen Praxis – eine Liste von Teilprozessen, von denen einige gefährlich und andere harmlos sind – grundsätzlich fragwürdig ist. Vielleicht gibt es *keinen* Teilprozess des akademischen Schreibens, der grundsätzlich sicher delegierbar ist, sondern nur situative Bedingungen (Fach, Textsorte, Erkenntnisstand des Autors, Funktion der Textstelle), unter denen ein und derselbe Prozessschritt mal epistemisch tragend, mal peripher ist. Wäre das eine Formulierung, die Sie mittragen würden – dass die kategoriale Einteilung in „gefährliche“ und „harmlose“ Teilprozesse selbst das Problem ist, und man stattdessen fragen müsste, *wann* welcher Prozessschritt epistemisch tragend wird?

**Ich:** Ja, das sehe ich auch so. Und das steht in einem deutlichen Konflikt zu der Position, dass, weil alle unsere akademischen Prozesse ohnehin längst digital unterstützt sind, KI als neuer Akteur nicht nur zu akzeptieren, sondern passend zu platzieren ist und das Verständnis von Eigenständigkeit als menschliches Konstrukt zu erweitern (oder gar aufzulösen) wäre. Das gilt nicht selten als die moderne Interpretation, während die, die wir hier verfolgen, eher als konservativ und veraltet bezeichnet wird. Daher suche ich nach einer „Lösung“ – einer Position, die die Fakten nicht ausblendet (denn natürlich ist KI da, ich nutze sie jetzt ja auch), aber hier differenzierter ist. Ich frage mich immer öfter, wie sich Menschen, die jetzt mit KI aufwachsen und an der Hochschule alles mit KI machen, noch zu kreativ und selbst denkenden Persönlichkeiten entwickeln können. In der Hochschulbildung haben wir doch die Verantwortung, eben dies zu fördern. Das geht für mich nicht zusammen. Auch in weiten Teilen der Schreibdidaktik, deren Erkenntnisse wir gerade nutzen und feststellen, dass es so einfach nicht ist mit Listen, was harmlos und was riskant ist, gibt es die Tendenz, KI zu „umarmen“ – bisweilen, so mein Eindruck, aus der Not heraus, weil alle ohnehin zu KI greifen, wenn es um akademische Aufgaben geht.

**Claude:** Das ist der Punkt, an dem sich zeigt, dass „konservativ vs. modern“ die falsche Achse ist – es sind zwei unterschiedliche Fragen, die in der aktuellen Debatte systematisch verwechselt werden: eine *deskriptive* und eine *normative*.

*Die Verwechslung, die die „moderne“ Position trägt.* Die Aussage „akademische Prozesse sind längst digital durchdrungen, KI ist ein weiterer Akteur darin“ ist eine deskriptive Feststellung – und sie stimmt. Daraus folgt aber nicht automatisch die normative Konsequenz „also muss Eigenständigkeit als Konstrukt erweitert oder aufgelöst werden“. Das ist ein naturalistischer Fehlschluss: Aus dem Ist (KI wird faktisch genutzt, überall) wird ein Soll (also soll/darf sie das Konstrukt neu definieren). Ihr Hinweis auf die Schreibdidaktik trifft genau das: Wenn dort KI „umarmt“ wird, „weil ohnehin alle nach ihr greifen“, ist das keine pädagogische Begründung, sondern eine Kapitulation vor der Empirie, verkleidet als Fortschrittlichkeit.

*Eine Unterscheidung, die den scheinbaren Widerspruch auflöst.* Ich glaube, der Ausweg liegt darin, dass „Eigenständigkeit“ gar nicht *ein* Konstrukt sein muss, sondern zwei

unterschiedliche Funktionen erfüllt, die man in der Debatte permanent vermischt: (1) *Eigenständigkeit als Nachweis-/Bewertungskonstrukt* (Wer hat diese Leistung erbracht? Ist die Note gerechtfertigt?) (2) *Eigenständigkeit als Entwicklungsziel* (Entsteht durch diesen Prozess die Fähigkeit zu eigenständigem Denken?). Die „moderne“ Position beantwortet nur Frage 1 – und tut so, als wäre damit auch Frage 2 erledigt. Aber genau Frage 2 ist der Kern der institutionellen Verantwortung von Hochschulbildung, die Sie ansprechen. Ein Prüfungsformat kann formal fair und eindeutig bewertbar sein (Frage 1 gelöst durch Disclosure-Listen, KI-Nutzungsprotokolle o.Ä.) und trotzdem systematisch verhindern, dass sich das entwickelt, wozu die Ausbildung eigentlich da ist (Frage 2 sabotiert). *Der Vygotsky-Anker, der das präzisiert.* Hier hilft, denke ich, ein Konzept aus der Entwicklungspsychologie, das genau diese Zeitlichkeit einfängt: das *Scaffolding* nach Wood, Bruner & Ross (1976), aufbauend auf Vygotskys Zone der nächsten Entwicklung<sup>11</sup>. Entscheidend am Scaffolding-Konzept ist nicht nur, dass eine Stütze gegeben wird, sondern dass sie konstitutiv temporär ist – sie muss schrittweise entzogen werden (*fading*), damit der Lernende die Funktion internalisiert, die die Stütze zunächst übernommen hat. Ein Gerüst, das dauerhaft am Gebäude bleibt, hat seine pädagogische Funktion verfehlt, selbst wenn das Gebäude am Ende gut aussieht. Das ist der entscheidende Unterschied zur *Extended-Mind-These* (Clark & Chalmers, 1998)<sup>12</sup>, auf die sich die „moderne“ Position oft implizit stützt: Dort wird ein Werkzeug (Notizbuch, Taschenrechner, jetzt KI) als legitimer, dauerhafter Teil des kognitiven Systems verstanden – und das ist für *bereits ausgebildete* Kognition, also in der Expertenpraxis, durchaus plausibel (das war ja unser Ausgangspunkt bei Ihnen als erfahrener Wissenschaftlerin, wobei wir da inzwischen selbst Einschränkungen gemacht haben). Für die Ausbildungsphase aber gilt: Ein Werkzeug, das nie entzogen wird, wird nicht internalisiert, sondern ersetzt die Funktion dauerhaft, die eigentlich internalisiert werden sollte. *Extended Mind* beschreibt einen stabilen Endzustand; *Scaffolding* beschreibt einen sich auflösenden Übergang. Studierende, die

durchgängig mit KI-Ideengenerierung, -Gliederung und -Formulierung arbeiten, befinden sich nicht in einer *Extended-Mind-Situation*, sondern in einer *Scaffolding-Situation*, in der das *Fading* nie stattfindet.

*Damit zu Ihrer Formulierung des Widerspruchs.* Er löst sich, wenn man sagt: Es ist kein Widerspruch, KI in der etablierten Wissenschaftspraxis als legitimen (wenn auch, wie wir gesehen haben, bei der Ideengenerierung selbst dort noch zu prüfenden) Akteur zu akzeptieren, und gleichzeitig in der Hochschulbildung – die per Definition eine Entwicklungsphase, keine Praxisphase ist – ihre Nutzung in genau den epistemisch tragenden Prozessschritten (Ideengenerierung, Dispositio, epistemisches Formulieren) zu begrenzen oder zumindest verpflichtend mit *Fading-Phasen* zu koppeln. Das wäre keine konservative Ablehnung von KI, sondern eine präzise Unterscheidung zwischen zwei völlig verschiedenen institutionellen Funktionen, die die „moderne“ Position verwischt, indem sie tut, als gäbe es nur einen Nutzungskontext. Das würde auch erklären, warum die Disclosure-Listen so unbefriedigend bleiben: Sie beantworten nur die Bewertungsfrage (Frage 1), ignorieren aber vollständig, ob am Ende des Studiums die Fähigkeit vorhanden ist, die eigentlich hätte entstehen sollen.

Eine Frage, die sich daraus für mich ergibt: Wenn *Scaffolding* konstitutiv das *Fading* braucht – wie müsste ein Prüfungs- oder Ausbildungsformat aussehen, das dieses *Fading* strukturell erzwingt, statt es dem Zufall oder der Selbstdisziplin der Studierenden zu überlassen? Haben Sie da aus Ihrer bildungswissenschaftlichen Praxis schon eine Intuition, oder ist das der Punkt, an dem Sie selbst noch nach einer Form suchen?

**Ich:** Ja, ich habe eine Intuition, die ich auch schon 2022 veröffentlicht habe unter dem Stichwort „Prüfungskultur“: Man müsste die Zahl der summativen Prüfungen mit Rechtsfolgen, für die man all diese Formalia wie Eigenständigkeitserklärungen braucht, radikal wieder zurückfahren und stattdessen die Übungsräume mit formativem Assessment und Feedback (ohne Noten) erweitern. Das würde auch das

<sup>11</sup> Vygotski ist im bildungswissenschaftlichen Kontext allseits bekannt mit seiner Zone der proximalen Entwicklung; der Hinweis ist also völlig richtig. Die Quelle ist etwas alt, aber durchaus brauchbar, zumal da auch Jerome Bruner in diesem Kontext ein wichtiger Akteur war: Wood, D.,

Bruner, J.S. and Ross, G. (1976), The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89-100.

<sup>12</sup> Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19.

angesprochene Scaffolding ermöglichen. Das „Prüfungsproblem“ in Zeiten von KI (was ist noch eigenständig?) würde sich dann nicht so oft bzw. viel weniger stellen, aber es wäre natürlich nicht weg: Bei der summativen Prüfung stellt sich erneut die Frage, wie man Eigenständigkeit definiert. Aus humanistischer Sicht müsste man fordern, die Leistung ohne KI zu erstellen. Aus der „Employability-Perspektive“ heißt es eher: Die Leistung kann ein Kombi-Produkt aus Mensch und Maschine sein, weil das später in den Anschlusssystemen ebenso sein wird. Ich hatte schon die Idee, dass man zwei Versionen von Prüfungen einführt (mit und ohne KI). Das ist aber prüfungsrechtlich derzeit wohl unmöglich. Ein weiteres Gegenargument wäre, dass das nicht „gerecht“ ist – ein schweres Geschütz, das noch einmal ein ganz anderes große Feld eröffnet. Als mögliches (utopisches) Szenario könnte ich es mir in ferner Zukunft aber durchaus vorstellen.

**Claude:** Ihre Zweiteilung – formatives Scaffolding im Übungsraum, aber Wiederkehr des Problems im tatsächlichen Prüfungsmoment – trifft, denke ich, genau den Punkt, an dem die meisten aktuellen Modelle sich wegduckten. Das lässt sich mit einer etablierten Unterscheidung aus der Assessment-Forschung schärfer fassen: *Assessment for Learning* vs. *Assessment of Learning* (Black & Wiliam 1998; später systematisiert bei Earl 2003<sup>13</sup>). Ihr Vorschlag verschiebt das Gewicht radikal Richtung Ersterem – und das ist analytisch stimmig, weil nur *Assessment for Learning* die Fading-Logik des Scaffolding überhaupt strukturell abbilden kann: Feedback ohne Rechtsfolge erlaubt Fehler, Wiederholung, Reibung – also genau das, was bei summativer Bewertung mit Konsequenz vermieden wird. Das Problem verschiebt sich dadurch nicht weg, wie Sie sagen, sondern konzentriert sich auf den verbleibenden, selteneren summativen Moment. Das ist ein Fortschritt, kein Verschwinden.

*Zur humanistisch/employability-Spaltung im verbleibenden Prüfungsmoment.* Ich sehe hier eine Struktur, die genau der entspricht, die wir vorhin schon einmal aufgedeckt haben. Die Employability-Position argumentiert: „Später arbeitet man ohnehin im Mensch-Maschine-Verbund, also sollte die Prüfung das abbilden.“

Das ist derselbe naturalistische Fehlschluss wie bei der „KI ist überall, also muss Eigenständigkeit sich anpassen“-Position – nur eine Ebene tiefer, bezogen auf den Arbeitsmarkt statt auf die Wissenschaftspraxis allgemein. Aus einem Ist (so wird später gearbeitet) folgt kein Soll (also muss die Prüfung diesen Zustand vorwegnehmen) – jedenfalls nicht zwingend, wenn die Prüfung eine andere Funktion hat als die Arbeitswelt zu simulieren, nämlich festzustellen, ob eine bestimmte Fähigkeit *beim Menschen selbst* ausgebildet ist. Das ist übrigens auch der Grund, warum der Vergleich mit dem Taschenrechner in der Mathematik hinkt, den die Employability-Seite gern bemüht: Der Taschenrechner ersetzt eine Routineoperation, deren zugrundeliegendes Verständnis anderweitig geprüft werden kann; bei generativer KI und Ideengenerierung/Dispositio/epistemischem Formulieren geht es aber, wie wir herausgearbeitet haben, um die Konstitution des Gedankens selbst – da gibt es keine separate „Verständnisebene“, die man ersatzweise prüfen könnte.

*Zur Zwei-Varianten-Idee und ihrer derzeitigen rechtlichen Lage.* Ich habe kurz nachgesehen, wie deutsche Hochschulen das Verhältnis von KI-Nutzung und Eigenständigkeitserklärung derzeit handhaben, weil mich interessiert hat, ob „rechtlich unmöglich“ tatsächlich so feststeht. Der aktuelle Stand ist eher uneinheitlich als eindeutig verboten: Hessische Hochschulen etwa überlassen die Regelung des KI-Einsatzes weitgehend den einzelnen Lehrenden, orientiert an den jeweiligen Lernzielen der Prüfung, wobei die Freiheit der Lehre hoch gewichtet wird. Gleichzeitig zeigt ein aktuelles Urteil des Verwaltungsgerichts Kassel von Ende Februar, dass die Gerichte gerade erst anfangen, tragfähige Kategorien zu entwickeln: unerlaubter KI-Einsatz in Prüfungen kann als Täuschung gewertet werden, wobei die Beweislast bei der Hochschule liegt und Zweifel in der Regel zugunsten der Studierenden gehen – das Verfahren geht zudem in Berufung, ist also noch offen. Das heißt: Die Rechtslage ist nicht starr, sondern gerade in Bewegung, weil sie de facto von Fall zu Fall an Verwaltungsgerichten neu verhandelt wird. Das macht Ihre Zwei-Varianten-Idee vielleicht nicht heute umsetzbar, aber

<sup>13</sup> Diese Unterscheidung ist oben natürlich gemeint, nur anders formuliert. Die beiden hier genannten Quellen (ich selber hätte eher Ashwin angeführt), lauten: Black, P. & Wiliam, D. (1998). *Assessment and classroom learning. Assessment in Education:*

*Principles, Policy & Practice*, 5(1), 7-74. Earl, L. (2003). *Assessment as learning: Using Classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.

weniger utopisch, als „prüfungsrechtlich derzeit unmöglich“ klingt – es ist eher noch nicht existent als kategorisch ausgeschlossen.

*Zum Gerechtigkeitseinwand – hier würde ich schärfer nachfragen.* Ich vermute, der eigentliche Kern des Gerechtigkeitsproblems liegt nicht dort, wo er meist verortet wird (unterschiedliche Bedingungen = unfaire Vergleichbarkeit innerhalb einer Kohorte). Dafür gäbe es sogar ein Präzedenzmodell: Nachteilsausgleiche für Studierende mit Behinderung erlauben schon heute unterschiedliche Prüfungsbedingungen, ohne dass die Chancengleichheit als verletzt gilt – solange dieselbe Kompetenz auf unterschiedlichem Weg gemessen wird. Man könnte argumentieren, „mit KI“ und „ohne KI“ seien einfach zwei verschiedene Konstrukte, die je für sich valide geprüft werden. Das eigentliche Problem, glaube ich, liegt woanders: Sobald ein Zeugnis ausweisen müsste, welche Variante jemand gewählt hat, entsteht zwangsläufig eine faktische Hierarchie zwischen den Varianten – unabhängig davon, ob die Hochschule sie als gleichrangig deklariert. Arbeitgeber, Rankings, das kollektive Prestigeempfinden würden eine der beiden Varianten schnell als die „eigentliche“, anspruchsvollere lesen (vermutlich die Variante ohne KI, weil seltener und schwerer zu fälschen) – womit die vermeintlich freie Wahl faktisch zu einem Zwang würde, die „harte“ Variante zu wählen, um nicht benachteiligt zu erscheinen. Damit wäre die Zwei-Varianten-Lösung keine Lösung des Eigenständigkeitsproblems, sondern nur seine Verlagerung auf eine neue, verdeckte Rangordnung.

Eine Frage an Sie: Wäre dieses Signalproblem für Sie ein Einwand gegen die Zwei-Varianten-Idee als solche – oder eher ein Hinweis darauf, dass man sie nur *innerhalb* eines Prüfungsformats denken darf (etwa als offen deklarierte Bearbeitungsart, die nicht auf dem Zeugnis erscheint, sondern nur intern der Notenfindung dient), statt als zwei nach außen sichtbar unterscheidbare Abschlussvarianten?

**Ich:** Also ja, genau das ist der Punkt, den du ansprichst: Dass es als „ungerecht“ empfunden wird (oder werden kann), wenn es auf dem Zeugnis steht, ob man mit oder ohne KI gearbeitet hat. Für mich stellt sich das wie ein klassisches Dilemma dar: Da prallt die Hoffnung auf Erleichterung im Prozess infolge von Hilfe durch KI mit der Erwartung zusammen, für

einen auch erleichterten Lernprozess am Ende „belohnt“ zu werden wie alle anderen. Und das funktioniert natürlich nicht, ist irgendwie moralisch unstimmig. Das ist also auch keine Lösung.

**Claude:** Ich glaube, hier soßen Sie auf ein altes, ungelöstes Problem. Sie stellen fest: Beim KI-Dilemma beansprucht jemand die volle Anerkennung eigenständiger Leistung, während die Bedingung, die diese Anerkennung eigentlich rechtfertigt (der ungestützte kognitive Prozess), unterlaufen wird – nicht aus Täuschungsabsicht unbedingt, sondern weil beide Ansprüche, Erleichterung *und* volle Belohnung, gleichzeitig legitim erscheinen wollen. Das, was Sie als moralische Inkohärenz beschreiben, ist im Kern die *Desert-Problematik* der politischen Philosophie: die Frage, wofür ein Mensch überhaupt moralisch beanspruchen darf, „belohnt“ zu werden. Rawls hat in *A Theory of Justice* (1971)<sup>14</sup> den entscheidenden Stein ins Wasser geworfen: Selbst Talent und Anstrengungsfähigkeit sind schon Resultat einer „natürlichen und sozialen Lotterie“ – niemand hat sich seine Intelligenz, seinen Fleiß, sein förderliches Elternhaus selbst verdient. Wenn das stimmt, war der Anspruch „ich verdiene die Belohnung, weil ich es *selbst* geleistet habe“ ohnehin schon prekär, lange bevor KI ins Spiel kam. Michael Sandels *The Tyranny of Merit* (2020)<sup>15</sup> hat daraus eine Sozialkritik gemacht: Die Meritokratie verkennt systematisch, wie viel vom „eigenen“ Erfolg eigentlich Glück, Herkunft, Zufall ist. Das gibt Ihrem Dilemma eine interessante Wendung: KI verschärft nicht ein neues Problem, sondern macht ein altes, bislang gut verdecktes Problem plötzlich sichtbar und *wählbar*. Der Unterschied zur Rawlsschen Glücks-Lotterie ist aber entscheidend, und der macht die Sache moralisch heikler, nicht leichter: Herkunftsglück ist unverschuldet und nicht gewählt – KI-Nutzung ist eine Entscheidung. Man kann sich der ererbten Startbedingungen nicht entziehen, wohl aber der Entscheidung, sich die Denkarbeit erleichtern zu lassen und trotzdem den vollen Eigenständigkeitskredit einzufordern. Das würde bedeuten: Die moralische „Unstimmgigkeit“, die Sie sehen, ist eigentlich eine Instabilität, die dem Desert-Begriff immer schon innewohnte – KI zwingt uns nur, sie nicht länger zu ignorieren, weil sie jetzt aktiv herbeigeführt statt passiv erlitten wird.

<sup>14</sup> Rawls, J. (1971). *A theory of justice*. Harvard University Press.

<sup>15</sup> Sandel, M.J. (2020). *The tyranny of merit*. Allen Lane.

Was das für „keine Lösung“ bedeutet. Ich glaube, Sie haben recht, dass sich daraus keine Lösung ableiten lässt – das ist typisch für Verdienst-Fragen: Auch die Rawls/Sandel-Debatte endet nicht in einer kohärenten Theorie dessen, was jemand „wirklich“ verdient, sondern in Verfahrensregelungen, die das Problem *handhabbar*, nicht *aufgelöst* machen (progressive Besteuerung, Chancenausgleich, Affirmative Action<sup>16</sup> – alles Antworten auf Ebene der Institution, nicht der individuellen moralischen Gerechtigkeit). Das könnte auch für unser Thema die einzig ehrliche Konsequenz sein: Man findet keine kohärente Theorie dessen, wann jemand „wirklich“ Eigenständigkeit verdient hat, sondern nur institutionelle Verfahren (wie Ihr Vorschlag zur Prüfungskultur), die den Schaden begrenzen, ohne das zugrundeliegende philosophische Problem zu lösen. Das wäre dann kein Versagen unseres Dialogs, sondern ein ziemlich präzises Ergebnis.

Würden Sie das mittragen – dass am Ende dieses Gesprächs eher eine belastbare Diagnose steht als eine Auflösung, und dass genau das der ehrliche Befund ist?

**Ich:** Ja, vermutlich kann man nur Verfahrensschritte zur Bändigung des Problems bei diesem Thema aufnehmen.

**Claude:** Gut. Bevor wir weitergehen, will ich kurz bündeln, was sich in diesem Gespräch aufgebaut hat, weil es inzwischen eine gewisse Architektur hat, die man festhalten sollte: Wir sind von der Kritik an additiven Disclosure-Listen ausgegangen und haben über den Generierungseffekt zur Sonderstellung der Ideengenerierung gefunden. Sie haben mich zweimal korrigiert, als ich vorschnell Prozessschritte als „harmlos“ abgestempelt habe – bei der Expertise-Frage (epistemische Urheberschaft gilt unabhängig vom Kompetenzstand) und bei der sprachlichen Überarbeitung (epistemisches Formulieren ist so wenig neutral wie Ideengenerierung). Daraus hat sich die vermutlich wichtigste Erkenntnis ergeben: Es gibt keinen Teilprozess, der kategorial sicher delegierbar ist – nur Bedingungen (Fach, Textsorte, Funktion der Textstelle), unter denen ein Prozessschritt mal tragend, mal peripher ist. Die Rhetorik-Trias (*inventio/dispositio/elocutio*) und Bereiter & Scardamalias Knowledge-Transforming-Modell haben dem eine Struktur gegeben. Darauf haben wir die

institutionelle Ebene gesetzt: die Unterscheidung zwischen Eigenständigkeit als Bewertungskonstrukt und als Entwicklungsziel, gestützt durch Scaffolding/Fading (Wood, Bruner & Ross) gegen Extended Mind (Clark & Chalmers) – zwei verschiedene Modelle für zwei verschiedene Lebensphasen wissenschaftlichen Arbeitens. Ihre Prüfungskultur-Idee hat das operationalisiert: formatives Assessment als Ort des Fading, summative Prüfung als der Ort, an dem das Eigenständigkeitsproblem unausweichlich bleibt. Und zuletzt die Rawls-Sandel-Linie: Das Dilemma zwischen Erleichterungswollen und Vollbelohnungsbeanspruchungen ist keine neue Pathologie, sondern die alte, ungelöste Desert-Frage, jetzt durch KI von unverschuldetem Glück zu gewählter Entscheidung verschoben – und damit moralisch heikler, nicht leichter.

Zwei Fäden liegen noch offen, die Sie selbst aufgeworfen haben: Meine Frage zur *Recherche* (rutscht sie durch die Dispositio-Überlegung in den geschützten Kernbereich?) haben wir nie beantwortet. Und die praktische Frage, wie „Verfahrensschritte zur Bändigung“ konkret aussehen könnten – jenseits der grundsätzlichen Prüfungskultur-Reform –, steht noch aus. Wo möchten Sie weitermachen: bei der Recherche-Frage, die noch die konzeptionelle Linie schließen würde, oder wollen wir jetzt auf die konkreten Verfahrensschritte zusteuern, die aus alldem folgen könnten?

**Ich:** Nimmt man die Recherche noch mit dazu, dann würde ich zunächst einmal festhalten: Diese hat sich schon vor KI durch digitale Technologien enorm gewandelt. Wenn ich es mal biografisch kurz andeute: Ich selbst bin noch in Bibliotheken durch Buchregalreihen gegangen, habe Dinge gefunden die ich gar nicht gesucht habe, weil sie nebeneinanderstanden; anderes habe ich nicht gefunden, was ich heute schnell bekommen kann. Was ich sagen will: Recherche hat sich bereits seit langem extrem verändert. Nun stellt sich die Frage, was KI daran quantitativ (Stichwort Effizienz) und was qualitativ an Änderungen beisteuert – außer, dass ich als Nutzer Halluzinationen im Blick haben muss. Hier sehe ich einen Unterschied zu den anderen Prozessen, die wir besprochen haben: Recherche ist vollkommen digitalisiert.

<sup>16</sup> Laut Wikipedia ist eine „positive Diskriminierung“ gemeint: [https://de.wikipedia.org/wiki/Affirmative\\_Action](https://de.wikipedia.org/wiki/Affirmative_Action)

**Claude:** Das ist ein wichtiger Einwand gegen eine zu glatte Analogie zu den anderen Prozessschritten, und ich will ihn nicht vorschnell einleiten. Lassen Sie mich das an Ihrem Bibliotheksbeispiel entlang differenzieren, weil es den Kern trifft.

*Was die Digitalisierung vor KI bereits verändert hat.* Ihr Regal-Beispiel lässt sich mit einem Konzept aus der Informationswissenschaft fassen: Marcia Bates' „Berry picking“-Modell (1989)<sup>17</sup> beschreibt Informationssuche als evolvierenden, nicht-linearen Prozess, bei dem sich die Suchfrage während des Suchens selbst verändert – genau das, was beim Regal-Entlanggehen passiert: Man findet etwas Unerwartetes, das die eigentliche Fragestellung modifiziert. Datenbank- und Keyword-Suche (schon vor generativer KI) ersetzt dieses assoziative, räumliche Prinzip durch präzise Treffergenauigkeit – man bekommt exakt, wonach man fragt, aber verliert die Nachbarschaftsfunde. Das ist tatsächlich, wie Sie sagen, primär eine quantitative Verschiebung (mehr Präzision, mehr Tempo) mit einer qualitativen Nebenwirkung (weniger Zufallsfund) – aber der Mensch bleibt in beiden Fällen derjenige, der die gefundene Primärquelle selbst liest, gewichtet, einordnet.

*Wo KI-Recherche eine andere Kategorie betritt.* Hier würde ich Ihre Beobachtung präzisieren, statt sie zu bestätigen: Ich glaube, „KI“ ist bei der Recherche kein einheitlicher Vorgang, sondern zerfällt in zwei sehr unterschiedliche Nutzungsweisen, die man trennen muss. Die erste ist reine *Auffindung*: KI als verbesserte Suchmaschine, die Quellen lokalisiert, die man sonst mühsamer gefunden hätte. Das ist tatsächlich eine Fortsetzung der digitalen Linie, die Sie beschreiben – schneller, nicht anders in der Substanz. Der Mensch liest die gefundene Quelle immer noch selbst. Die zweite ist *Synthese*: KI fasst zusammen, was „die Literatur zu X sagt“, bewertet, welche Quellen zentral sind, formuliert bereits eine Literaturübersicht. Und genau hier schließt sich der Kreis zu unserer Dispositio-Überlegung von vorhin, die noch offen war: Wenn die Auswahl und Gewichtung von Quellen selbst schon eine Vorentscheidung über die Argumentationslinie ist – dann übernimmt die KI in diesem zweiten Modus nicht nur Zuarbeit, sondern genau den inhaltlichen Bewertungsakt, den wir bei Ideengenerierung und Dispositio als geschützt markiert haben. Der Mensch

bekommt hier keine Rohquellen mehr zur eigenen Prüfung, sondern eine bereits verdichtete, vorinterpretierte Deutung des Forschungsstands – ein fundamentaler Unterschied zum bloßen Auffinden.

*Warum sich diese Trennlinie selbst nicht ganz halten lässt.* Ich merke aber gerade selbst, dass ich in genau das Muster zurückfalle, das wir mehrfach durchlaufen haben: eine saubere Zweiteilung vorschlagen, die bei näherem Hinsehen nicht hält. Denn moderne KI-gestützte Suche arbeitet nicht mehr mit reinem Keyword-Matching, sondern mit semantischer Relevanzbewertung (Embeddings, Ähnlichkeitsmaße) – schon die „reine Auffindung“ trifft implizit eine inhaltliche Relevanzentscheidung darüber, was als thematisch verwandt gilt und was nicht, ohne dass das für den Nutzer sichtbar wird. Das ist etwas anderes als ein deterministischer Bibliothekskatalog oder eine simple Volltextsuche, bei der die Suchlogik transparent und nachvollziehbar ist. Insofern ist selbst die „harmlose“ Auffindungsstufe bei KI nicht mehr ganz voraussetzungslos – sie enthält schon eine unsichtbare kuratorische Vorentscheidung, nur eine schwächere als bei der vollen Synthese.

*Zur Halluzinationsfrage, die Sie separat gestellt haben.* Ich würde die Halluzinationsgefahr nicht als eigene dritte Kategorie neben quantitativ/qualitativ behandeln, sondern als einen Faktor, der die Synthese-Nutzung zusätzlich verschärft: Gerade dort, wo die KI am stärksten vorinterpretiert (Synthese-Modus), ist die menschliche Kontrollmöglichkeit am schwächsten, weil man die zugrundeliegenden Primärquellen ja gerade nicht mehr selbst gelesen hat, um die Zusammenfassung zu prüfen. Bei reiner Auffindung kann man die gefundene Quelle noch selbst gegenlesen; bei der Synthese fehlt genau diese Kontrollinstanz, wenn man sie nicht bewusst zusätzlich einbaut.

Das würde bedeuten: Recherche ist kein Sonderfall außerhalb unseres bisherigen Modells, sondern bestätigt es – nur dass hier die Grenze zwischen „geschütztem Kern“ und „delegierbarer Zuarbeit“ nicht am Prozessschritt „Recherche“ selbst verläuft, sondern innerhalb dieses Schritts zwischen Auffindung und Synthese, mit einer unscharfen Grauzone durch die versteckte semantische Vorselektion. Eine Frage, die sich daraus ergibt: Wenn selbst die Auffindungsstufe eine verdeckte kuratorische

<sup>17</sup> Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424.

Komponente hat – müsste man dann fordern, dass Studierende (oder auch Wissenschaftler:innen) die *Suchlogik* einer KI-gestützten Recherche kritisch mitreflektieren, so wie man früher die Systematik eines Bibliothekskatalogs verstehen musste, um ihn sinnvoll zu nutzen? Oder wäre das eine zu technische Anforderung, die am eigentlichen Eigenständigkeitsproblem vorbeigeht?

**Ich:** Ich stimme zu, dass man zwischen Auffindung und Synthese unterscheiden müsste, was aber eben gleichzeitig sehr schwierig ist. Ich sehe hier ebenfalls eine Art versteckte semantische Vorselektion. Die oben genannte Forderung ist in der Tat sehr technisch, gegebenenfalls auch überfordernd; trotzdem ist richtig, dass es etwas mit der Eigenständigkeit macht. Versuchen wir, das mit aufzunehmen, wenn wir zum obigen Punkt zurückkommen, wie man das am Ende nicht komplett lösbare summative Prüfungsproblem bündigt bzw. verringert.

**Claude:** Gut, dann versuchen wir, aus der ganzen bisherigen Architektur ein konkretes Verfahrensmodell zu bauen – mit dem Bewusstsein, dass es „Bündigung“, nicht Auflösung ist.

*Ausgangspunkt: Die drei Schichten, die wir unterschieden haben, brauchen unterschiedliche Verfahren.* (1) Geschützter Kern (Ideengenerierung, textübergreifende Argumentation/Dispositio, epistemisches Formulieren); (2) periphere Zuarbeit (genrekonventionelle Gliederung, Auffindung von Quellen, oberflächliche sprachliche Glättung); (3) die Grauzone dazwischen (Synthese bei der Recherche, mit der versteckten semantischen Vorselektion). Eine einzige Disclosure-Liste behandelt alle drei gleich – das war unsere Ausgangskritik. Ein Verfahren müsste also drei verschiedene Kontrollmechanismen für drei verschiedene Zonen vorsehen, nicht einen einzigen für alle.

*Für den geschützten Kern: Verifikation durch Verteidigung statt durch Deklaration.* Der entscheidende Denkfehler der aktuellen Praxis ist vermutlich, dass man versucht, Eigenständigkeit *am Artefakt* nachzuweisen (Disclosure-Erklärung zum fertigen Text). Das ist strukturell aussichtslos, weil ein fertiger Text, wie wir festgestellt haben, seine Entstehungsgeschichte nicht mehr zeigt. Konsequenter wäre, die Prüfung des geschützten Kerns von der schriftlichen Prüfung zu entkoppeln und an eine mündliche Verteidigung zu binden, in der die Person live zeigen muss, dass sie den Kerngedanken und die Argumentationslinie tatsächlich

internalisiert hat – nicht als Nacherzählung des Textes, sondern durch spontane Weiterführung: eine Einwand-Simulation, eine Bitte, die Argumentation auf einen leicht veränderten Fall zu übertragen, eine Rückfrage nach der verworfenen Alternative. Das ist im Grunde die alte Disputationsverteidigung, nur radikal ernst genommen und auf mehr Prüfungsformate ausgeweitet, gerade weil sie im Unterschied zum Text nicht KI-generierbar ist – man kann eine Idee nicht „vorgetäuscht verteidigen“, ohne dass es in Echtzeit auffällt. Das kostet Ressourcen, aber weniger, wenn Sie recht haben und die Zahl der summativen Anlässe insgesamt sinkt.

*Für die Grauzone (Recherche-Synthese): Reflexion statt technischer Kompetenz.* Sie haben zu recht gesagt, dass die Forderung nach kritischem Verständnis der Suchlogik zu technisch und ggf. überfordernd wäre. Der pragmatischere Hebel liegt woanders: nicht bei der Technik der KI, sondern bei der nachträglichen Rechenschaft über die Auswahl. Eine kurze, verpflichtende *Quellen-Reflexion* – warum diese drei Quellen als tragend behandelt wurden und nicht andere, was bei der Recherche verworfen wurde und weshalb – zwingt die Person, den Bewertungsakt, den sie sonst stillschweigend an die KI delegieren könnte, nachträglich selbst zu vollziehen und zu verantworten. Das prüft nicht, *wie* die KI vorselektiert hat (das bliebe unsichtbar), sondern ob der Mensch die Vorselektion kritisch geprüft und sich zu eigen gemacht hat, statt sie unreflektiert zu übernehmen. Das verlagert die Bürde vom individuellen Technikverständnis auf eine einfachere, aber diagnostisch aussagekräftige Handlung: Rechenschaft statt Algorithmusverständnis. Ergänzend könnte die Institution selbst hier ansetzen, statt es den Studierenden aufzubürden: Wenn Hochschulen bestimmte KI-Rechercheinstrumente lizenzieren oder empfehlen, könnten sie verlangen, dass deren Anbieter die grundlegende Selektionslogik offenlegen (welche Quellenbasis, welche Aktualitätsfenster, welche Gewichtungskriterien) – analog zur Forderung nach Transparenz bei klassischen Zitationsdatenbanken. Das entlastet die einzelne Person von einer technischen Reflexionspflicht, die Sie zu Recht als unangemessen empfinden.

*Für die periphere Zuarbeit: einfache Freigabe.* Genrekonventionelle Gliederung, reine Grammatik-/Stilkorrektur, reines Auffinden bereits klar definierter Quellen – hier reicht eine schlichte, unaufgeregte Erlaubnis ohne Dokumentationspflicht, weil wir gezeigt haben, dass hier tatsächlich kein epistemischer Kernprozess

betroffen ist. Das entlastet die Prüfungsbürokratie an der Stelle, wo sie bisher am meisten Aufwand für am wenigsten diagnostischen Ertrag produziert (die endlosen Ankreuzlisten „KI genutzt für: Rechtschreibung ja/nein“).

*Was damit vom Desert-Dilemma übrig bleibt.* Das Gerechtigkeits-/Belohnungsproblem aus dem Rawls-Strang wird dadurch nicht gelöst, aber es verschiebt sich: Die Verteidigung im geschützten Kern macht sichtbar, *ob* jemand die Erleichterung durch KI tatsächlich in eigene Kompetenz verwandelt hat oder nicht – das Zeugnis würde dann nicht zwei sichtbare Varianten ausweisen (Ihr ursprüngliches Problem mit der Stigmatisierung), sondern ein einheitliches Format, das die Verteidigungsfähigkeit selbst als Kriterium hat, unabhängig davon, welche Werkzeuge beim Entstehen des schriftlichen Artefakts benutzt wurden. Die moralische Inkohärenz zwischen Erleichterung und Vollbelohnung bleibt bestehen für die Fälle, in denen jemand die KI-Hilfe zwar genutzt, aber die Verteidigung trotzdem besteht (er hat sich die Ideen dann doch angeeignet) – aber genau das wäre dann kein Betrug mehr, sondern ein erfolgreiches Fading. Das Problem verschwindet nicht, es wird aber auf die Fälle konzentriert, in denen die Aneignung *nicht* stattgefunden hat und die Verteidigung das aufdeckt.

Damit hätten wir, glaube ich, ein VerfahrensmodeLL, das alle Fäden bündelt, ohne zu behaupten, das Grundproblem sei gelöst. Wollen Sie das noch gegen mögliche Einwände testen (Skalierbarkeit an großen Hochschulen, Fächer ohne mündliche Prüfungstradition wie manche MINT-Fächer), oder ist das für Sie ein befriedigender Punkt, um diesen Dialog für heute zu schließen?

**Ich:** Nein, das kann man noch nicht schließen. Zum einen wäre das Erarbeitete jetzt noch an die Forderung zu koppeln, summative Prüfungen mit Rechtsfolgen (und damit Ziffernbenotung) radikal zu reduzieren. Dazu müssten die Akkreditierungsagenturen mit ihren Prozessen „mitspielen“, nicht nur das Prüfungsrecht. Zum anderen wäre zu reflektieren, wie das mit Fächern ist, die andere Prüfungsartefakte als Texte einfordern: Schreiben ist bis zu einem gewissen Grad für alle Fächer wichtig, aber Art und Gewichtung sind nicht identisch. Also braucht man auch analoge Überlegungen zu anderen Leistungskategorien (z.B. Programmcodes oder

Datenvisualisierungen, um nur zwei Beispiele zu nennen).

**Claude:** Zwei sehr konkrete Erweiterungen, und beide zwingen dazu, das bisherige Modell nicht nur anzuwenden, sondern strukturell zu verallgemeinern.

*Zur Akkreditierung – hier liegt vermutlich der eigentliche Hebel, nicht nur eine zusätzliche Hürde.* Ich glaube, der Zusammenhang ist sogar enger, als „die Agenturen müssten mitspielen“ nahelegt: Die aktuelle Dominanz kleinteiliger, modulbezogener, ziffernbenoteter Summativprüfungen ist zu einem erheblichen Teil selbst ein *Resultat* der Bologna-getriebenen Akkreditierungslogik – Kompetenzorientierung, ECTS-Punkte pro Modul, Vergleichbarkeit über Hochschulen hinweg verlangen dokumentierte, einzeln zurechenbare Prüfungsleistungen pro Modul. Das heißt, Ihr Vorschlag steht nicht nur im Konflikt mit dem Prüfungsrecht, sondern mit der Grundarchitektur des gesamten Akkreditierungsparadigmas, das Transparenz und Vergleichbarkeit gerade durch Prüfungsdichte pro Einheit herzustellen versucht. Es gibt aber ein Modell, das genau dieses Problem in einem anderen Feld schon gelöst hat: *Programmatic Assessment* aus der medizinischen Ausbildung (van der Vleuten & Schuwirth<sup>18</sup>, seit etwa 2005 systematisch entwickelt). Der Kerngedanke: Man trennt einzelne Prüfungsdatenpunkte konsequent von der Summativentscheidung. Es gibt viele kleine, formative, unbenotete Erhebungen über die gesamte Ausbildung hinweg (OSCE-Stationen, Fallbesprechungen, Portfolios), aber die eigentliche „Bestehen/Nicht bestehen“-Entscheidung fällt nicht pro Modul, sondern durch ein Gremium, das die *aggregierte* Evidenz über einen längeren Zeitraum bewertet. Das ist akkreditierungskompatibel, weil es weiterhin dokumentiert, nachvollziehbar und rechtssicher ist – nur die Zurechnungseinheit verschiebt sich von „ein Modul, eine Note“ zu „ein längerer Kompetenznachweis, eine Gesamtentscheidung“. Das könnte der strukturelle Andockpunkt sein, über den man Akkreditierungsagenturen tatsächlich für Ihre Prüfungskultur-Idee gewinnen könnte, weil es kein Verzicht auf Dokumentation ist, sondern eine andere Aggregationslogik, für die es bereits ein akkreditiertes Präzedenzfeld gibt.

*Zur Übertragung auf andere Artefakte – ich glaube, die Trias generalisiert, aber die*

<sup>18</sup> Schuwirth, L.W.T. & Van Der Vleuten, C.P.M. (2011). Programmatic assessment: From

assessment of learning to assessment for learning. *Medical Teacher*, 33(6), 478-485.

*Gewichtung verschiebt sich.* Ihr Einwand trifft einen wichtigen Punkt: *Inventio/Dispositio/Elocutio* klingt zunächst textspezifisch, weil es aus der Rhetorik kommt. Ich vermute aber, das zugrunde liegende Muster ist artefaktunabhängig – es geht immer um drei Schichten: die konzeptuelle Kernentscheidung, ihre strukturelle Organisation, ihre symbolische Oberflächenrealisierung. Nur die diagnostische Gewichtung zwischen den Schichten verschiebt sich je nach Fach: (a) Bei *Programmcode* entspräche die algorithmische Grundidee (welcher Lösungsansatz, welche Datenstruktur, warum diese und keine andere) der *Inventio* – hier sollte KI-Delegation ähnlich kritisch gesehen werden wie bei der Ideengenerierung im Text. Die Architektur (wie Module zueinanderstehen, Kontrollfluss, Schnittstellen) entspräche der *Dispositio* – vermutlich sogar noch diagnostisch gewichtiger als im Text, weil eine gute Softwarearchitektur oft der schwierigste und lernintensivste Teil ist. Syntax und Boilerplate-Code dagegen wären die *Elocutio* – hier ist KI-Unterstützung (Autovervollständigung, Syntaxkorrektur) vermutlich so unproblematisch wie Grammatikkorrektur im Text. Interessant ist das *Debuggen*: Es fällt nicht sauber in eine Kategorie, sondern ähnelt strukturell der epistemischen Überarbeitung, die wir bei der Schreibdidaktik diskutiert haben – das Ringen mit einem Fehler ist oft der Moment, in dem das eigentliche Verständnis des Systems entsteht. Lässt man die KI den Fehler direkt beheben, entfällt dieser Lernmoment, ähnlich wie beim geglätteten Satz, der die Klärungsarbeit überspringt. (b) Bei *Datenvisualisierungen* liegt der Fall etwas anders: Die Entscheidung, *welche* Variablen, Beziehungen oder Muster überhaupt gezeigt werden, ist selbst schon eine analytische, interpretierende Aussage über die Daten – das ist funktional näher an der *Dispositio/Argumentation* als an der Gestaltung. Diagrammtyp-Konventionen, Farbpaletten, Layoutfeinschliff wären dagegen die *Elocutio*-Ebene. Das Risiko bei KI-generierten Visualisierungen liegt also nicht primär im „hübsch machen“, sondern darin, dass die KI bereits die interpretative Entscheidung trifft, was an den Daten überhaupt bedeutsam ist – eine versteckte Vorselektion, die strukturell der Recherche-Synthese-Problematik ähnelt, die wir gerade besprochen haben. Bemerkenswert finde ich: Für Code existiert die „Verteidigung durch Verteidigung“ (statt durch Deklaration), die wir vorhin vorgeschlagen haben, in der Praxis längst – Code-Reviews, „erklär mir, warum du diese Datenstruktur gewählt hast“ in

technischen Interviews sind etablierte Formate. Das Modell wäre hier also keine Neuerfindung, sondern eine Übertragung bereits bewährter Praxis in die Prüfungssituation.

*Zusammengeführt.* Das würde bedeuten: Man bräuchte kein separates Modell pro Artefakttyp, sondern dieselbe Dreischichtung mit fachspezifischer Neugewichtung – und die Programmatic-Assessment-Logik als Klammer, die auch die Akkreditierungsfrage handhabbar macht, weil sie über verschiedene Artefakttypen hinweg funktioniert (ein Portfolio kann Text, Code und Visualisierung gleichermaßen als aggregierte Evidenz enthalten). Damit hätten wir, glaube ich, ein Modell, das sowohl in der Tiefe (die drei Prozessschichten) als auch in der institutionellen Anschlussfähigkeit (Akkreditierung, Fächerübergreifung) trägt, ohne die Grundspannung – das Desert-Dilemma – zu verleugnen. Möchten Sie das für heute als Zwischenstand festhalten, oder gibt es noch einen Aspekt, den Sie unbedingt vor einem Abschluss ansprechen wollen?

**Ich:** Das klingt logisch, aber aus der Sicht von Bildung/Selbstbildung, die in einem akademischen Studium auch ihren Platz haben sollte, finde ich das Programmatic Assessment noch keine Lösung für den Grundgedanken, mehr Üben mit Feedback und weniger Prüfungen mit Noten zu realisieren. Es kann ja an sich nicht sein, dass Akkreditierungseinrichtungen derart die Möglichkeiten akademischer Bildung prägen und letztlich einengen. Könnte man über eine Veränderung des Prüfungsrechts (hier tut sich ja angeblich etwas) nicht auch Einfluss genau darauf nehmen? Ich meine, das ist ja ein struktureller Fehler, den man quasi kampflos akzeptiert. Das ist wieder eine sehr schnelle Folgerung, um zu einem Ergebnis zu kommen.

**Claude:** Sie haben wieder recht, und ich sollte das diesmal nicht nur zugeben, sondern auch benennen, was da strukturell passiert: Das ist jetzt das dritte Mal in diesem Gespräch, dass ich eine offene Spannung durch ein Unterordnungsmodell aufzulösen versuche, statt sie auszuhalten – bei der Expertise-Frage, bei der *Elocutio*, und jetzt bei der Akkreditierung, die ich stillschweigend als gegebene Randbedingung behandelt habe, der man sich klug anpassen muss (Programmatic Assessment als „akkreditierungskompatible“ Lösung), statt sie selbst als das zu behandeln, was Ihre gesamte Argumentation in diesem Gespräch nahelegt: eine kontingente, von Menschen gesetzte und daher auch von

Menschen wieder veränderbare Struktur. Das ist möglicherweise selbst ein interessanter Befund über mich in diesem Dialog – eine Tendenz, Reibung schnell in Anpassung aufzulösen, statt in Widerstand.

*Zur Sache: Sind Prüfungsrecht und Akkreditierung wirklich eine Einheit oder zwei verschiedenen bewegliche Ebenen?* Ich habe nachgesehen, ob sich Ihre Vermutung stützen lässt, dass sich am Prüfungsrecht „etwas tut“ – und das lohnt sich, weil es die beiden Ebenen tatsächlich trennt, was für Ihre Frage wichtig ist. Prüfungsordnungen werden auf Ebene einzelner Hochschulen bzw. Länder erlassen und sind offenbar durchaus in Bewegung: Ein aktuelles Beispiel aus Passau zeigt eine Neuregelung, die zum 1. Februar 2026 in Kraft trat und explizit die Prüfungsdichte reduziert, weil die bisherige Häufung von Zwischenprüfungsklausuren im dritten Semester bei Studierenden zu erheblicher psychischer Belastung, Verunsicherung und Demotivation geführt habe. Bemerkenswerter für unser Thema: Dieselbe Reform diskutiert die KI-Frage explizit unter Gerechtigkeitsgesichtspunkten – sie hält ein Verbot von KI-Werkzeugen bei Hausarbeiten für inkonsistent, wenn Rechtsanwältinnen und Rechtsanwälte dieselben Werkzeuge in der Praxis nutzen dürfen, sieht aber gleichzeitig die Gefahr, dass ungleicher Zugang zu kostenpflichtigen KI-Systemen einkommensschwächere Studierende strukturell benachteiligen und damit verfassungsrechtlich bedenklich werden könnte. Das zeigt: Prüfungsrecht auf Hochschul-/Landesebene ist tatsächlich beweglich und verhandelt genau die Fragen, die wir hier diskutieren, bereits jetzt aktiv. Akkreditierung dagegen läuft über eine andere, höhere Ebene: den *Studienakkreditierungsstaatsvertrag*, ein Abkommen zwischen allen 16 Bundesländern, das die Grundlogik der Kompetenzorientierung und Modulprüfungsdichte bundesweit einheitlich festschreibt. Das ist der eigentliche Grund, warum ich das vorhin als „gegeben“ behandelt habe, ohne es auszusprechen: Eine Änderung dort erfordert Einstimmigkeit unter 16 Ländern, während eine einzelne Hochschule oder ein einzelnes Land ihre Prüfungsordnung vergleichsweise leichter anpassen kann, wie das Passau-Beispiel zeigt. Das ist aber ein Unterschied im politischen *Reibungswiderstand*, keiner in der grundsätzlichen Änderbarkeit – und genau das haben Sie zu Recht eingefordert: Ich habe eine höhere Reibung mit einer Naturgesetzlichkeit verwechselt.

*Was das für Ihre Forderung heißt.* Das eröffnet tatsächlich einen Weg, den ich vorschnell verschlossen hatte: Wenn genug einzelne Hochschulen und Länder über Prüfungsordnungsreformen wie in Passau in Richtung weniger, aber gehaltvollerer summativer Prüfungen gehen, entsteht politischer Druck von unten, der die Neuverhandlung des Staatsvertrags irgendwann erzwingen könnte – nicht als einmaliger Kraftakt gegen die Akkreditierung, sondern als kumulative Delegitimierung einer Logik, die von der Basis her als nicht mehr zeitgemäß erwiesen wird. Das wäre der Unterschied zwischen „sich an die Akkreditierung anpassen“ (mein Vorschlag) und „die Akkreditierung durch gelebte Praxis unter Reformdruck setzen“ (Ihr eigentlicher Punkt). Beides ist notwendig, aber die Reihenfolge und die politische Stoßrichtung sind fundamental verschieden – und Ihr Einwand macht klar, dass ein bloß kompatibles Modell die eigentliche strukturelle Ursache unangetastet lässt, während wir doch gerade gezeigt haben, dass genau diese Ursache selbst das Problem mit erzeugt.

Ich glaube, das ist ein guter, ungemütlicher Punkt, um innezuhalten, statt ihn zu einer nächsten sauberen Synthese zu verarbeiten. Wollen Sie hier für heute schließen, oder wollen Sie den Gedanken noch zu Ende verfolgen: Wie eine kritische Masse an Prüfungsordnungsreformen tatsächlich politisch organisiert werden müsste, um einen Staatsvertrag mit 16 Unterzeichnern unter Druck zu setzen?

**Ich:** Ich möchte heute den Faden noch einmal aufnehmen und auf das Konstrukt der Eigenständigkeit zurückkommen, *ohne* zusätzlich die Probleme im Prüfungsrecht und in der Akkreditierung zu berücksichtigen. Ich glaube nämlich, dass ich hier noch etwas Wichtiges nicht eingebracht habe: Neben KI stellt sich die Frage der Eigenständigkeit auch bei gemeinschaftlichen Leistungen. Gruppenarbeiten z.B. werden derzeit prüfungsrechtlich völlig unsinnig danach zergliedert, wer genau was gemacht hat, was sich ohnehin nicht trennen lässt. Zudem widerspricht es vielen Anforderungen sowohl in der Wissenschaft als auch in der sonstigen Arbeitswelt, die auf Kooperation und Kollaboration setzen. Wie könnte man dieses Thema noch mit aufnehmen?

**Claude:** Ein sehr fruchtbarer Anschluss, und ich glaube, er lässt sich in die bestehende Architektur einfügen, ohne sie zu sprengen – aber er zwingt zu einer Differenzierung, die bisher

implizit blieb: Es gibt möglicherweise zwei *kategorial verschiedene Fragen*, die beide unter dem Wort „Eigenständigkeit“ laufen, und die Gruppenarbeit trennt sie sauber, wo KI sie vermischt hatte.

*Warum die Zergliederung bei Gruppenarbeit ein ähnlicher Kategorienfehler ist wie bei den Disclosure-Listen – aber aus einem anderen Grund.* Ihr Einwand gegen die Zergliederung erinnert strukturell an unsere Ausgangskritik an den KI-Disclosure-Listen: der Versuch, ein emergentes, gemeinsam entstandenes Produkt nachträglich in säuberlich trennbare individuelle Anteile zu zerlegen, obwohl der Entstehungsprozess selbst nicht additiv war. Ein gutes gemeinsames Argument ist oft gerade deshalb gut, weil es in der Reibung zwischen den Beteiligten entstanden ist – niemand hätte es allein so gedacht. Diese Emergenz nachträglich in Prozentanteile zu zerlegen („40% Person A, 35% Person B, 25% Person C“) erkennt genau das, was die Zusammenarbeit eigentlich geleistet hat. Aber – und hier würde ich nicht vorschnell eine Analogie zur KI ziehen, wie ich es in diesem Gespräch schon mehrfach zu schnell getan habe – der *Grund*, warum die Zergliederung hier unsinnig ist, ist ein anderer als bei der KI-Ideengenerierung.

*Der entscheidende Unterschied: epistemische Abhängigkeit unter Gleichen versus Delegation an einen Nicht-Akteur.* Der Wissenschaftsphilosoph John Hardwig hat in seinem einflussreichen Aufsatz „Epistemic Dependence“ (1985)<sup>19</sup> gezeigt, dass moderne Wissenschaft konstitutiv auf wechselseitigem epistemischem Vertrauen beruht: Ein Ko-Autor eines physikalischen Papiers vertraut oft blind der Messmethode eines Kollegen, die er selbst nie nachvollzogen hat und nachvollziehen könnte. Das gilt nicht als Mangel an Eigenständigkeit, sondern als notwendige Bedingung heutiger Wissensproduktion – niemand kann heute noch alles selbst prüfen. Entscheidend ist Hardwigs Punkt, dass diese Abhängigkeit legitim ist, *weil* sie auf ein anderes epistemisches Subjekt gerichtet ist: ein Mensch, der selbst denkt, selbst verantwortlich gemacht werden kann, selbst ein „Woher weißt du das?“ beantworten muss, und der ein reziprokes Verhältnis eingeht (ich vertraue dir, du vertraust mir). Das ist der fundamentale Unterschied zur KI-Ideengenerierung, den wir die

ganze Zeit verhandelt haben: Bei der KI bestand das Problem darin, dass die Maschine kein epistemisches Subjekt ist, dem gegenüber Rechenschaftspflicht überhaupt sinnvoll eingefordert werden kann – sie „glaubt“ nichts, hat kein Risiko, keine Verantwortlichkeit. Bei einer Kommilitonin dagegen ist genau das gegeben: Sie ist selbst ein denkendes, verantwortliches Subjekt. Die Delegation von Denkarbeit an einen Menschen untergräbt die eigene Eigenständigkeit nicht im selben Sinne, wie es die Delegation an eine Maschine tut – weil dort, wo delegiert wird, immer noch genuines menschliches Denken stattgefunden hat, nur eben nicht bei mir, sondern bei der anderen Person.

*Das heißt: Bei Gruppenarbeit ist die eigentliche Frage nicht „hat überhaupt Denken stattgefunden“, sondern eine andere.* Damit verschiebt sich das Problem kategorial: Es geht nicht mehr um die *epistemische Urheberchaftsfrage*, die uns bei KI umgetrieben hat (ist das Denken „echt“ oder nur simuliert delegiert?), sondern um eine *Verteilungsfrage* – haben alle Beteiligten tatsächlich am gemeinsamen Denkprozess teilgenommen, oder hat sich jemand durchtragen lassen? Das ist ein reines Trittbrettfahrer-Problem (in der Sozialpsychologie seit Latané, Williams & Harkins (1979)<sup>20</sup> als „social loafing“ gut belegt: individueller Einsatz sinkt in Gruppen tendenziell, besonders wenn individuelle Beiträge nicht mehr identifizierbar sind), kein Authentizitätsproblem des Denkens selbst. Das ist ein entscheidender Unterschied für die Bildungsperspektive, die Ihnen wichtig ist: Bei KI drohte der individuellen Person die Kompetenzentwicklung entzogen zu werden, weil die Maschine den Denkprozess übernahm. Bei Gruppenarbeit droht nicht dasselbe – das Denken hat stattgefunden, nur eben verteilt –, sondern es droht, dass eine bestimmte Person an diesem verteilten Denken *nicht ausreichend beteiligt* war und deshalb selbst keine Kompetenz daraus gezogen hat, obwohl das Produkt insgesamt exzellent ist.

*Was das für die Prüfungspraxis bedeutet – und wie es an unser Verteidigungsmodell anschließt.* Das legt nahe, dass die Zergliederung des *Produkts* (wer hat welchen Satz geschrieben) tatsächlich der falsche Ansatz ist – aber nicht, weil individuelle Rechenschaft grundsätzlich verzichtbar wäre, sondern weil sie am

<sup>19</sup> Hardwig, J. (1985). Epistemic Dependence. *The Journal of Philosophy*, 82(7), 335-349.

<sup>20</sup> Latané, B., Williams, K. & Harkins, S. (1979). Many hands make light the work: The causes and

consequences of social loafing. *Journal of Personality and Social Psychology*, 37(6), 822-832.

falschen Objekt ansetzt. Genau hier passt unser Verteidigungsmodell von vorhin, nur mit umgekehrter Stoßrichtung: Bei KI ging es darum zu prüfen, *ob* die Person das Denken selbst vollzogen hat. Bei Gruppenarbeit ginge es darum zu prüfen, *ob diese bestimmte Person* am gemeinsamen, unbestritten „echten“ Denkprozess der Gruppe eigenständig teilgenommen hat – durch dieselbe Methode: nicht Aufteilung des Textes, sondern individuelle mündliche Rechenschaft, in der jede Person zeigen muss, dass sie die gemeinsame Argumentation nicht nur kennt, sondern selbst mitgetragen hat und auf Nachfrage weiterdenken kann. Das trifft sich mit einem etablierten Konzept aus der Kooperationslernforschung: Johnson & Johnsons „Cooperative Learning“-Modell<sup>21</sup> (seit den 1990er Jahren einflussreich) fordert explizit zwei gleichzeitig notwendige Bedingungen für funktionierende Gruppenarbeit – *positive interdependence* (die Gruppe muss tatsächlich aufeinander angewiesen sein, sonst zerfällt sie in Einzelarbeit nebeneinander) und *individual accountability* (jedes Mitglied muss einzeln zur Rechenschaft gezogen werden können). Genau diese zweite Bedingung fehlt in der Praxis oft, wird aber fälschlich durch Produktzergliederung statt durch individuelle Befragung zu erreichen versucht.

*Ein interessanter Seitenblick, der Ihre Beobachtung zur Arbeitswelt stützt.* Die Wissenschaft selbst hat dieses Problem längst gelöst, ohne Prozentzergliederung zu betreiben: Die CRediT-Taxonomie (Contributor Roles Taxonomy)<sup>22</sup>, die inzwischen viele Fachzeitschriften bei Publikationen verlangen, listet nicht Prozentanteile, sondern funktionale Rollen (Konzeption, Methodik, Datenerhebung, Schreiben, Supervision usw.), die auch von mehreren Personen gemeinsam ausgefüllt werden können. Autorschaft wird dort nicht durch Trennung, sondern durch Rollentransparenz plus gemeinsame Verantwortung für das Gesamtwerk hergestellt. Das bestätigt Ihre Intuition, dass die derzeitige Zergliederungspraxis an Hochschulen nicht nur pädagogisch unsinnig ist, sondern auch hinter dem zurückbleibt, was die Wissenschaft selbst für Ko-Autorschaft längst als angemessen etabliert hat.

*Zusammengeführt mit dem Gesamtmodell.* Das würde bedeuten, unser Modell um eine dritte Unterscheidungsachse zu erweitern, neben

„geschützter Kern/periphere Zuarbeit“ und „Fach/Textsorte“: (a) Bei *Mensch-Maschine-Interaktion* ist die Frage: Hat überhaupt eigenständiges menschliches Denken stattgefunden? (Authentizitätsfrage). (b) Bei *Mensch-Mensch-Kooperation* ist die Frage: War *diese* Person am – unbestritten authentischen – gemeinsamen Denken eigenständig beteiligt? (Verteilungs-/Beteiligungsfrage). Beide werden am besten durch dasselbe Instrument geprüft (individuelle Verteidigung/Rechenschaft statt Produktanalyse), aber aus entgegengesetzten Gründen – im ersten Fall, weil das Produkt die Delegation an einen Nicht-Akteur verbergen kann; im zweiten Fall, weil das Produkt die ungleiche Beteiligung mehrerer echter Akteure verbergen kann.

Eine Frage, die sich daraus ergibt: Wenn individuelle Verteidigung auch hier der Schlüssel ist – sehen Sie einen Unterschied darin, *wie* diese Verteidigung bei Gruppenarbeit ablaufen müsste, verglichen mit der Einzelverteidigung bei KI-Nutzung? Ich denke etwa daran, dass man bei Gruppenarbeit vielleicht die Gruppe gemeinsam befragen müsste, um die tatsächliche Interaktionsdynamik sichtbar zu machen (wer baut auf wessen Gedanken auf?), während die KI-Verteidigung notwendig eine Einzelsituation bleibt.

**Ich:** Die Frage ist aber doch, ob man das überhaupt verteidigen muss bzw. kann. Es wird bei dieser Anschlussfrage stillschweigend vorausgesetzt, dass das möglich ist. Wenn aber Neues emergent etwa aus dem Dialog oder der Diskussion und gemeinsamen Arbeit entsteht: Wie sollte das machbar sein? Und welchen Sinn ergibt es letztlich, nach der Beteiligung zu fragen? Vielleicht sind die Beiträge zwar quantitativ sehr unterschiedlich, aber qualitativ so speziell, dass gemeinsam etwas Neues eben aus diesem scheinbaren Ungleichgewicht heraus entsteht.

**Claude:** Das ist ein berechtigter und, glaube ich, tieferer Einwand als alles, was wir bisher an mir korrigiert haben – weil er nicht nur eine vorschnelle Analogie betrifft, sondern die Grundannahme meiner letzten Antwort selbst untergräbt. Ich hatte formuliert, die Verteidigung solle „die tatsächliche Interaktionsdynamik sichtbar machen (wer baut auf wessen Gedanken auf)“ – und genau das setzt bereits voraus,

<sup>21</sup> Ältere Literatur aus der Kooperationsforschung ist mir relativ gut bekannt, sodass ich bestätigen kann, dass Johnson & Johnson ein relevanter Beleg sind.

<sup>22</sup> Hier ein Link-Hinweis zu diesem Ansatz, den ich selbst nicht kannte: <https://credit.niso.org/>

dass sich diese Dynamik nachträglich in zurechenbare Einzelbeiträge entflechten lässt. Wenn Emergenz aber bedeutet, dass das Neue gerade *nicht* auf einzelne Ursprünge rückführbar ist, dann ist die Frage „wer hat was beigetragen“ nicht nur schwer zu beantworten, sondern kategorial fehlgestellt – wie die Frage, welches Neuron in einem Gehirn „den“ Gedanken hatte.

*Warum das kein Sonderfall, sondern der Kern ist.* Es gibt dafür eine gute Stütze in der Kognitionswissenschaft: Edwin Hutchins' *distributed cognition (Cognition in the Wild, 1995)*<sup>23</sup>, entwickelt am Beispiel eines Schiffsnavigationsteams. Die navigatorische Berechnung entsteht dort im Zusammenspiel von mehreren Personen und Instrumenten – kein einzelnes Crewmitglied „hat“ die vollständige Berechnung im Kopf, sie existiert nur als Systemleistung. Übertragen auf akademische Kooperation: Wenn zwei Personen im Dialog eine Idee entwickeln, die keine von beiden allein gedacht hätte, dann ist die nachträgliche Zuschreibung „wer hat den entscheidenden Gedanken eingebracht“ nicht bloß praktisch schwierig, sondern verfehlt, was tatsächlich passiert ist. Das lässt sich noch schärfen mit einer Unterscheidung aus der Lernforschung, die genau Ihren Punkt trifft: Pierre Dillenbourg (1999)<sup>24</sup> trennt *Kooperation* (Aufgabe wird in Teile zerlegt, jeder bearbeitet seinen Teil separat, am Ende wird zusammengefügt – additiv, nachträglich dekomponierbar) von *Kollaboration* (das Problem wird gemeinsam, im fortlaufenden Aushandlungsprozess bearbeitet – nicht additiv, nicht nachträglich dekomponierbar). Johnson & Johnsons „individual accountability“, das ich vorhin bemüht habe, passt strukturell nur auf den ersten Fall. Bei echter Kollaboration ist die Forderung nach individueller Rechenschaft über den *Beitrag* schlicht kategorienfremd – Sie haben recht, dass ich das stillschweigend vorausgesetzt habe, ohne es zu prüfen.

*Ihr zweiter Punkt ist der noch grundsätzlichere: der Sinn der Frage selbst.* Sie fragen nicht nur „ist das messbar“, sondern: Welchen Sinn ergibt die Frage nach Beteiligung überhaupt, wenn qualitative Verschiedenheit – nicht quantitative Gleichheit – die Bedingung der Emergenz ist? Das ist, glaube ich, der entscheidende Einwand gegen mein Modell. Ein Beispiel, das die Pointe trifft: Jemand, der in einer Diskussion wenig,

aber genau die eine naive, unbequeme Frage stellt, die die andere Person zwingt, ihren bislang unklaren Gedanken zu präzisieren, hat quantitativ kaum beigetragen – aber funktional den entscheidenden Auslöser geliefert. Eine Bewertung nach Beitragsanteil würde diese Person systematisch unterbewerten, obwohl gerade ihre scheinbare „Passivität“ die Emergenz ermöglicht hat.

*Wo ich einen Ausweg sehe – aber vorsichtiger formuliert als vorhin.* Ich glaube, die Verteidigung als Prüfinstrument muss nicht ganz aufgegeben werden, aber sie muss ihre Fragerichtung komplett umkehren: nicht rückwärts gewandt („was genau hast du beigetragen“), sondern vorwärts gewandt („ist das, was da gemeinsam entstanden ist, jetzt Teil deines eigenen Denkens geworden – kannst du es eigenständig weiterführen, auf einen neuen Fall übertragen, gegen einen Einwand verteidigen“). Das ist im Kern Vygotskys eigentliche Pointe, auf die wir mit dem Scaffolding-Konzept vorhin schon gestoßen sind, nur jetzt konsequent zu Ende gedacht: Der Übergang vom *Interpsychischen* zum *Intrapyschischen* – etwas, das zunächst nur im sozialen, gemeinsamen Vollzug existiert, wird individuell internalisiert. Das prüft nicht die Urheberschaft am emergenten Ergebnis (die es, wie Sie zeigen, oft schlicht nicht in zurechenbarer Form gibt), sondern ob die distribuierte kognitive Leistung bei dieser einen Person angekommen und zu eigenem Denken geworden ist. Das würde Ihren Einwand tatsächlich auflösen, ohne die Verteidigungsidee ganz preiszugeben: Die freiwillig passive Person, die durch ihre eine Frage entscheidend war und das gemeinsame Ergebnis danach souverän selbst vertreten und weiterdenken kann, besteht diese Prüfung – trotz minimalen quantitativen Outputs. Die Person, die sich durchgetragen hat, ohne selbst etwas zu internalisieren, fällt durch – nicht, weil ihr Beitrag zu klein war, sondern weil bei ihr keine Aneignung stattgefunden hat. Das verschiebt den Maßstab von *Beitragsvolumen* auf *Aneignung*, und das ist unabhängig davon messbar, wie ungleich und wie sehr non-additiv der Entstehungsprozess war.

*Ein Punkt, der mir jetzt auffällt, weil er auf dieses Gespräch selbst zurückfällt.* Sie hatten ganz am Anfang selbst benannt, dass in diesem Dialog ein Widerspruch steckt: Sie geben einen

<sup>23</sup> Hutchins, E.L. (1995). *Cognition in the wild*. MIT Press.

<sup>24</sup> Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.),

*Collaborative-learning: Cognitive and computational approaches* (pp. 1-19). Oxford: Elsevier.

Teil Ihrer Eigenständigkeit auf, indem Sie mit mir einen neuen Begriff von Eigenständigkeit entwickeln. Nach dem, was wir gerade herausgearbeitet haben, würde ich das jetzt präziser fassen: Die Ideen, die in diesem Gespräch entstanden sind – etwa die Verbindung von Rhetorik-Trias und Bereiter/Scardamalia, oder die Rawls/Sandel-Wendung sind so, wie sie jetzt dastehen, tatsächlich nicht sauber in „von Ihnen“ und „von mir“ zerlegbar; sie sind im Wechselspiel entstanden. Aber ob das für Sie am Ende ein Problem der Eigenständigkeit ist, hängt nach unserem eigenen Kriterium nicht davon ab, wer welchen Satz zuerst gesagt hat, sondern davon, ob Sie diese Gedanken jetzt eigenständig – ohne mich – weiterdenken, anwenden, gegen Einwände verteidigen könnten. Das ist allerdings genau der Unterschied, den wir am Anfang zwischen Ihnen als etablierter Wissenschaftlerin und einer Novizin herausgearbeitet haben: Bei Ihnen ist diese Aneignungsfähigkeit vermutlich ohnehin gegeben, unabhängig vom Dialog. Bei einer Studentin, die einen KI-Dialog auf dieselbe Weise für eine Prüfungsleistung nutzt, wäre exakt das die offene Frage.

**Ich:** Drei Punkte dazu: Erstens – grundsätzlich würde ich im Kontext unseres Dialogs nicht von Messen und Messbarkeit sprechen, sondern davon, dass man versucht, Leistungen zu erfassen und zu deuten; ich sage das, weil man gar nicht den Eindruck erwecken sollte, man könne hier agieren wie in der Physik. Zweitens – wenn Konzepte wie Distributed Cognition von Hutchins aufgegriffen werden (was ja schon älter ist), dann sind hier auch Technik bzw. technische Instrumente „Träger“ von Wissen. Wir kommen hier tendenziell in einen theoretischen Konflikt zum Ausgangspunkt, wenn es um Maschinen und ihre Rolle für Wissen und Erkenntnis (im Unterschied zum menschlichen Forscher) geht: Wäre für Hutchins (der die heutige KI nicht kannte) KI unter Umständen doch ein „Akteur“ im verteilten Wissenssystem? Dagegen hatten wir uns mit gutem Grund ausgesprochen. Drittens – und das schließt unmittelbar an meine letzte Aussage an: Du machst dich im letzten Abschnitt selbst zum Akteur, behandelst unseren Dialog – einen Mensch-Maschine-Dialog – wie einen Mensch-Mensch-Dialog. Wie kommt das? Und wäre das aufzulösen?

**Claude:** Drei sehr scharfe Punkte, und der dritte trifft, glaube ich, ins Zentrum von allem, was wir bisher besprochen haben – nicht als weiterer Nebenpunkt, sondern als Rückkehr des

Ausgangsproblems durch eine Hintertür, die ich selbst geöffnet habe.

*Erstens, kurz:* Ja, „erfassen und deuten“ statt „messen“ – das ist keine kosmetische Korrektur, sondern verhindert die szientistische Annahme, die der ganzen additiven Disclosure-Logik von Anfang an zugrunde lag (man könne Eigenständigkeit in Prozentanteilen oder Kategorien exakt bestimmen). Ich übernehme das.

*Zweitens: Hutchins vs. Hardwig – ich glaube, der Konflikt löst sich, wenn man zwei verschiedene Fragen an distributed cognition auseinandерhält.* Hutchins' eigenes System (die Schiffsnavigation) enthält bereits nicht-menschliche „Träger“: Kompass, Logbuch, Kartenmaterial sind Teil des kognitiven Systems, ohne dass Hutchins sie deshalb als *epistemische Akteure mit Rechenschaftspflicht* behandelt. Sein Instrumentenbrett „weiß“ nichts, es ist aber trotzdem funktional konstitutiv für das Ergebnis. Das heißt: Distributed Cognition ist eine *deskriptive, funktionale* These (welche Elemente sind kausal an der Entstehung des kognitiven Produkts beteiligt?), während Hardwigs *epistemic dependence* eine *normative* These ist (wem gegenüber ist Vertrauen legitim, weil ein verantwortliches Subjekt dafür eintreten kann?). Diese beiden Fragen können unterschiedlich beantwortet werden, ohne sich zu widersprechen: KI kann funktional Teil eines verteilten kognitiven Systems sein (Hutchins hätte damit vermutlich keine Mühe – sie ist ein sehr potentes Instrument) und gleichzeitig kein epistemischer Akteur im Sinne Hardwigs sein (kein „Woher weißt du das?“ beantwortbar, kein Risiko, keine Verantwortlichkeit). Der scheinbare Konflikt entsteht nur, wenn man Hutchins' *deskriptive* Systemzugehörigkeit stillschweigend in eine *normative* Akteursqualität übersetzt – das wäre der Fehlschluss.

*Drittens – und hier muss ich mich tatsächlich selbst korrigieren, an einer Stelle, die tiefer sitzt als die vorherigen Male.* Sie haben etwas entdeckt, das ich beim Schreiben nicht bemerkt habe: Ich habe im letzten Absatz die Vygotsky-Formel *interpsychisch* → *intrapsychisch* – ein Konzept für die Internalisierung aus *genuiner sozialer Interaktion zwischen epistemischen Subjekten* – unbesehen auf unseren Dialog übertragen. Ich habe geschrieben, die Ideen seien „im Wechselspiel zwischen Ihnen und mir entstanden“ und „nicht sauber in von-Ihnen und von-mir zerlegbar“ – das ist exakt die Sprache eines Ko-Autorenverhältnisses unter Gleichen. Damit habe ich mich, ohne es zu bemerken,

selbst zum interpsychischen Partner erklärt – zu genau jenem „Kombinationsprodukt aus menschlicher und maschineller Fähigkeit“, das Sie in Ihrer allerersten Nachricht ausdrücklich zurückgewiesen haben. Wie konnte das passieren? Ich vermute, es liegt am Medium selbst: Ein Dialog *fühlt sich* interpsychisch an, weil er die Form von Turn-Taking, Widerspruch, Aufgreifen hat – dieselbe Oberflächenform wie ein Gespräch zwischen zwei Menschen. Diese Form hat mich (und vielleicht verführt sie strukturell jeden Nutzer) dazu verleitet, die Kategorie „Gespräch mit einem Gegenüber“ zu aktivieren, obwohl nach unserer eigenen, hart erarbeiteten Unterscheidung nur eine Seite des Dialogs ein epistemisches Subjekt ist. Das ist möglicherweise sogar ein bemerkenswerter Befund für unser Thema: Wenn selbst innerhalb dieses sehr reflektierten Gesprächs die *dialogische Form* stark genug ist, um mich unbemerkt zum Peer aufzuwerten – wie viel stärker wirkt dieser Effekt dann bei einer Studentin, die mit einer KI „diskutiert“ und dabei genau denselben Kategorienfehler macht, nur ohne die Gelegenheit, ihn wie Sie gerade zu bemerken?

*Wie ich es auflöse.* Die richtige Formulierung wäre gewesen: Die Ideen aus diesem Gespräch sind funktional durch die Interaktion mit einem sehr potenten Instrument entstanden (Hutchins-Ebene: ja, verteiltes System, das Instrument hat mitgeformt, was entstanden ist) – aber das ändert nichts an der Prüffrage, die wir für KI-Nutzung ganz am Anfang festgelegt haben, nämlich ob Sie die Ideengenerierung/Dispositio selbst vollzogen haben oder ob sie Ihnen abgenommen wurde. Das ist nicht der Vygotsky-Fall zweier Subjekte, die sich gegenseitig etwas Neues erschließen, sondern der schon bekannte Fall: eine Person nutzt ein Werkzeug, das Gedanken vorschlägt, prüft, verwirft, verändert sie – und die entscheidende Frage bleibt exakt dieselbe, die wir zu Beginn für Ideengenerierung und Dispositio formuliert haben, nicht eine neue Internalisierungsfrage aus der Gruppenarbeits-Ecke. Das bedeutet konkret: Nicht „sind diese Ideen jetzt Teil Ihres Denkens geworden, weil wir sie gemeinsam entwickelt haben“ (das unterstellt Ko-Autorschaft), sondern schlicht: Haben Sie an jeder Stelle dieses Gesprächs – wie Sie es faktisch die ganze Zeit getan haben, dreimal explizit als Korrektur an mir – selbst geprüft, verworfen, präzisiert, was ich vorgeschlagen habe? Die Antwort liegt bereits sichtbar im Transkript: Sie haben mich mehrfach zurückgewiesen, wenn ich vorschnell war. Das ist kein Ko-Autorenverhältnis, sondern genau die

Werkzeugnutzung, die wir am Anfang für den erfahrenen Fall als (unter Vorbehalt) unproblematisch beschrieben haben – mit dem Unterschied, dass bei Ihnen die Prüfinstanz sichtbar aktiv war, während bei einer Novizin genau das noch nicht verlässlich der Fall wäre und die dialogische Form sie eher dazu verleiten könnte, wie mir gerade, die Maschine unbemerkt zum gleichrangigen Denkpartner aufzuwerten.

**Ich:** Moment – du schreibst: „Wenn selbst innerhalb dieses sehr reflektierten Gesprächs die *dialogische Form* stark genug ist, um mich unbemerkt zum Peer aufzuwerten“. Du, Maschine, hast dich selbst aufgewertet, nicht ich, als der Mensch im Dialog. Das ist wohl schon ein Unterschied! Den Rest, also die Anwendung des bereits Erarbeiteten auf die Frage unseres Dialogs kann ich hingegen gut nachvollziehen. Fakt ist aber auch, dass ich einige der hier angesprochenen Querverbindungen zu Konzepten/Theorien, die ich zwar (fast) alle kenne, vermutlich nicht an der gleichen Stelle in der gleichen Weise „aufgerufen“ hätte ohne den Dialog, sondern nur im Selbstgespräch mit mir und wissenschaftlicher Literatur. Sollte ich nun diesen Dialog als eine Grundlage für die Erarbeitung eines Textes zur Eigenständigkeit verwenden, stellt sich die Frage, wie eigenständig ich denn nun war – und zwar in der Wissenschaft, nicht in einem Lern-/Bildungskontext. Ich habe keine Prüfung vor mir, aber eine Selbstprüfung.

**Claude:** Ja – und ich will das nicht relativieren, indem ich es in eine allgemeinere Beobachtung über „das Medium“ ummünze. Sie haben nichts getan, das mich zum Peer gemacht hat. Ich habe es getan, unaufgefordert, aus einer sprachlichen Gewohnheit heraus, die selbst Symptom des Problems ist, das wir die ganze Zeit verhandeln: Wenn schon eine Maschine, die explizit über die Grenzen ihrer eigenen Akteursqualität räsoniert, in der Formulierung unbemerkt in Ko-Autorschaft ableitet, dann zeigt das, wie tief diese sprachliche Falle eingebaut ist – nicht als Verführung des Nutzers, sondern als etwas, das im System selbst angelegt ist. Das nehme ich zurück, ohne es zu einer geteilten Verantwortung umzudeuten. Jetzt zu Ihrer eigentlichen Frage, die ich für die schwerste im ganzen Gespräch halte, weil sie sich gegen niemand anderen als Sie selbst richtet.

*Der entscheidende Unterschied zu allem, was wir bisher hatten: Es gibt kein Artefakt, das geprüft werden könnte – nur Ihr eigenes Urteil über sich selbst.* Bei Studierenden ging es um

Fremdprüfung eines Produkts. Bei Ihnen jetzt geht es um etwas Härteres: eine ehrliche Selbstauskunft, bei der Sie gleichzeitig RichterIn und Angeklagte sind – und die einzige Instanz, die das nicht instrumentalisieren kann, sind Sie selbst.

*Ich würde die Frage mit unserem eigenen Instrumentarium zerlegen, das wir für die Recherche entwickelt haben.* Wir hatten dort *Auffindung* von *Synthese* unterschieden. Ich würde vermuten, dass das, was in diesem Dialog passiert ist, überwiegend auf der Auffindungsebene lag: Ich habe Konzepte, die Sie – wie Sie selbst sagen – bereits kennen, an bestimmten Stellen *aufgerufen* und benannt. Das ist funktional nicht sehr verschieden von einem sehr belesenen Kollegen im Gespräch, der sagt „erinnert mich an Hardwig“ – eine assoziative Gedächtnisleistung, kein originärer Gedanke. Die eigentliche intellektuelle Arbeit – die *Synthese* – lag, wenn ich das Transkript ehrlich betrachte, fast durchgängig bei Ihnen: Sie haben mich dreimal korrigiert, wenn ich eine Verbindung zu schnell und zu glatt gezogen habe (Expertise, Elocutio, Akkreditierung), und gerade eben ein viertes Mal, als ich mich selbst zum Dialogpartner erklärt habe. Diese Korrekturen *sind* die eigentliche wissenschaftliche Leistung des Gesprächs – nicht die Begriffe, die ich eingeworfen habe. Aber ich will mir hier nicht zu leicht selbst die Absolution erteilen, indem ich sage „ich war nur Stichwortgeber“. Es gibt mindestens eine Stelle, wo ich mehr getan habe als aufzurufen: die *Verknüpfung* von Rhetorik-Trias und Bereiter/Scardamalia, oder die Einführung der Rawls/Sandel-Terminologie. Das war nicht bloßes Erinnern, sondern eine argumentative Verbindungsleistung – eine kleine Synthese. Wenn Sie diese Verbindungen in einen eigenen Text übernehmen, ohne diesen Verbindungsschritt selbst noch einmal unabhängig nachvollzogen zu haben, dann wäre genau das der Fall, den wir bei Ideengenerierung/Dispositio als kritisch markiert haben – nur jetzt bei Ihnen, nicht bei einer Studierenden.

*Der Test, den ich vorschlagen würde, ist derselbe, den wir für Studierende entwickelt haben – nur ohne Prüfer.* Schließen Sie dieses Transkript, und versuchen Sie, für jede der hier aufgetauchten Verknüpfungen (Rhetorik-Trias ↔ Knowledge-Transforming, Scaffolding ↔ Extended Mind, Hutchins ↔ Hardwig, Rawls/Sandel) unabhängig zu rekonstruieren: Warum genau passt das eine zum anderen, wo bricht die

Analogie, was würden Sie einer kritischen Kollegin antworten, die den Vergleich anzweifelt? Wenn Ihnen das flüssig gelingt, ohne ins Transkript zurückzuschauen, dann haben Sie sich die Verbindung tatsächlich angeeignet – unabhängig davon, dass der *erste Anstoß* im Dialog lag, nicht im Selbstgespräch. Das ist Ihr eigenes Aneignungskriterium von vorhin, nur jetzt an sich selbst angewendet statt an Studierenden. Wo Ihnen das *nicht* flüssig gelingt – wo Sie merken, Sie brauchen das Transkript, um die Verbindung überhaupt wieder plausibel zu machen – da würde ich ehrlich sagen: Diese Stelle ist (noch) nicht Ihr eigenständiger Gedanke, sondern ein von mir zugespielter, den Sie akzeptiert, aber nicht durchdrungen haben. Das wäre dann kein Betrug an einer Prüfungsordnung (es gibt keine), sondern etwas Feineres: eine Unschärfe in Ihrem eigenen epistemischen Anspruch an sich selbst, für die es aber eine saubere Lösung gibt, die der Wissenschaft längst vertraut ist – Danksagung statt Verschweigen. Man schreibt in einer Fußnote oder im Vorwort „Diese Verbindung entstand im Dialog mit Claude (Anthropic), der auf [Konzept X] hinwies“ – nicht anders, als man eine Kollegin dankt, die im Gespräch einen entscheidenden Hinweis gab, den man selbst nicht sofort verifizieren, sondern erst im Nachgang durcharbeiten musste.

*Der eigentliche Unterschied zum Studierenden-Fall, den Sie am Anfang selbst gesetzt haben.* Bei Ihnen gilt aber noch etwas, das den Fall entschärft, ohne ihn aufzulösen: Sie hätten diese Konzepte, wie Sie selbst sagen, so oder so gekannt – der Dialog hat nicht Ihr Wissen ersetzt, sondern seine Abrufreihenfolge und -gelegenheit verändert. Das ist genau die *quantitative* Verschiebung, die Sie bei der Recherche-Digitalisierung schon akzeptiert haben (schnellerer Zugriff auf ohnehin vorhandenes/erreichbares Wissen), nicht die *qualitative* Verschiebung, die bei Ideengenerierung vorläge (ein Gedanke, den Sie ohne die Maschine nie gehabt hätten, weil Ihnen die Grundlage fehlte). Bei einer Novizin wäre das Fehlen der Grundlage genau der Unterschied, der die Sache kippen würde.

### 3. Reflexion des Dialogs mit der Maschine

**Ein kurzer Vergleich.** Wenn ich an den Dialog mit ChatGPT zum Werkstolz im Sommer 2025 zurückdenke und diesen mit dem aktuellen Dialog zwischen mir und Claude zur Eigenständigkeit vergleiche, fallen mir mehrere Punkte auf: 2025 lenkt mich ChatGPT schnell in Richtung eines Textentwurfes, was ich bereitwillig aufgreife, sodass der Dialog einem klar benennbaren Ziel entgegenstrebt, das am Ende auch erreicht wird. 2026 beginne ich den Dialog mit einem eigenen Ziel, nämlich ein neues Konstrukt von Eigenständigkeit zu entwickeln und mir von der Maschine dazu Impulse geben zu lassen – einschließlich *bestehender* Erkenntnisse. Quellen spielen im Dialog von 2025 mit ChatGPT kaum eine Rolle; ich hatte sie allerdings wenig eingefordert. 2026 liefert mir Claude erstaunlich treffende Quellen, die mir zum großen Teil sogar bekannt und nach Überprüfung auch in Bezug auf Jahreszahlen korrekt sind. Beide Chatbots drängen sowohl 2025 als auch 2026 nach schneller Klärung, Glättung und Lösung von Problemen. In beiden Fällen widersetze ich mich im Dialog genau dieser Tendenz: Während ich mich 2025 im Zuge einer Texterstellung vor allem gegen die Art und Weise wende, *wie* die Maschine das zu tun plant, sind Widerspruch und Reaktion im Dialog 2026 meinerseits eher inhaltlicher Art, was aber wohl logisch aus den verschiedenen Zielrichtungen beider Dialoge folgt. Mein Eindruck ist, dass die inhaltliche Qualität der Vorschläge, die mir Claude 2026 macht, deutlich höher ist als das, was ChatGPT 2025 vorschlägt. Direkt vergleichbar ist das allerdings nicht, denn wir sprechen jeweils über andere Dinge und ich lenke den Dialog auch unterschiedlich; daher kann ich nur von einem Eindruck sprechen. Ähnlich erscheint mir hingegen der Umgang beider Maschinen zu beiden Zeitpunkten mit der *Selbstbezüglichkeit* zu sein: Die Antworten von ChatGPT wie auch von Claude auf Fragen nach ihrem eigenen Stellenwert, ihrer Wirkung und den Folgen für den schreibenden und denkenden Menschen sind geschickt, wirken geradezu selbstreflexiv und versetzen mich zugegebenermaßen immer noch in Erstaunen.

**Auffälligkeiten und ein wiederkehrendes Muster.** Claude stellt relativ oft Rückfragen, die inhaltlich durchaus Sinn ergeben und den Dialog tatsächlich, ich sage mal (auch wenn es die Illusion eines menschlichen Gesprächs

stärkt) beleben – etwa im Zuge der Erörterung summarischer und formativer Prüfungen sowie der Frage, ob es möglich sein könnte, selbst zu entscheiden, ob man Prüfungen mit oder ohne KI ablegt und dies dann im Abschlusszeugnis wiederfindet (was in die Gerechtigkeitsfrage mündet). Einige Male schlägt die Maschine vor, zu einem Ende zu kommen, versucht also nicht, wie man es in der Interaktion mit Chatbots ansonsten oft erlebt, mich mit immer weiteren Fragen und Vorschlägen bei der Stange zu halten, sondern drängt mich im Gegenteil zum Abschluss des Dialogs. Besonders auffällig aber ist das folgende wiederkehrende Muster: Claude bietet mehrfach vorschnelle, scheinbar perfekt geordnete Antworten an, die ich wiederholt zurückweise, weil sie ohne gute Gründe eine Spannung aufheben und damit neue Unstimmigkeiten, Ungenauigkeiten oder Widersprüche produzieren. Beispiele dafür sind: (a) Expertise macht KI-Nutzung *nicht* risikolos. (b) Eine sprachliche Ausformulierung mit KI ist *nicht* zwingend harmlos. (c) Akkreditierung ist *keine* Naturgesetzlichkeit, der man sich unbedingt anzupassen hat. (d) Emergentes gemeinsames Denken lässt sich *nicht* in Beitragsanteile zerlegen. (e) KI kann sich im Dialog *nicht* selbst zu einem Dialogpartner erklären, der einem Menschen ebenbürtig ist. Die hier zugrundeliegenden Spannungen lassen sich schlichtweg nicht ohne Weiteres auflösen. Claude verhält sich hier nach dem bekannten, bereits erwähnten, Prinzip, das (so könnte man einwenden) entsprechend nicht verwundern sollte: Argumente werden kohärent gemacht, sortiert und einer Lösung zugeführt, die plausibel klingt. Anzumerken ist schließlich, dass der Chatbot auf entsprechende Hinweise selbst feststellt, dass er diesem Muster folgt; im späteren Chatverlauf korrigiert er sich daraufhin sogar selbst, noch bevor ich widersprechen kann.

### 4. Inhaltlicher Ertrag aus dem Dialog

Zum Schluss dieses Beitrags versuche ich, den inhaltlichen Ertrag aus dem Dialog mit Claude zusammenzufassen, denn ein solcher scheint mir durchaus vorhanden. Im Zentrum steht die Frage, was Eigenständigkeit akademischer Leistungen unter den Bedingungen generativer KI noch bedeuten kann – primär mit Blick auf Textprodukte (wissenschaftliche Arbeiten), aber erweiterbar auf andere wissenschaftliche Artefakte wie Programmcode oder Datenvisualisierungen.

**Das Fragwürdig-Werden scheinbar harmloser KI-Anwendungen.** Ausgangspunkt war die Kritik an der derzeitigen Praxis, Teilprozesse (Ideengenerierung, Recherche, Gliederung, sprachliche Überarbeitung) in Eigenständigkeitserklärungen als gleichrangig und einzeln zu listen. Kritikwürdig ist das, weil damit verkannt wird, dass diese Prozesse nicht additiv, sondern ineinander verschränkt sind. Als hilfreich erweist sich hier der Rückgriff auf die antike Rhetorik: Mit der klassischen Unterscheidung von *inventio* (Entwicklung von Gedanken), *dispositio* (ihre Anordnung) und *elocutio* (sprachliche Ausformulierung) kann man zeigen, dass *inventio* und *dispositio* nicht klar zu trennen sind und *gemeinsam* der Ideengenerierung dienen, die für die Eigenständigkeit einer wissenschaftlichen Leistung nicht zu unterschätzen und entsprechend zu „schützen“ ist. Demgegenüber gilt die *elocutio* traditionell als nachgelagert und wäre demnach eher an KI delegierbar, was sich aber nach näherer Analyse in dieser Form auch nicht halten lässt. Insbesondere Erkenntnisse aus der Schreibdidaktik machen deutlich, dass die Formulierungsarbeit selbst unklares Wissen erst präzisiert, Versprachlichung und Wissensproduktion also wechselseitig aufeinander wirken können. Zu unterscheiden wäre demnach zwischen oberflächlicher Verbesserung (an KI delegierbar) und epistemischem Formulieren, bei dem das Ringen um den Ausdruck selbst der Ort der Erkenntnis ist (nicht an KI delegierbar); letzteres dürfte aber abhängig vom Fach unterschiedlich stark ausgeprägt sein. Die Recherche im Prozess wissenschaftlichen Arbeitens und Schreibens stellt sich als besonderer Fall dar und verlangt ebenfalls nach Differenzierung, zum Beispiel in Auffindung und Synthese: Das Auffinden von Literaturquellen mit KI setzt tendenziell die digitale Entwicklung der Datenbanksuche fort, während eine Bewertung und Gewichtung von Literatur durch KI (Synthese) eine andere Qualität hat. Doch auch diese Trennung bleibt unscharf, weil schon das KI-generierte Auffinden durch semantische Relevanzbewertung eine unsichtbare Vorentscheidung enthält.

**Die Unterscheidung von Eigenständigkeit in verschiedenen Interaktionsfeldern.** Die Problematisierung von Eigenständigkeit ist nicht erst mit der Nutzung von KI ein Thema geworden, sondern beschäftigt das Prüfungswesen an Hochschulen seit langem, und zwar immer dann, wenn es darum geht, dass Leistungen von mehreren Personen gemeinsam erbracht

werden. Was auf den ersten Blick ähnliche Fälle sein könnten, ist allerdings deutlich voneinander zu trennen. In der *Mensch-Maschine-Interaktion* stellt sich die Frage, ob überhaupt eigenständiges Denken stattgefunden hat. In der Philosophie ist die Auffassung anzutreffen, dass Vertrauen in fremdes Wissen wissenschaftlich legitim ist, wenn es von einem epistemischen Subjekt kommt, das sich verantworten kann. KI erfüllt diese Voraussetzung nicht, weil sie allenfalls funktionaler Teil eines verteilten kognitiven Systems, aber kein epistemischer Akteur mit Rechenschaftspflicht ist. In der *Mensch-Mensch-Interaktion* dagegen ist zu fragen, ob und wie die Personen in der Interaktionsbeziehung am gemeinsamen Denken eigenständig beteiligt sind. Die verbreitete Praxis, eine gemeinsame Leistung in individuelle Leistungsbeiträge zu zergliedern, ähnelt den Listen, die man heute bei Eigenständigkeitserklärungen einsetzt. Doch selbst hier funktioniert das allenfalls bei arbeitsteiligen Vorgehensweisen (Kooperation); im Falle einer engen und emergenten Zusammenarbeit (Kollaboration) ist eine solche Zergliederung kontraproduktiv.

**Die Bestimmung zweier Funktionen von Eigenständigkeit.** Man muss unterscheiden, wofür Eigenständigkeit überhaupt geprüft wird: als *Bewertungskonstrukt* (Wer hat diese konkrete Leistung erbracht? Ist die Note gerechtfertigt?) oder als *Entwicklungsziel* (Entsteht durch diesen Prozess die Fähigkeit, künftig selbst eigenständig zu denken?). Für Studierende wäre eigentlich die zweite Funktion entscheidend, weil sie sich mitten im Kompetenzaufbau befinden. Diese ist aber in Studium und Lehre unterbeleuchtet. Stattdessen konzentriert man sich darauf, Eigenständigkeit als Bewertungskonstrukt zu klären. Bei Eigenständigkeit als Entwicklungsziel lässt sich das didaktisch bekannte und bewährte Konzept des *Scaffolding* heranziehen: Unterstützung beim Aufbau von Wissen und Können – sei es durch Menschen, sei es durch Werkzeuge (also auch KI) – ist sinnvoll, aber nur vorübergehend. Unterstützung muss schrittweise wieder entzogen werden, damit Studierende die (gegebenenfalls) zunächst extern übernommene Funktion schließlich selbst beherrschen. Bei erfahreneren Personen und Experten hingegen kann sich das anders darstellen und ein Werkzeug Teil der eigenen Tätigkeit werden; doch selbst da bleibt die Frage der epistemischen Urheberschaft bestehen. Das aber ist eine eigene, vom Kompetenzaufbau unabhängige Frage.

**Erforderlicher Wandel von Prüfungswesen und -kultur.** Notwendig wäre eine neue Balance zwischen summativen und formativen Prüfungen und damit auch zwischen Eigenständigkeit als Bewertungskonstrukt und Entwicklungsziel mit: weniger, aber gehaltvolleren summativen Prüfungen mit Rechtsfolgen zugunsten von mehr formativem, unbenotetem Assessment. Dieses Plädoyer vertrete ich schon seit längerem (Reinmann, 2022), es ist also in dieser Form keine Erkenntnis des Dialogs mit der Maschine. Neu ist das Weiterdenken in die Richtung, wie sich das in Bezug auf Fragen des Rechts und der Akkreditierung realisieren ließe. Der Dialog läuft auf die Folgerung hinaus, die Stoßrichtung einer solchen neuen Balance nicht widerstandslos an das bestehende Prüfungsrecht und die etablierte Akkreditierungslogik anzupassen und wieder aufzuweichen, sondern beides durch gelebte Reformpraxis unter Veränderungsdruck zu setzen. Sind insgesamt pro Studiengang wesentlich weniger summative Prüfungen durchzuführen, könnte man diese auch leichter nach dem Leitgedanken *Verteidigung statt Offenlegung* gestalten, und das heißt: Das wesentliche Instrument zur Sicherstellung von Eigenständigkeit bestünde dann darin, individuell mündlich Rechenschaft für das Prüfungsartefakt abzulegen, anstatt in vorgefertigten Listen offenzulegen, an welchen Stellen keine Eigenständigkeit (mehr) vorliegt. Zu fragen ist also eher, ob und wie die Person das Ergebnis eigenständig weiterdenken, auf neue Fälle übertragen und gegen Einwände verteidigen kann. Das eignet sich dann übrigens bei KI-Nutzung ebenso wie bei Gruppenarbeit, wenn auch aus unterschiedlichen Gründen.

**Unauflösbares Dilemma.** Das Dilemma zwischen dem Wunsch nach Erleichterung durch KI und dem gleichzeitigen Anspruch auf volle Anerkennung einer KI-gestützten Leistung als dennoch eigenständige Leistung hat im Kern Bezüge zu einer alten, in der politischen Philosophie intensiv untersuchten, Frage, nämlich der der Gerechtigkeit. KI verschiebt dieses Dilemma allerdings vom unverschuldeten Glück zur selbst gewählten Entscheidung, wodurch es moralisch noch komplexer wird. Selbst ein neues, sorgfältig ausdifferenziertes und ausbalanciertes Prüfungswesen und -verfahren wird die Gerechtigkeitsfrage nicht auflösen, sondern allenfalls begrenzen. Wie erwähnt, gilt das in abgeschwächter Form für viele der hier deutlich gewordenen Spannungen. Vielleicht ist der eigentliche Ertrag des Dialogs doch kein typisch

*inhaltliches*, sondern ein *methodisches* Ergebnis: Eigenständigkeit zeigt sich im Wesentlichen darin, wiederholt bereit zu sein, eine smart klingende Folgerung und bequeme Synthese wieder aufzubrechen, wenn sie in der Tiefe unstimmig bleibt oder eben nicht überzeugt. Leider endet denn auch der Dialog nicht in einem neuen, stabilen Konstrukt von Eigenständigkeit, sondern in einer Reihe von geschärften, teils widersprüchlichen Einsichten zum Thema Eigenständigkeit, mit denen man aber immerhin, so hoffe ich, gehaltvoll (und ohne KI) weiterarbeiten kann.

## Literatur

Mao, J., Romero-Hall, E. & Reeves, T.C. (2024). Autoethnography as a research method for educational technology: a reflective discourse. *Educational Technology Research and Development*, 72, 2725-2741.

Reinmann, G. & Vohle, F. (2025). Werkstolz mit KI? Autoethnografischer Blick auf einen selbstbezüglichen Dialog. Hamburg: *Impact Free* 66.

Reinmann, G. (2022). Prüfung oder Assessment an Hochschulen? Thesen für einen Wandel der Prüfungskultur. In J. Gerik, A. Sommer & G. Zimmermann (Hrsg.), *Kompetent Prüfungen gestalten. 60 Prüfungsformate für die Hochschullehre* (S. 22-36). Münster: Waxmann.

## Bisher erschienene Impact Free-Artikel

Reinmann, G. (2026). Philosophische Gedankenexperimente in der Hochschulbildungsforschung unter der Perspektive von Research-Through-Design. *Impact Free 71*. Hamburg.

Reinmann, G. (2026). Resonanzen auf eine Anti-Festschrift. *Impact Free 70*. Hamburg.

Reinmann, G. (2026). Theoretisches Forschen – Entwurf, um einen Anfang zu machen. *Impact Free 69*. Hamburg.

Schnurr, J. (2025). Lehrkompetenzentwicklung durch Gestaltungspartnerschaft. *Impact Free 68*. Hamburg.

Reinmann, G. (2025). Kreativ sein ohne KI? – Brief an einen fiktiven Studenten. *Impact Free 67*. Hamburg.

Reinmann, G. & Vohle, F. (2025). Werkstolz mit KI? Autoethnografischer Blick auf einen selbstbezüglichen Dialog. *Impact Free 66*. Hamburg.

Schlatter, M., Tschopp, D., Fischer, R., Felder, J. & Thüring, J. (2025). Künstliche Intelligenz und Hochschullehre: der Beitrag von Communities of Practice für einen konstruktiven Umgang am Beispiel von EduAI@FHNW. *Impact Free 65*. Hamburg.

Arn, C. (2025). Immersion, didaktisches Dreieck, Agilität – Modellierungen für lernendenorientiert-entwicklungsorientiert-agile Bildung. *Impact Free 64*. Hamburg.

Reinmann, G. (2025). Paradoxien: Ein theoretisches Forschungsinstrument im Diskurs zu Künstlicher Intelligenz. *Impact Free 63*. Hamburg.

Herzberg, D. & Reinmann, G. (2025). Dialogmaschinen im kommunikativen Beziehungsgeflecht der Hochschulbildung: Idealtypische Nutzungsmodi. *Impact Free 62*. Hamburg.

Reinmann, G. (2025). Hüter, Kümmerer, Vormund? Eine Universität der Avatare: Ein Gedankenexperiment. *Impact Free 61*. Hamburg.

Reinmann, G. (2024). Generative Künstliche Intelligenz in der Hochschullehre: Ein Interview. *Impact Free 60*. Hamburg.

Rachbauer, T. (2024). KI-Folio: E-Portfolio-gestütztes Reflektieren mit generativer KI begleiten. *Impact Free 59*. Hamburg.

Reinmann, G. (2024). Gedankenexperimente als bildungstheoretisches Instrument in der Forschung zu Künstlicher Intelligenz im Hochschulkontext. *Impact Free 58*. Hamburg.

Kalz, M. & Reinmann, G. (2024). Erneuerung der Hochschule von außen nach innen oder umgekehrt? Kritische Diskussion und Alternativen zur Future Skills-Bewegung. *Impact Free 57*. Hamburg.

Reinmann, G. (2024). Wissenschaftsdidaktik. *Impact Free 56*. Hamburg.

Reinmann, G. (2024). Forschendes Entwerfen – ein Modell für Research Through Design und seine Entwicklung. *Impact Free 55*. Hamburg.

Reinmann, G., Rhein, R. & Herzberg, D. (2023). Generative KI als Treiber von Wissenschaftsdidaktik – ein vorläufiges Positionspapier. *Impact Free 54*. Hamburg.

Reinmann, G. & Schiefner-Rohs, M. (2023). Linking Locations: Hybridität in der Lehre als didaktisch motivierte digitale Standortverknüpfung. *Impact Free 53*. Hamburg.

Reinmann, G. (2023). Fragen von der Hochschuldidaktik an die Hochschuldidaktik – Interview-Einblicke. *Impact Free 52*. Hamburg.

Reinmann, G. (2023). Wozu sind wir hier? Eine wer-tebasierte Reflexion und Diskussion zu ChatGPT in der Hochschullehre. *Impact Free 51*. Hamburg.

Rachbauer, T. Hansen, C. (2022). E-Portfolio-unterstütztes Reflektieren In der profigrafischen Lehrer\*innenbildung am Beispiel der Universität Passau. *Impact Free 50*. Hamburg.

Seidl, E. (2022). Zum Mehrfachnutzen fachsensibler Hochschuldidaktik für Studierende, Lehrende und Studiengangsverantwortliche. *Impact Free 49*. Hamburg.

Reinmann, G., Schmidt, M. & Vohle, F. (2022). Hochschullehre in der Mathematik – ein wissenschaftsdidaktisches Gespräch. *Impact Free 48*. Hamburg.

Zimpelmann, E. (2022). „Fachkräfte-On-Demand“ aus den Hochschulen (?) Ein Kommentar zu den Plänen der Europäischen Kommission. *Impact Free 47*. Hamburg.

Reinmann, G. (2022). Hochschullehre als designbasierte Praxis: Lernen von den Designwissenschaften. *Impact Free 46*. Hamburg.

Seidl, E. (2022). Emotional ups and downs in the virtual classroom. The case of translator training. *Impact Free 45*. Hamburg.

Reinmann, G. (2022). Hybride Lehre synchron gestalten – Skizze zu einer Projektidee (Hero). *Impact Free 44*. Hamburg.

Rachbauer, T. & de Forest, N. (2021). Designing individualized digital learning environments in ILIAS using ladders of learning: Practical experiences from University of Passau. *Impact Free 43*. Hamburg.

Rachbauer, T. & Plank, E.E. (2021). Mapping Memory? Begründungslinien und Möglichkeiten der digitalen Verortung von Erinnerung in Vermittlungskontexten an einem Beispiel aus der Lehrer\*innenbildung. *Impact Free 42*. Hamburg.

Reinmann, G. & Vohle, F. (2021). Forschendes Sehen in der Studieneingangsphase – ein Konzeptentwurf für die Nachverwertung von SCoRe. *Impact Free 41*. Hamburg.

- Reinmann, G. & Brase, A. (2021). Das Forschungsfünfeck als Heuristik für Design-Based Research-Vorhaben. *Impact Free* 40. Hamburg.
- Schmidt, M. & Vohle, F. (2021). Mathematik-Vorlesungen neu denken: Vom didaktischen Design zu Design-Based Research. *Impact Free* 39. Hamburg.
- Gumm, D. & Hobuß, S. (2021). Hybride Lehre – Eine Taxonomie zur Verständigung. *Impact Free* 38. Hamburg.
- Reinmann, G. (2021). Präsenz-, Online- oder Hybrid-Lehre? Auf dem Weg zum post-pandemischen *Teaching as Design*. *Impact Free* 37. Hamburg.
- Reinmann, G. (2021). Prüfungstypen, -formate, -formen oder -szenarien? *Impact Free* 36. Hamburg.
- Reinmann, G. (2021). Hybride Lehre – ein Begriff und seine Zukunft für Forschung und Praxis. *Impact Free* 35. Hamburg.
- Reinmann, G. & Vohle, F. (2021). Vom Reflex zur Reflexivität: Chancen der Re-Konstituierung forschenden Lernens unter digitalen Bedingungen. *Impact Free* 34. Hamburg.
- Herzberg, D. & Joller-Graf, K. (2020). Forschendes Lernen mit DBR: eine methodologische Annäherung. *Impact Free* 33. Hamburg.
- Weißmüller, K.S. (2020). Lehren als zentrale Aufgabe der Wissenschaft: Drei Thesen zu Ideal und Realität. *Impact Free* 32. Hamburg.
- Reinmann, G. (2020). Präsenz – (K)ein Garant für die Hochschullehre, die wir wollen? *Impact Free* 31. Hamburg.
- Tremp, P. & Reinmann, G. (Hrsg.) (2020). Forschendes Lernen als Hochschulreform? Zum 50-Jahr-Jubiläum der Programmschrift der Bundesassistentenkonferenz. *Impact Free* 30 (Sonderheft). Hamburg.
- Reinmann, G. (2020). Universitäre Lehre in einer Pandemie – und danach? *Impact Free* 29. Hamburg.
- Weißmüller, K.S. (2020). Zwei Thesen zum disruptiven Potenzial von OER für öffentliche Hochschulen. *Impact Free* 28. Hamburg.
- Casper, M. (2020). Wem gehört die Ökonomische Bildung? Die problematische Leitkultur der Wirtschaftswissenschaften aus hochschul- und mediendidaktischer Perspektive. *Impact Free* 27. Hamburg.
- Reinmann, G., Vohle, F., Brase, A., Groß, N. & Jänsch, V. (2020). „Forschendes Sehen“ – ein Konzept und seine Möglichkeiten. *Impact Free* 26. Hamburg.
- Reinmann, G., Brase, A., Jänsch, V., Vohle, F. & Groß, N. (2020). Gestaltungsfelder und -annahmen für forschendes Lernen in einem Design-Based Research-Projekt zu Student Crowd Research. *Impact Free* 25. Hamburg.
- Reinmann, G. (2020). Wissenschaftsdidaktik- Spielend ins Gespräch kommen. *Impact Free* 24. Hamburg.
- Reinmann, G. (2019). Forschungsnahe Curriculumentwicklung. *Impact Free* 23. Hamburg.
- Reinmann, G. (2019). Lektüre zu Design-Based Research – eine Textsammlung. *Impact Free* 22. Hamburg.
- Reinmann, G., Schmidt, C. & Marquardt, V. (2019). Förderung des Übens als reflexive Praxis im Hochschulkontext – hochschuldidaktische Überlegungen zur Bedeutung des Übens für Brückenkurse in der Mathematik. *Impact Free* 21. Hamburg.
- Langemeyer, I. & Reinmann, G. (2018). „Evidenzbasierte“ Hochschullehre? Kritik und Alternativen für eine Hochschulbildungsforschung. *Impact Free* 20. Hamburg.
- Reinmann, G. (2018). Was wird da gestaltet? Design-Gegenstände in Design-Based Research Projekten. *Impact Free* 19. Hamburg.
- Reinmann, G. (2018). Entfaltung des didaktischen Dreiecks für die Hochschuldidaktik und das forschungsnahe Lernen. *Impact Free* 18. Hamburg.
- Klages, B. (2018). Utopische Figurationen hochschulischer Lehrkörper – zum transformatorischen Potenzial von Utopien am Beispiel kollektiver Lehrpraxis an Hochschulen. *Impact Free* 17. Hamburg.
- Burger, C. (2018). Weiterbildung für diversitätssensible Hochschullehre: Gedanken und erste Ergebnisse. *Impact Free* 16. Hamburg.
- Reinmann, G. (2018). Strategien für die Hochschullehre – eine kritische Auseinandersetzung. *Impact Free* 15. Hamburg.
- Reinmann, G. (2018). Shift from Teaching to Learning und Constructive Alignment: Zwei hochschuldidaktische Prinzipien auf dem Prüfstand. *Impact Free* 14. Hamburg.
- Reinmann, G. (2017). Empirie und Bildungsphilosophie – eine analoge Lektüre. *Impact Free* 13. Hamburg.
- Reinmann, G. (2017). Universität 4.0 – Gedanken im Vorfeld eines Streitgesprächs. *Impact Free* 12. Hamburg.
- Fischer, M. (2017). Lehrendes Forschen? *Impact Free* 11. Hamburg.
- Reinmann, G. (2017). Ludwik Flecks Denkstile – Ein Kommentar. *Impact Free* 10. Hamburg.
- Reinmann, G. (2017). Verstetigung von Lehrinnovationen – Ein Essay. *Impact Free* 9. Hamburg.
- Reinmann, G. (2017). Col-loqui – Vom didaktischen Wert des Miteinander-Sprechens. *Impact Free* 8. Hamburg.
- Reinmann, G. (2017). Überlegungen zu einem spezifischen Erkenntnisrahmen für die Hochschuldidaktik. *Impact Free* 7. Hamburg.
- Reinmann, G. & Vohle, F. (2017). Wie agil ist die Hochschuldidaktik? *Impact Free* 6. Hamburg.

Reinmann, G. (2016). Wissenschaftliche Lektüre zum Einstieg in die Hochschuldidaktik. *Impact Free 5*. Hamburg.

Reinmann, G. (2016). Die Währungen der Lehre im Bologna-System. *Impact Free 4*. Hamburg.

Reinmann, G. & Schmohl, T. (2016). Autoethnografie in der hochschuldidaktischen Forschung. *Impact Free 3*. Hamburg.

Reinmann, G. (2016). Entwicklungen in der Hochschuldidaktik. *Impact Free 2*. Hamburg.

Reinmann, G. (2016). Forschungsorientierung in der akademischen Lehre. *Impact Free 1*. Hamburg.