Alice E. Rosslenbroich
Denken unter Maschinen
Wie generative künstliche Intelligenz menschliche Kognition verändert

Einleitung

Die rasche Verbreitung generativer künstlicher Intelligenz markiert keinen gewöhnlichen Technologiesprung. Erstmals steht ein Werkzeug zur Verfügung, das nicht nur körperliche oder administrative Arbeit automatisiert, sondern in Bereiche menschlicher Kognition eingreift, die lange als genuin geistige Eigenleistung galten: Formulieren, Strukturieren, Recherchieren, Zusammenfassen, Argumentieren und Problemlösen.

Systeme wie ChatGPT, Claude oder Gemini erzeugen innerhalb weniger Sekunden Texte, Analysen und Antworten, die in zahlreichen Alltagssituationen den Eindruck intellektueller und emotionaler Kompetenz vermitteln. Dadurch verändert sich die Beziehung zwischen Mensch und Maschine grundlegend. Digitale Systeme dienen nicht länger ausschließlich als Werkzeuge zur Informationsverarbeitung, sondern zunehmend als sprachlich adressierbare Assistenten kognitiver Arbeit.

Die gesellschaftliche Relevanz dieser Entwicklung reicht weit über technische Fragen hinaus. Bildungseinrichtungen diskutieren den Verlust eigenständiger Schreib- und Lernprozesse. Unternehmen integrieren KI-Systeme in Wissensarbeit und Entscheidungsstrukturen. Gleichzeitig wächst die Sorge, dass die dauerhafte Auslagerung geistiger Prozesse langfristig Konzentrationsfähigkeit, kritische Reflexion und selbständige Urteilsbildung beeinflussen könnte.

Die Geschichte technischer Innovationen ist zugleich eine Geschichte kognitiver Entlastung. Schrift reduzierte die Notwendigkeit mündlicher Erinnerung. Der Buchdruck externalisierte Wissen in bislang ungekanntem Umfang. Taschenrechner automatisierten mathematische Operationen, Suchmaschinen verkürzten Informationssuche auf wenige Sekunden. Jede dieser Technologien veränderte die Art und Weise, wie Menschen Informationen speichern, verarbeiten und abrufen.

Generative künstliche Intelligenz unterscheidet sich jedoch in einem entscheidenden Punkt von früheren Technologien. Während klassische digitale Werkzeuge vor allem Informationszugang beschleunigten, erzeugen große Sprachmodelle Antworten, Argumentationsskizzen, Zusammenfassungen und Formulierungen. Sie liefern nicht nur Daten, sondern präsentieren sprachlich geordnete Ergebnisse, die häufig den Eindruck eigenständigen Verstehens vermitteln.

ChangeCockpit Coaching & Consulting

Gerade diese sprachliche Form von Problemlösung macht die Technologie gesellschaftlich relevant. Die Interaktion erfolgt nicht mehr primär über technische Bedienlogik, sondern über natürliche Sprache. Der Mensch beschreibt ein Problem, die Maschine erzeugt einen Lösungsvorschlag. Dadurch sinkt die Schwelle zur Auslagerung geistiger Arbeit erheblich.

Bereits frühere Forschung zur Nutzung digitaler Medien deutete darauf hin, dass technische Systeme menschliche Gedächtnis- und Aufmerksamkeitsprozesse beeinflussen können. Die sogenannte Google-Effekt-Studie zeigte 2011, dass Menschen Informationen schlechter erinnern, wenn sie davon ausgehen, diese später digital abrufen zu können. Entscheidend war dabei weniger ein allgemeiner Wissensverlust als die Veränderung der Gedächtnisstrategie: Erinnert wurde primär der Speicherort, nicht mehr der Inhalt selbst.

Quelle: https://www.science.org/doi/10.1126/science.1207745

Mit generativer KI erreicht diese Form kognitiver Auslagerung eine neue Qualität. Erstmals können nicht nur Informationssuche oder Gedächtnisleistungen, sondern auch sprachliche, analytische und strukturierende Denkprozesse an externe Systeme delegiert werden. Damit rückt eine zentrale Frage in den Mittelpunkt der Bildungs- und Kognitionsforschung: Welche Auswirkungen hat es, wenn Menschen immer seltener gezwungen sind, bestimmte geistige Prozesse selbst aktiv zu durchlaufen? Lernen, Problemlösen und kritische Reflexion entstehen nicht ausschließlich durch das fertige Ergebnis, sondern wesentlich durch die mentale Verarbeitung während des Denk- und Arbeitsprozesses selbst.

Die Neurowissenschaften betrachten das Gehirn seit Jahrzehnten nicht als statisches Organ, sondern als dynamisches System kontinuierlicher Anpassung. Neuroplastizität beschreibt die Fähigkeit neuronaler Netzwerke, sich durch wiederholte Nutzung strukturell und funktionell zu verändern. Vereinfacht formuliert kann häufige Aktivierung bestimmte neuronale Verbindungen stärken, während selten genutzte Netzwerke funktionell abgeschwächt werden können. Eric Kandel erhielt den Nobelpreis unter anderem für Arbeiten zu Signalübertragung im Nervensystem und zu biologischen Grundlagen von Lernen und Gedächtnis.

Quelle: https://www.nobelprize.org/prizes/medicine/2000/kandel/facts/

Vor diesem Hintergrund untersucht die vorliegende Abhandlung nicht primär die technische Leistungsfähigkeit künstlicher Intelligenz, sondern deren Einfluss auf menschliche Kognition und auf die neurobiologischen Voraussetzungen des Lernens. Im Zentrum steht die Frage, ob generative KI lediglich Arbeitsabläufe effizienter gestaltet oder ob sie die Bedingungen verändert, unter denen das menschliche Gehirn erinnert, lernt und denkt.

Dabei ist eine differenzierte wissenschaftliche Einordnung notwendig. Für die Annahme unmittelbarer neurologischer Schäden durch KI-Nutzung existieren gegenwärtig keine belastbaren Langzeitdaten. Gleichzeitig deutet die aktuelle Forschung darauf hin, dass intensive KI-Nutzung Prozesse des Cognitive Offloading - also der Auslagerung geistiger Arbeit - verstärken kann. Die entscheidende Frage lautet daher nicht, ob künstliche Intelligenz grundsätzlich schädlich oder nützlich ist, sondern unter welchen Bedingungen sie aktive geistige Verarbeitung unterstützt oder ersetzt.

ChangeCockpit Coaching & Consulting

1: Neuroplastizität und kognitive Anpassung

Die Fähigkeit des menschlichen Gehirns zur Veränderung bildet die neurobiologische Grundlage jedes Lernprozesses. Moderne Neurowissenschaften verstehen das Gehirn nicht als fest verdrahtete Struktur, sondern als hochdynamisches System, dessen neuronale Netzwerke sich kontinuierlich an Erfahrungen, Umweltreize und Nutzungsverhalten anpassen. Dieser Prozess wird als Neuroplastizität bezeichnet.

Neuroplastische Veränderungen entstehen durch wiederholte Aktivierung bestimmter neuronaler Verbindungen. Donald Hebb formulierte 1949 in The Organization of Behavior ein bis heute grundlegendes Prinzip assoziativen Lernens: Wenn eine Nervenzelle wiederholt an der Aktivierung einer anderen beteiligt ist, kann sich die Effizienz dieser Verbindung erhöhen. Die populäre Kurzformel „neurons that fire together wire together“ ist eine spätere Verdichtung dieses Gedankens, nicht die wörtliche Formulierung Hebbs.

Wiederholung kann neuronale Netzwerke stabilisieren; ausbleibende Nutzung kann dagegen zu funktioneller Abschwächung beitragen.

Quelle: https://pure.mpg.de/pubman/item/item_2346268_3/component/file_2346267/Hebb_1949_The_Organization_of_Behavior.pdf

Lernen ist daher kein abstrakter geistiger Vorgang, sondern ein biologischer Anpassungsprozess. Aufmerksamkeit, Gedächtnisbildung, sprachliche Verarbeitung und analytisches Denken beruhen auf wiederholter neuronaler Aktivierung. Besonders relevant ist dabei die aktive kognitive Anstrengung. Studien der Lernforschung zeigen seit Jahrzehnten, dass dauerhafte Wissenskonsolidierung eng mit mentaler Verarbeitungstiefe verbunden ist. Informationen werden insbesondere dann langfristig gespeichert, wenn Menschen Inhalte eigenständig strukturieren, formulieren, abrufen und kritisch verarbeiten.

Die Cognitive-Load-Theory des Psychologen John Sweller beschreibt in diesem Zusammenhang die begrenzte Verarbeitungskapazität des Arbeitsgedächtnisses. Lernen entsteht nicht allein durch Informationsaufnahme, sondern durch aktive Verarbeitung innerhalb dieser begrenzten Kapazität. Wird ein Problem vollständig extern gelöst, kann sich jene kognitive Belastung verringern, die unter bestimmten Bedingungen für tieferes Lernen produktiv wäre. Zugleich ist zu beachten, dass zu hohe Belastung Lernprozesse ebenfalls behindern kann; die lernrelevante Frage betrifft daher nicht Belastung an sich, sondern deren angemessene Dosierung.

Quelle: https://doi.org/10.1207/s15516709cog1202_4

Diese Zusammenhänge besitzen besondere Relevanz für den Einsatz generativer KI. Systeme wie ChatGPT reduzieren nicht nur mechanische Arbeitsprozesse, sondern können sprachliche und analytische Teilschritte übernehmen. Dadurch entsteht ein grundlegender Unterschied zu früheren digitalen Werkzeugen. Während Suchmaschinen primär Informationszugang erleichtern, erzeugen Sprachmodelle vorstrukturierte Antworten, Argumente und Formulierungen. Die aktive geistige Verarbeitung kann dadurch teilweise umgangen werden.

Neurowissenschaftlich betrachtet ist dies relevant, weil neuronale Stabilisierung an Nutzung gekoppelt bleibt. Die Hirnforschung zeigt seit Jahren, dass anspruchsvolle kognitive Aktivität zur Ausbildung stabiler funktioneller neuronaler Netzwerke beitragen kann. Der Neurowissenschaftler Stanislas Dehaene betont in seinen Arbeiten zur Lern- und Leseforschung, dass konzentrierte, aktive Verarbeitung die funktionelle Organisation des Gehirns formt. Lernen entsteht dabei nicht passiv, sondern durch Prozesse gezielter Aufmerksamkeit, aktiver Beteiligung, Rückmeldung über Fehler und anschließender Konsolidierung neuer Inhalte. Lesen, Schreiben und Problemlösen sind demnach keine rein kulturellen Praktiken, sondern greifen in neuronale Lernprozesse ein.

Quelle: https://www.penguin.co.uk/books/310169/how-we-learn-by-dehaene-stanislas/9780141989303

Ähnliche Argumente formuliert die Kognitionswissenschaftlerin Maryanne Wolf im Zusammenhang digitaler Mediennutzung. In Reader, Come Home beschreibt sie die Gefahr einer zunehmend fragmentierten Informationsverarbeitung, in der tiefes Lesen durch schnellere und oberflächlichere Rezeptionsformen verdrängt werden könnte. Wolf argumentiert dabei nicht technikfeindlich, sondern

ChangeCockpit Coaching & Consulting

neurokognitiv: Das Gehirn passt sich an dominante Nutzungsformen an. Werden langsame analytische Prozesse seltener trainiert, könnten entsprechende Fähigkeiten langfristig an Stabilität verlieren.

Quelle: https://www.harpercollins.com/products/reader-come-home-maryanne-wolf

Eine weitere, in der Kognitionswissenschaft zunehmend beachtete Perspektive ergänzt diese Überlegungen: das Konzept der verkörperten Kognition (*Embodied Cognition*). Vertreter dieses Ansatzes — darunter George Lakoff, Mark Johnson und Francisco Varela — argumentieren, dass Denken nicht ausschließlich ein neuronaler Prozess ist, sondern wesentlich durch körperliche Erfahrung, Motorik und sensorisches Feedback geformt wird. Kognition entsteht demnach nicht allein im Gehirn, sondern im Zusammenspiel von Körper, Gehirn und Umwelt.

Für die Frage nach den Auswirkungen generativer KI ist dies aus einem konkreten Grund relevant: Schreiben ist nicht nur ein geistiger, sondern auch ein körperlicher Vorgang. Die motorische Praxis des Schreibens — insbesondere der Handschrift — ist neuronal eng mit Lern- und Gedächtnisprozessen verknüpft. Eine vielzitierte Studie von Pam Mueller und Daniel Oppenheimer aus dem Jahr 2014 zeigte, dass Studierende, die Notizen handschriftlich anfertigten, Inhalte tiefer verarbeiteten und besser behielten als jene, die tippten — vermutlich, weil Handschrift zur eigenständigen Selektion und Umformulierung zwingt, während Tippen schnelleres, aber oberflächlicheres Mitschreiben begünstigt. Wird die sprachliche Produktion zunehmend an generative Systeme abgegeben, entfällt nicht nur die geistige, sondern auch die körperlich-motorische Dimension des Formulierens — ein Aspekt, der in der gegenwärtigen KI-Debatte kaum Beachtung findet.

Diese Überlegungen zur verkörperten Dimension von Kognition unterstreichen, dass die Frage nach den Auswirkungen generativer KI nicht allein auf neuronale Prozesse im engeren Sinne zu reduzieren ist.

Gleichzeitig wäre es wissenschaftlich unhaltbar, aus diesen Mechanismen eine unmittelbare Schädigung des Gehirns durch künstliche Intelligenz abzuleiten. Neuroplastizität ist kein ausschließlich negativer Prozess. Technologische Werkzeuge können kognitive Leistungen ebenso erweitern wie reduzieren. Zahlreiche Studien zur computergestützten Lernunterstützung zeigen, dass interaktive digitale Systeme Lernfortschritte verbessern können, sofern Nutzer aktiv in Problemlösungsprozesse eingebunden bleiben.

Die entscheidende Variable ist daher nicht die Existenz künstlicher Intelligenz selbst, sondern der Grad eigenständiger kognitiver Beteiligung. Aus neurowissenschaftlicher Perspektive ist nicht entscheidend, ob Menschen Werkzeuge verwenden, sondern welche mentalen Prozesse sie währenddessen weiterhin selbst ausführen.

ChangeCockpit Coaching & Consulting

2: Cognitive Offloading und die Auslagerung menschlicher Denkleistung

Die Nutzung externer Hilfsmittel zur Entlastung geistiger Prozesse ist kein neues Phänomen. Bereits Schrift, Bibliotheken und später digitale Speichersysteme ermöglichten es Menschen, Informationen außerhalb des biologischen Gedächtnisses zu speichern. Die moderne Kognitionspsychologie bezeichnet diesen Prozess

als Cognitive Offloading. Gemeint ist die Verlagerung kognitiver Aufgaben auf externe Systeme, um Gedächtnis, Aufmerksamkeit oder Arbeitsgedächtnis zu entlasten.

Grundsätzlich stellt Cognitive Offloading keinen pathologischen Vorgang dar. Im Gegenteil gilt die Fähigkeit zur intelligenten Auslagerung von Informationen als zentraler Bestandteil menschlicher Kulturentwicklung. Problematisch wird der Prozess dort, wo nicht lediglich Speicher- oder Routinefunktionen ausgelagert werden, sondern aktive Denkprozesse von externen Systemen übernommen werden und dadurch seltener aktiv trainiert werden.

Einen frühen empirischen Hinweis auf veränderte Gedächtnisstrategien lieferte die bereits erwähnte Google-Effekt-Studie von Betsy Sparrow, Jenny Liu und Daniel Wegner aus dem Jahr 2011. Die Forscher konnten zeigen, dass Versuchspersonen Informationen schlechter erinnerten, wenn sie glaubten, diese später digital abrufen zu können. Das Gedächtnis fokussierte sich dabei mehr auf den Speicherort als auf den Inhalt.

Diese Studie ist für die heutige KI-Debatte relevant, weil sie zeigt, dass digitale Werkzeuge nicht nur Informationszugang verändern, sondern auch Gedächtnisstrategien beeinflussen können. Sie beweist allerdings nicht, dass digitale Systeme Wissen generell abbauen; sie beschreibt eine spezifische Veränderung des Erinnerns unter bestimmten experimentellen Bedingungen.

Evan Risko und Sam Gilbert beschrieben Cognitive Offloading 2016 als grundlegende Form verteilter Kognition. Menschen nutzen Körper, Umwelt und technische Artefakte, um interne kognitive Belastung zu reduzieren. Dabei spielen wahrgenommene Schwierigkeit und metakognitive Einschätzung eine zentrale Rolle: Je aufwendiger eine Aufgabe erscheint, desto eher greifen Menschen auf externe Hilfen zurück.

Quelle: https://doi.org/10.1016/j.tics.2016.07.002

Im digitalen Alltag hat dieses Prinzip eine neue Reichweite erhalten. Smartphones, Suchmaschinen und Navigationssysteme übernehmen Aufgaben, die zuvor aktives Erinnern oder Orientieren erforderten. Eine 2017 veröffentlichte Studie von Adrian Ward, Kristen Duke, Ayelet Gneezy und Maarten Bos fand Hinweise darauf, dass bereits die bloße Anwesenheit des eigenen Smartphones verfügbare kognitive Kapazität reduzieren kann – “brain drain”. Auch dieser Befund ist vorsichtig zu lesen: Untersucht wurde ein spezifisches experimentelles Setting, nicht eine allgemeine oder irreversible Veränderung der Aufmerksamkeit.

Quelle: https://www.journals.uchicago.edu/doi/10.1086/691462

Generative künstliche Intelligenz erweitert Cognitive Offloading auf komplexe sprachliche und analytische Prozesse. Während frühere digitale Systeme primär Informationen bereitstellten, erzeugen große Sprachmodelle eigenständig formulierte Antworten, Argumentationsstrukturen und Problemlösungen. Dadurch entsteht eine qualitativ neue Form kognitiver Entlastung: Nicht nur Erinnerung oder Suche werden ausgelagert, sondern Teile des Prozesses, in dem ein Gedanke sprachlich geordnet und argumentativ geprüft wird.

Für Lern- und Wissensprozesse ist dies relevant, weil nachhaltiges Verstehen nicht nur vom Ergebnis abhängt. Das eigenständige Formulieren eines Gedankens, das Strukturieren eines Arguments oder das

ChangeCockpit Coaching & Consulting

selbständige Lösen eines Problems erzeugen andere kognitive Anforderungen als die bloße Rezeption fertiger Antworten. Wenn ein Sprachmodell vollständige Essays, Zusammenfassungen oder Analysen erzeugt, kann genau jene geistige Anstrengung reduziert werden, die für tieferes Verständnis bedeutsam ist.

Aktuelle Forschung deutet darauf hin, dass dieser Effekt insbesondere bei hoher Vertrauensbildung gegenüber KI-Systemen relevant wird. Eine 2025 in *Societies* veröffentlichte Studie von Michael Gerlich untersuchte den Zusammenhang zwischen KI-Nutzung, Cognitive Offloading und kritischem Denken. Die Studie fand eine negative Beziehung zwischen intensiver KI-Nutzung und kritischer Reflexionsfähigkeit sowie einen möglichen Zusammenhang mit verstärktem Cognitive Offloading.

Die Aussagekraft dieser Studie ist jedoch begrenzt. Die Ergebnisse sind überwiegend korrelativ. Aus ihnen lässt sich nicht unmittelbar ableiten, dass künstliche Intelligenz kritisches Denken direkt reduziert. Ebenso denkbar ist, dass Personen mit geringerer analytischer Reflexionsneigung häufiger auf KI-Systeme zurückgreifen oder dass weitere Faktoren wie Bildungsgrad, Medienkompetenz, Arbeitsbelastung oder Nutzungszweck eine Rolle spielen.

Quelle: https://www.mdpi.com/2075-4698/15/1/6

Auch die Untersuchung von Microsoft Research und der Carnegie Mellon University beschreibt keine generelle kognitive Verschlechterung durch generative KI. Die Autoren analysierten den Einsatz generativer KI in Wissensarbeit und fanden, dass sich menschliche kognitive Beteiligung verschieben kann. Nutzer tendierten insbesondere bei hohem Vertrauen in KI-generierte Inhalte dazu, weniger eigenständige Kontroll- und Verifikationsprozesse auszuführen. Die Studie untersucht vor allem Selbstberichte und Beispiele aus der Wissensarbeit; sie erlaubt keine Aussagen über langfristige neurologische Veränderungen.

Quelle: https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf

Cognitive Offloading ist deshalb nicht als grundsätzlich schädlicher Prozess zu verstehen. Externe Entlastung kann kognitive Ressourcen freisetzen und komplexere Problemlösungen ermöglichen. Historisch betrachtet beruhen viele wissenschaftliche und gesellschaftliche Fortschritte auf Formen ausgelagerter Kognition - von Schrift bis Computertechnik.

Entscheidend ist daher die Frage, welche kognitiven Prozesse ausgelagert werden und welche weiterhin aktiv trainiert bleiben. Werden Routineaufgaben automatisiert, kann dies höhere geistige Leistungen unterstützen. Werden dagegen grundlegende Fähigkeiten wie Analyse, Strukturierung oder kritische Prüfung dauerhaft ersetzt, könnte langfristig eine Reduktion aktiver kognitiver Praxis entstehen.

Damit verschiebt generative KI die historische Dimension technologischer Entlastung. Sie speichert nicht nur Informationen und beschleunigt nicht nur den Zugang zu Wissen, sondern simuliert sprachliche und analytische Teilschritte menschlicher Kognition. Die Frage ist daher nicht mehr nur, wie schnell Menschen Informationen finden, sondern in welchem Umfang sie die geistige Arbeit der Strukturierung und Prüfung noch selbst vollziehen.

ChangeCockpit Coaching & Consulting

3: Generative künstliche Intelligenz und die Transformation kognitiver Prozesse

Generative Sprachmodelle erweitern das Prinzip des Cognitive Offloading auf eine neue Ebene. Während frühere digitale Systeme vor allem Informationssuche, Datenspeicherung oder repetitive Arbeitsabläufe unterstützten, greifen große Sprachmodelle unmittelbar in sprachliche und analytische Denkprozesse ein. Texte, Argumentationen, Zusammenfassungen und Problemlösungen können heute teilweise automatisiert erzeugt werden — also genau jene Formen geistiger Verarbeitung, die bislang eng mit eigenständigem Denken und Lernen verbunden waren.

Sprache ist dabei nicht nur ein Mittel zur Kommunikation. In vielen Lern- und Arbeitsprozessen ist sie selbst Teil der kognitiven Leistung. Wer formuliert, muss auswählen, gewichten, verbinden, begründen und prüfen. Wird dieser Prozess teilweise an ein Sprachmodell abgegeben, verändert sich nicht nur die Produktionsweise eines Textes, sondern auch die Verteilung kognitiver Arbeit zwischen Mensch und System.

Besonders sichtbar wird dies beim Schreiben. Ein eigener Text entsteht normalerweise durch mehrere kognitive Teilschritte: Problemverständnis, Recherche, Auswahl, Strukturierung, Formulierung, Revision und Selbstprüfung. Generative KI kann einzelne oder mehrere dieser Schritte übernehmen. Der Mensch bleibt dann zwar weiterhin Entscheider, muss aber nicht zwangsläufig alle Zwischenschritte selbst vollziehen.

Diese Verschiebung ist nicht automatisch negativ. In professionellen Kontexten kann sie Rechercheprozesse beschleunigen, Rohfassungen erzeugen oder alternative Argumentationswege sichtbar machen. Problematisch wird sie dort, wo der erzeugte Text nicht mehr als Arbeitsmaterial, sondern als Ersatz eigener Analyse verwendet wird. In diesem Fall reduziert sich nicht nur Schreibzeit, sondern möglicherweise auch jene metakognitive Kontrolle, die beim Formulieren und Überarbeiten entsteht.

Eine viel diskutierte Untersuchung des Massachusetts Institute of Technology aus dem Jahr 2025 versuchte, diese Frage experimentell zu fassen. Die unter dem Titel Your Brain on ChatGPT veröffentlichte Arbeit analysierte mithilfe elektroenzephalographischer Messungen (per EEG) die Gehirnaktivität von Personen beim Schreiben von Essays mit und ohne Unterstützung generativer KI. Die Teilnehmer wurden unter anderem einer LLM-Gruppe (Large Language Model), einer Suchmaschinen-Gruppe und einer Gruppe ohne externe Hilfsmittel zugeteilt.

Die Forscher berichteten über Hinweise auf geringere funktionelle Konnektivität während der Nutzung von ChatGPT. Zusätzlich erinnerten sich Teilnehmer später teilweise schlechter an selbst verfasste Inhalte und beschrieben ein reduziertes Gefühl persönlicher Autorschaft. Diese Befunde passen zu der Annahme, dass stark automatisierte Textproduktion mit geringerer aktiver Verarbeitung einhergehen kann.

Die Aussagekraft der Untersuchung ist jedoch begrenzt. Die Studie lag zunächst als Preprint ohne abgeschlossenes Peer-Review-Verfahren vor, die Stichprobe war vergleichsweise klein und die Untersuchungsdauer kurz. In einer wissenschaftsjournalistischen Einordnung auf der Plattform *Nature* wurde zudem vor einer Überinterpretation der Ergebnisse gewarnt.

ChangeCockpit Coaching & Consulting

Aus den Daten lassen sich weder langfristige neurologische Veränderungen noch dauerhafte kognitive Schäden ableiten.

Quelle: https://arxiv.org/abs/2506.08872

Zusätzliche wissenschaftliche Einordnung: https://www.nature.com/articles/d41586-025-02005-y

Trotz dieser Einschränkungen ist die Studie für die Debatte relevant. Sie liefert keinen Beweis für eine Schädigung des Gehirns, aber einen frühen Hinweis darauf, dass unterschiedliche Schreibbedingungen mit unterschiedlicher kognitiver Beteiligung einhergehen können. Genau diese Frage - nicht eine pauschale Behauptung über „KI und Gehirnschäden“ - ist wissenschaftlich anschlussfähig.

Auch außerhalb neurowissenschaftlicher Laborstudien zeigen sich Hinweise auf veränderte Arbeits- und Denkprozesse. Die bereits zuvor erwähnte Untersuchung von Microsoft Research und der Carnegie Mellon University analysierte, wie Wissensarbeiter generative KI in realen Aufgaben einsetzen. Die Autoren fanden Hinweise darauf, dass sich kritisches Denken teilweise von der eigentlichen Problemlösung hin zur Bewertung maschinell erzeugter Ergebnisse verschiebt. Gleichzeitig nahm bei hohem Vertrauen in KI-generierte Ausgaben die berichtete eigene Prüfaktivität ab.

Für Bildungsprozesse ist diese Verschiebung besonders bedeutsam. Lernen verlangt nicht nur den Zugang zu korrekten Lösungen, sondern die aktive Konstruktion von Verständnis. Wer einen Rechenweg nachvollzieht, eine These formuliert oder einen Text argumentativ ordnet, trainiert andere Fähigkeiten als jemand, der ein fertiges Ergebnis lediglich übernimmt.

Die empirische Bildungsforschung liefert hierzu ein gemischtes Bild. Eine 2025 in Scientific Reports veröffentlichte Untersuchung zu einem KI-gestützten Tutorensystem fand, dass Studierende mit dem AI Tutor in einem bestimmten Setting mehr lernten und weniger Zeit benötigten als in einer Vergleichsbedingung mit aktiver Präsenzlehre. Die Autoren betonten zugleich, dass der Tutor nach pädagogischen Best Practices gestaltet war. Die Studie spricht daher nicht für beliebige KI-Nutzung, sondern für sorgfältig designte, interaktive Lernumgebungen.

Quelle: https://www.nature.com/articles/s41598-025-97652-6

Eine 2025 in den *Proceedings of the National Academy of Sciences* (PNAS) veröffentlichte Untersuchung zum Einsatz von GPT-4 im Mathematikunterricht zeigt die Gegenseite. Schülerinnen und Schüler erzielten während der Nutzung von GPT-4 kurzfristig bessere Leistungen. Wurde der Zugang später entzogen, schnitten insbesondere jene schlechter ab, die zuvor stark auf KI-generierte Lösungswege angewiesen waren. Die Autoren interpretieren dies als Hinweis darauf, dass generative KI kurzfristig Performance verbessern kann, nachhaltiger Kompetenzerwerb jedoch davon abhängt, ob Lernende weiterhin aktiv in den Problemlösungsprozess eingebunden bleiben.

Quelle: https://www.pnas.org/doi/10.1073/pnas.2422633122

Aus diesen Studien ergibt sich kein einheitliches Urteil über generative KI. Vielmehr wird eine Unterscheidung notwendig: KI kann Lernprozesse fördern, wenn sie zum Fragen, Prüfen, Begründen und schrittweisen Verstehen anleitet. Sie kann Lernprozesse behindern, wenn sie zentrale Denkoperationen ersetzt und Lernende nur noch Ergebnisse übernehmen.

ChangeCockpit Coaching & Consulting

Die langfristigen neurokognitiven Folgen dieser Entwicklung sind gegenwärtig noch nicht abschließend erforscht. Der aktuelle Forschungsstand erlaubt weder alarmistische Behauptungen über eine generelle Verdummung durch KI noch technologische Euphorie über uneingeschränkte kognitive Erweiterung. Die empirische Evidenz spricht für ein differenziertes Bild: Generative künstliche Intelligenz verändert die Bedingungen menschlicher Denkprozesse. Ob daraus kognitive Schwächung oder Erweiterung entsteht, hängt wesentlich von der Form menschlicher Beteiligung ab.

4: Aufmerksamkeit, Belohnungssysteme und digitale Effizienz

Die Wirkung generativer künstlicher Intelligenz auf menschliche Kognition lässt sich nicht allein über Lern- und Gedächtnisprozesse erklären. Ebenso relevant ist die Frage, wie KI-Systeme grundlegende Mechanismen von Aufmerksamkeit, Motivation und kognitivem Aufwand beeinflussen. Moderne Neurowissenschaften verstehen Denken nicht als unbegrenzt verfügbare Ressource, sondern als biologisch aufwendigen Prozess. Konzentration, Problemlösung und analytische Reflexion gehen mit erhöhtem neuronalen Energieverbrauch einher.

Das menschliche Gehirn entwickelte sich unter Bedingungen knapper Ressourcen. Evolutionsbiologisch entstand daher eine Tendenz zur Effizienzoptimierung. Menschen bevorzugen in vielen Situationen jene Verhaltensweisen, die mit möglichst geringem kognitivem Aufwand verwertbare Ergebnisse erzeugen. Genau an diesem Punkt treffen generative KI-Systeme auf grundlegende Mechanismen menschlicher Informationsverarbeitung.

Der Psychologe Daniel Kahneman beschreibt in seiner Forschung zwei unterschiedliche Formen menschlicher Informationsverarbeitung. Eine davon arbeitet schnell, intuitiv und mit vergleichsweise geringem kognitivem Aufwand, während die andere langsamere, analytische und bewusst kontrollierte Denkprozesse umfasst. Komplexes Problemlösen, kritische Prüfung und tiefes Verständnis beruhen wesentlich auf dieser zweiten, reflektierten Form der Verarbeitung.

Quelle: https://us.macmillan.com/books/9780374533557/thinkingfastandslow

Generative KI verstärkt diese Dynamik, weil sie analytische Arbeit scheinbar mühelos verfügbar macht. Der Nutzer erhält innerhalb weniger Sekunden Zusammenfassungen, Argumente oder Problemlösungen, ohne längere Phasen konzentrierter Verarbeitung durchlaufen zu müssen. Dadurch sinkt die Schwelle, komplexe geistige Prozesse zu externalisieren.

Neurowissenschaftliche Forschung zeigt, dass dopaminerge Systeme eine zentrale Rolle bei Motivation, Erwartungsverarbeitung und Belohnungslernen spielen. Schnelle Problemlösungen und reduzierte Unsicherheit können daher als kognitiv entlastend und kurzfristig belohnend erlebt werden.

Generative künstliche Intelligenz überträgt die Logik digitaler Effizienz erstmals unmittelbar auf menschliche Denkprozesse. Nicht mehr nur Unterhaltung oder soziale Rückmeldung werden beschleunigt, sondern Schreiben, Strukturieren und Problemlösen. Die Technologie reduziert sogenannte kognitive

ChangeCockpit Coaching & Consulting

Friktion — also jene mentale Reibung, die traditionell mit tiefer Analyse, Unsicherheit und geistiger Anstrengung verbunden war.

Gerade diese Friktion besitzt für Lern- und Erkenntnisprozesse Bedeutung. Die Lernpsychologie beschreibt unter dem Begriff *desirable difficulties* produktive Schwierigkeiten, die langfristige Wissenskonsolidierung fördern können. Robert Bjork und Kolleginnen und Kollegen argumentieren, dass anspruchsvolle mentale Verarbeitung häufig zu stabileren Gedächtnisspuren führt als reibungslose Informationsaufnahme. Die Stärke dieses Effekts hängt jedoch stark von Kontext, Schwierigkeit, Vorwissen und individueller Lernvoraussetzung ab.

Quelle: https://bjorklab.psych.ucla.edu/research/

Ein verwandter, bislang wenig diskutierter Aspekt betrifft die psychologische Dimension kognitiver Abhängigkeit. Der Psychologe Albert Bandura beschrieb mit dem Konzept der Selbstwirksamkeitserwartung (*Self-Efficacy*) die subjektive Überzeugung eines Menschen, bestimmte Aufgaben aus eigener Kraft bewältigen zu können. Diese Überzeugung entsteht wesentlich durch eigene Erfolgserfahrungen — also durch das wiederholte Erleben, dass anspruchsvolle Probleme durch eigene Anstrengung lösbar sind.

Generative KI kann diesen Mechanismus in beide Richtungen beeinflussen. Einerseits kann sie Erfolgserlebnisse ermöglichen, indem sie Hürden reduziert und Lernende zu Ergebnissen führt, die sie allein nicht erreicht hätten. Andererseits besteht das Risiko, dass Menschen, die wiederholt erleben, wie KI-Systeme Aufgaben schneller, umfassender oder sprachlich überzeugender lösen als sie selbst, langfristig an Vertrauen in die eigene kognitive Leistungsfähigkeit verlieren. Eine geschwächte Selbstwirksamkeitserwartung kann in der Folge Motivation, Lernbereitschaft und die Bereitschaft reduzieren, sich eigenständig mit schwierigen Problemen auseinanderzusetzen — unabhängig davon, ob die eigentliche kognitive Leistungsfähigkeit tatsächlich abgenommen hat. Die psychologische Wirkung wahrgenommener Überlegenheit der Maschine kann damit reale Verhaltensfolgen erzeugen, die über rein neurokognitive Prozesse hinausgehen.

Vor diesem Hintergrund — und über rein neurokognitive Prozesse hinaus — entsteht eine zentrale Spannung moderner KI-Nutzung: Technologien, die Denken effizienter machen, können zugleich jene kognitiven Belastungen reduzieren, die für tiefes Lernen unter bestimmten Bedingungen notwendig sind.

Besonders deutlich zeigt sich diese Entwicklung im Umgang moderner digitaler Medien mit menschlicher Aufmerksamkeit. Bereits vor dem Aufstieg generativer KI diskutierten Neurowissenschaftler und Medienforscher die Auswirkungen permanenter digitaler Reizsysteme auf Konzentrationsfähigkeit und Informationsverarbeitung. Nicholas Carr argumentierte in *The Shallows*, dass intensive Internetnutzung fragmentierte Aufmerksamkeitsmuster fördern könne. Diese These ist einflussreich, aber nicht als abschließender neurowissenschaftlicher Konsens zu behandeln; sie bildet vielmehr einen wichtigen Beitrag zur Debatte über digitale Lese- und Aufmerksamkeitspraktiken.

Quelle: https://wwnorton.com/books/the-shallows

Auch Maryanne Wolf beschreibt in ihrer Forschung zur Lesekognition, dass tiefes Lesen komplexe neuronale und kognitive Prozesse aktiviert, die Analyse, Perspektivübernahme und abstraktes Denken

ChangeCockpit Coaching & Consulting

unterstützen. Verkürzte digitale Rezeptionsformen könnten diese Prozesse langfristig abschwächen, wenn intensive Lesepraxis zunehmend durch fragmentierte Informationsaufnahme ersetzt wird.

Generative KI könnte diese Entwicklung weiter verstärken, allerdings auf andere Weise als soziale Medien. Während soziale Medien primär Aufmerksamkeit fragmentieren können, besteht bei Sprachmodellen die Möglichkeit zunehmender kognitiver Passivität: Nutzer erhalten nicht nur Informationsfragmente, sondern bereits synthetisierte Denkprodukte. Die mentale Distanz zwischen Frage und fertiger Antwort schrumpft erheblich.

Gerade im Bildungsbereich erzeugt dies neue Herausforderungen. Lernen beruht nicht ausschließlich auf korrekten Ergebnissen, sondern auf der aktiven Auseinandersetzung mit Unsicherheit, Fehlern und kognitiver Belastung. Wird dieser Prozess zu stark automatisiert, könnte langfristig die Bereitschaft abnehmen, sich eigenständig mit komplexen Problemen auseinanderzusetzen und geistige Anstrengung auszuhalten. Einige Bildungsdebatten beschreiben dies als mögliche Abschwächung der Fähigkeit zur kritischen und eigenständigen Problembearbeitung („epistemische Resilienz“).

Gleichzeitig wäre eine rein negative Interpretation wissenschaftlich nicht haltbar. Zahlreiche neurowissenschaftliche und kognitionspsychologische Modelle gehen davon aus, dass das Gehirn zu energieeffizienter Informationsverarbeitung tendiert. Daraus darf jedoch nicht geschlossen werden, dass jede Form technologischer Entlastung automatisch kognitiv schädlich wirkt. Historisch beruhen kulturelle und wissenschaftliche Fortschritte wesentlich auf der Nutzung externer Werkzeuge zur Erweiterung menschlicher Fähigkeiten.

Die wissenschaftlich relevante Frage lautet daher nicht, ob Menschen Denkprozesse auslagern, sondern welche kognitiven Prozesse weiterhin aktiv trainiert werden. Gegenwärtig existieren nur begrenzte Langzeitdaten darüber, wie sich dauerhafte Nutzung generativer KI auf Aufmerksamkeit, Gedächtnis oder Problemlösung langfristig auswirkt. Mehrere aktuelle Arbeiten beschreiben daher eher mögliche Risiken und theoretische Mechanismen als bereits eindeutig nachgewiesene neurologische Folgen.

Genau hierin liegt die besondere neurobiologische Relevanz generativer KI. Erstmals entsteht ein System, das menschliche Kognition nicht nur ergänzt, sondern sprachliche und analytische Teilschritte simuliert. Dadurch verschiebt sich das Gleichgewicht zwischen eigener geistiger Aktivität und externer Problemlösung. Ob daraus langfristig kognitive Erweiterung oder geistige Abhängigkeit entsteht, hängt wesentlich davon ab, ob Menschen KI als Werkzeug aktiver Reflexion nutzen - oder als dauerhaften Ersatz eigenständigen Denkens.

ChangeCockpit Coaching & Consulting

5: Bildung, Wissensarbeit und die gesellschaftlichen Folgen kognitiver Automatisierung

Die Auswirkungen generativer künstlicher Intelligenz beschränken sich nicht auf individuelle Lern- oder Aufmerksamkeitsprozesse. Mit der zunehmenden Integration von Sprachmodellen in Bildung, Arbeitswelt und Informationssysteme verändert sich die gesellschaftliche Organisation von Wissen selbst. Besonders

betroffen sind jene Bereiche, die traditionell auf eigenständiger Analyse, sprachlicher Verarbeitung und kritischer Reflexion beruhen: Schule, Hochschule, Wissenschaft, Journalismus und wissensbasierte Berufe.

Im Bildungsbereich zeigt sich die Wirkung generativer KI besonders deutlich. Schreiben galt über Jahrzehnte nicht nur als Prüfungsleistung, sondern als Werkzeug geistiger Entwicklung. Der Prozess des eigenständigen Formulierens zwingt Lernende dazu, Informationen zu strukturieren, Zusammenhänge herzustellen und argumentative Kohärenz zu erzeugen. Genau diese Prozesse können durch KI-Systeme teilweise automatisiert werden.

Dadurch entsteht ein pädagogisches Problem. Wenn Schülerinnen, Studierende oder Lernende fertige Antworten übernehmen, entfällt häufig nicht nur die Schreibarbeit selbst, sondern auch ein Teil des zugrunde liegenden Denkprozesses. Bildungsforschung und Bildungspolitik diskutieren deshalb zunehmend, ob generative KI traditionelle Aufgabenformate unterläuft und welche neuen Prüfungs- und Lernformen erforderlich sind.

Die UNESCO veröffentlichte 2023 Leitlinien zum Einsatz generativer KI im Bildungsbereich. Darin betont die Organisation, dass KI-Systeme Lernprozesse unterstützen können, zugleich jedoch Risiken für Autonomie, Datenschutz, Chancengleichheit, kritisches Denken und die menschliche Verantwortung in Bildungsprozessen entstehen. Entscheidend ist eine menschenzentrierte Integration, nicht eine unregulierte Verlagerung von Lern- und Bewertungsprozessen an Maschinen.

Quelle: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Besondere Aufmerksamkeit verdient in diesem Zusammenhang die Frage nach den Auswirkungen generativer KI auf Kinder und Jugendliche. Entwicklungspsychologisch und neurobiologisch unterscheidet sich diese Gruppe grundlegend von erwachsenen Nutzern. Der präfrontale Kortex, der für Planung, Impulskontrolle, kritisches Urteilen und komplexes Problemlösen zentral verantwortlich ist, reift beim Menschen bis weit in die dritte Lebensdekade. Kognitive Fähigkeiten, die in dieser Phase nicht regelmäßig aktiv trainiert werden, können strukturell weniger stabil ausgeprägt bleiben als bei bereits ausgereiften neuronalen Netzwerken.

Daraus folgt keine pauschale Warnung vor KI-Nutzung im Kindes- und Jugendalter. Wohl aber ergibt sich die Notwendigkeit einer entwicklungssensiblen Betrachtung: Was für erwachsene Wissensarbeiter als produktive kognitive Entlastung gelten kann, stellt sich für Lernende in kritischen Reifungsphasen möglicherweise anders dar. Gerade dann, wenn grundlegende Fähigkeiten wie Argumentieren, strukturiertes Schreiben oder analytisches Lesen erstmals erworben werden sollen, kann die frühzeitige Auslagerung dieser Prozesse an generative Systeme die notwendige Übung dieser Fähigkeiten verringern. Bildungspolitik und Pädagogik stehen daher vor der Aufgabe, altersdifferenzierte Rahmenbedingungen für den Einsatz generativer KI zu entwickeln — jenseits pauschaler Verbote ebenso wie jenseits unreflektierter Integration.

Auch die *Organisation for Economic Co-operation and Development* (OECD) untersucht im Projekt *Artificial Intelligence and the Future of Skills*, wie KI-Fähigkeiten mit menschlichen Kompetenzen in Bildung und Arbeit verglichen werden können. Der Schwerpunkt liegt dabei weniger auf unmittelbaren

ChangeCockpit coaching & consulting

Gehirnveränderungen als auf der Frage, welche Fähigkeiten künftig besonders relevant bleiben und wie Bildungssysteme auf die wachsende Leistungsfähigkeit von KI reagieren sollten.

Quelle: https://www.oecd.org/en/publications/ai-and-the-future-of-skills-volume-2_a9fe53cb-en.html

Aktuelle empirische Studien stützen diese Sorge teilweise. Die bereits erwähnte PNAS-Studie zum Einsatz von GPT-4 im Mathematikunterricht deutet darauf hin, dass nachhaltiger Kompetenzerwerb wesentlich davon abhängt, ob Lernende trotz KI-Unterstützung weiterhin aktiv in den Problemlösungsprozess eingebunden bleiben.

Quelle: https://www.pnas.org/doi/10.1073/pnas.2422633122

Diese Dynamik betrifft nicht nur formale Bildung, sondern zunehmend auch Wissensarbeit. Berufe, die traditionell auf Recherche, Analyse und sprachlicher Präzision beruhen, erleben derzeit eine tiefgreifende Transformation.

Die Auswirkungen generativer KI auf Kognition sind dabei nicht ausschließlich als individuelles Phänomen zu verstehen. Denken vollzieht sich in Wissenschaft, Organisationen und demokratischen Gesellschaften wesentlich auch als kollektiver Prozess. Teams entwickeln gemeinsame Problemlösungen, wissenschaftliche Gemeinschaften erzeugen Wissen durch den Widerstreit unterschiedlicher Perspektiven, demokratische Öffentlichkeiten bilden Urteile durch Auseinandersetzung und Deliberation. Generative KI verändert auch diese kollektiven Prozesse, ohne dass dies bislang systematisch untersucht worden wäre.

Ein erstes Risiko betrifft die Homogenisierung von Denken und Argumentation. Wenn große Bevölkerungsgruppen dieselben Sprachmodelle für Analyse, Recherche und Formulierung nutzen, könnte die Vielfalt der erzeugten Perspektiven abnehmen. Sprachmodelle werden auf ähnlichen Datengrundlagen trainiert und tendieren zu statistisch wahrscheinlichen, also häufig konsensuellen Antworten. Damit besteht die Gefahr einer algorithmisch erzeugten Verengung des Meinungs- und Argumentationsraums — einer Form kollektiven Groupthink, die nicht durch soziale Dynamiken, sondern durch technische Systeme befördert wird.

Ein zweites Risiko betrifft demokratische Urteilsbildung. Deliberative Demokratie setzt voraus, dass Bürgerinnen und Bürger in der Lage sind, Informationen eigenständig zu prüfen, Argumente abzuwägen und politische Urteile auf Grundlage eigener Reflexion zu bilden. Wird dieser Prozess zunehmend durch KI-generierte Meinungsgerüste ersetzt oder vorgeprägt, entsteht eine neue Form epistemischer Abhängigkeit — nicht auf individueller, sondern auf gesellschaftlicher Ebene. Die Fähigkeit zur kritischen Kognition wird damit zu einer demokratietheoretisch relevanten Frage.

ChangeCockpit Coaching & Consulting

Journalistische Arbeit besteht nicht allein aus der Produktion von Texten. Sie umfasst Quellenbewertung, Kontextualisierung, Erkennen von Widersprüchen, Prüfung von Interessenlagen und die Einordnung gesellschaftlicher Machtverhältnisse. Generative KI kann viele dieser Prozesse sprachlich simulieren, ersetzt jedoch nicht automatisch deren erkenntnistheoretische Qualität.

Gerade darin liegt eine zentrale Gefahr. Sprachmodelle erzeugen plausible Ausgaben, ohne über tatsächliches Verständnis oder eigenes Wahrheitsbewusstsein zu verfügen. Die Systeme berechnen sprachliche Wahrscheinlichkeiten und können dabei faktisch inkorrekte oder erfundene Aussagen

produzieren. OpenAI beschreibt solche “Halluzinationen” als plausible, aber falsche Aussagen, die selbst bei scheinbar einfachen Fragen auftreten können.

Quelle: https://openai.com/index/why-language-models-hallucinate/

Für Journalismus und Wissenschaft besitzt dies erhebliche Bedeutung. Wenn KI-generierte Inhalte ungeprüft übernommen werden, verschiebt sich die Verantwortung für Prüfung und Einordnung zunehmend vom aktiven Recherchierenden hin zur Maschine. Gleichzeitig wird die massenhafte Produktion oberflächlicher Inhalte durch generative Systeme erheblich erleichtert.

Mehrere Medienforscher warnen deshalb vor einer möglichen Industrialisierung synthetischer Information. Gemeint ist eine digitale Umgebung, in der große Mengen sprachlich überzeugender, aber inhaltlich unsicherer Inhalte entstehen. In solchen Systemen wird kritische Quellenprüfung nicht weniger, sondern wichtiger.

Hinzu kommt ein weiterer Aspekt: Generative KI verändert nicht nur die Produktion von Wissen, sondern auch die Beziehung des Menschen zu Autorität und Glaubwürdigkeit von Informationen. Historisch wurden Wissen und Expertise mit menschlicher Erfahrung, Ausbildung und institutioneller Glaubwürdigkeit verbunden. Sprachmodelle erzeugen dagegen den Eindruck kompetenter Autorität unabhängig von tatsächlicher Verifikation.

Internationale Bildungsorganisationen betrachten AI Literacy inzwischen zunehmend als grundlegende Zukunftskompetenz. Gemeint ist nicht allein die technische Nutzung von KI-Systemen, sondern die Fähigkeit, maschinell erzeugte Inhalte kritisch zu prüfen, Unsicherheiten zu erkennen und die Grenzen algorithmischer Systeme einschätzen zu können.

Diese Forderung nach AI Literacy gewinnt zusätzliche Dringlichkeit, wenn man die globale Verteilung von KI-Zugang und Medienkompetenz in den Blick nimmt. Generative KI wird gegenwärtig weder technisch noch bildungspolitisch gleichmäßig zugänglich entwickelt. Hochwertige Sprachmodelle, leistungsfähige Infrastrukturen und kritisch-reflexive Nutzungskompetenzen konzentrieren sich bislang überwiegend in wirtschaftlich starken Ländern und privilegierten Bevölkerungsgruppen. Damit entsteht eine neue Dimension digitaler Ungleichheit — nicht mehr allein im Sinne unterschiedlichen Informationszugangs, sondern im Sinne unterschiedlicher kognitiver Werkzeuge und der Fähigkeit, diese kritisch einzusetzen.

Dabei sind zwei gegenläufige Risiken zu unterscheiden. Einerseits könnten Länder und Bevölkerungsgruppen mit gutem Zugang zu leistungsfähigen KI-Systemen Wissens- und Produktivitätsvorsprünge entwickeln, die bestehende globale Ungleichheiten verschärfen. Andererseits besteht in Kontexten mit schwach ausgeprägter Medien- und AI Literacy die Gefahr, dass generative Systeme unkritischer und unreflektierter eingesetzt werden — mit entsprechend größeren Risiken für epistemische Autonomie und Bildungsqualität. UNESCO und OECD betonen in diesem Zusammenhang, dass eine menschenzentrierte KI-Integration globale Gerechtigkeitsfragen nicht ausblenden darf. Die kognitive Dimension generativer KI ist damit zugleich eine verteilungspolitische.

Im OECD *Digital Education Outlook 2026* wird in diesem Zusammenhang auf ein mögliches „AI learning paradox“ hingewiesen: Generative KI kann kurzfristig bessere Leistungen und effizientere Ergebnisse

ermöglichen, zugleich jedoch den tatsächlichen Kompetenzerwerb abschwächen, wenn zentrale Denk- und Problemlösungsprozesse nicht mehr eigenständig durchlaufen werden. Diskutiert wird dabei auch die Gefahr einer „false mastery“ — also des Eindrucks hoher Kompetenz bei gleichzeitig begrenztem eigenständigen Verständnis.

Quelle: https://www.oecd.org/content/dam/oecd/en/publications/reports/2026/01/oecd-digital-education-outlook-2026_940e0dd8/062a7394-en.pdf

Das Europäische Parlament griff diese Problematik 2026 in einem wissenschaftlichen Briefing zur Rolle künstlicher Intelligenz im Klassenzimmer auf. Hervorgehoben wurden kognitive Dimensionen wie reasoning, memory consolidation, metacognitive regulation, attention und epistemic judgement. Das Briefing warnt davor, digitale Bedienkompetenz mit einer unbeeinträchtigten kognitiven Entwicklung gleichzusetzen.

Quelle: https://www.europarl.europa.eu/thinktank/en/document/IUST_BRI%282026%29784575

Die gegenwärtige Forschungslage spricht daher weder für eine generelle kognitive Verarmung noch für uneingeschränkte technologische Optimierung. Vielmehr deutet die Evidenz darauf hin, dass die Auswirkungen generativer KI wesentlich von Bildungsumgebung, Nutzungsintensität, Aufgabenformat und Grad aktiver Beteiligung abhängen.

Die gesellschaftliche Relevanz dieser Entwicklung reicht weit über individuelle Lernprozesse hinaus. Demokratien sind auf Bürger angewiesen, die Informationen prüfen, Widersprüche erkennen und komplexe Zusammenhänge eigenständig bewerten können. Wenn zentrale Mechanismen kritischer Kognition zunehmend automatisiert werden, entsteht nicht nur eine technologische, sondern auch eine gesellschaftliche Herausforderung.

Gleichzeitig wäre es wissenschaftlich unangemessen, generative KI ausschließlich als Bedrohung zu interpretieren. KI-Systeme können Zugang zu Wissen erleichtern, Lernbarrieren reduzieren und kreative Prozesse unterstützen. Besonders für Menschen mit sprachlichen, organisatorischen oder kognitiven Einschränkungen eröffnen sich erhebliche Potenziale.

Die entscheidende Frage lautet daher nicht, ob künstliche Intelligenz in Bildung und Wissensarbeit eingesetzt werden sollte, sondern unter welchen Bedingungen sie menschliche Urteilskraft stärkt statt ersetzt. Genau an dieser Grenze entscheidet sich, ob generative KI zu einem Werkzeug intellektueller Erweiterung wird - oder zu einem System schleichender kognitiver Abhängigkeit.

ChangeCockpit
Coaching & Consulting

Schlussbetrachtung: Zwischen kognitiver Erweiterung und geistiger Abhängigkeit

Die gegenwärtige Debatte über künstliche Intelligenz ist von starken Gegensätzen geprägt. Auf der einen Seite stehen technologische Fortschrittserzählungen, die generative KI als Beginn einer neuen Ära menschlicher Produktivität betrachten. Auf der anderen Seite dominieren Warnungen vor geistiger Verarmung, sinkender Konzentrationsfähigkeit und dem Verlust eigenständigen Denkens. Die wissenschaftliche Evidenz der vergangenen Jahre spricht jedoch weder für uneingeschränkten Optimismus noch für vereinfachende Warnszenarien. Vielmehr zeichnet sich ein deutlich komplexeres Bild ab.

Für die Annahme, generative künstliche Intelligenz verursache unmittelbare neurologische Schäden am menschlichen Gehirn, existieren gegenwärtig keine belastbaren wissenschaftlichen Langzeitdaten. Weder Neurowissenschaften noch Kognitionspsychologie liefern bislang Hinweise auf eine pauschale Degeneration durch KI-Nutzung. Zahlreiche öffentliche Aussagen über eine angebliche Verdummung der Gesellschaft überschreiten daher die empirische Datenlage.

Gleichzeitig wäre es wissenschaftlich ebenso unhaltbar, die Auswirkungen generativer KI auf menschliche Kognition zu unterschätzen. Die vorliegenden Studien deuten darauf hin, dass Sprachmodelle Bedingungen geistiger Aktivität verändern können. Prozesse wie Schreiben, Recherchieren, Strukturieren und Problemlösen lassen sich heute in einem Umfang externalisieren, der historisch neu ist. Damit verschiebt sich nicht nur die Effizienz menschlicher Arbeit, sondern potenziell auch die Art, wie Wissen verarbeitet und Denken organisiert wird.

Besonders relevant erscheint dabei das Prinzip des *Cognitive Offloading*. Das Gehirn neigt grundsätzlich dazu, kognitive Belastung zu reduzieren und externe Systeme zur Entlastung zu nutzen. Generative KI erweitert dieses Prinzip auf komplexe sprachliche und analytische Prozesse. Die Technologie liefert nicht nur Informationen, sondern erzeugt bereits vorstrukturierte Denkprodukte. Genau dadurch kann die Notwendigkeit aktiver mentaler Verarbeitung sinken.

Neurowissenschaftlich ist dies bedeutsam, weil Lernen und kognitive Stabilisierung an Nutzung gekoppelt bleiben. Aufmerksamkeit, Gedächtnisbildung und analytisches Denken beruhen auf wiederholter Aktivierung neuronaler Netzwerke. Werden bestimmte geistige Prozesse dauerhaft ausgelagert, reduziert sich potenziell auch deren Training. Besonders betroffen sind Fähigkeiten, die intensive kognitive Beteiligung erfordern: kritische Reflexion, tiefes Lesen, eigenständige Argumentation und komplexe Problemlösung.

Die aktuelle Forschungslage deutet allerdings darauf hin, dass nicht die Existenz künstlicher Intelligenz selbst entscheidend ist, sondern die Form ihrer Nutzung. Systeme, die Menschen aktiv in Denk- und Lernprozesse einbinden, können kognitive Leistungen unterstützen und erweitern. KI-gestützte Tutorensysteme, adaptive Lernumgebungen oder assistive Technologien zeigen, dass generative Modelle unter bestimmten Bedingungen pädagogische und kreative Potenziale besitzen.

Problematisch wird KI dort, wo sie nicht ergänzt, sondern ersetzt. Wenn Sprachmodelle dauerhaft die Rolle eigener Analyse, Formulierung oder Quellenprüfung übernehmen, könnte langfristig eine Verschiebung kognitiver Gewohnheiten entstehen. Besonders relevant ist dies in Bildung, Wissenschaft und Journalismus — also in Bereichen, deren gesellschaftliche Funktion wesentlich auf kritischer Prüfung und verantwortungsvollem Umgang mit Wissen beruht.

Gerade im Journalismus zeigt sich die Ambivalenz generativer KI besonders deutlich. Sprachmodelle können Rechercheprozesse beschleunigen, Daten strukturieren und Arbeitsabläufe effizienter gestalten. Gleichzeitig erzeugen sie sprachlich plausible Inhalte ohne eigenes Wahrheitsverständnis. Damit wächst die Gefahr einer Informationsumgebung, in der sprachliche Überzeugungskraft zunehmend von faktischer

ChangeCockpit Coaching & Consulting

Verlässlichkeit entkoppelt wird. Die Fähigkeit zur Quellenkritik wird dadurch nicht obsolet, sondern wichtiger denn je.

Zudem lässt sich gegenwärtig häufig nur schwer unterscheiden, welche Veränderungen spezifisch durch generative künstliche Intelligenz entstehen und welche bereits Teil allgemeiner digitaler Mediennutzung sind. Zahlreiche beobachtete Veränderungen - etwa fragmentierte Informationsverarbeitung oder verstärkte kognitive Auslagerung - wurden bereits lange vor dem Aufstieg generativer Sprachmodelle im Zusammenhang mit Suchmaschinen, sozialen Medien und permanenter Smartphone-Nutzung diskutiert. Die isolierte wissenschaftliche Untersuchung spezifischer KI-Effekte steht daher noch am Anfang.

Regulatorisch sind erste Antworten bereits erkennbar. Mit dem EU AI Act, der 2024 in Kraft trat, verabschiedete die Europäische Union das weltweit erste umfassende Regelwerk zur Klassifikation und Regulierung künstlicher Intelligenz nach Risikoklassen. KI-Systeme, die in Bildungskontexten eingesetzt werden, unterliegen darin erhöhten Anforderungen an Transparenz, Nachvollziehbarkeit und menschliche Aufsicht. Der Rechtsrahmen allein beantwortet jedoch keine pädagogischen Fragen: Wie KI-Nutzung in Lernprozesse integriert werden sollte, bleibt eine bildungswissenschaftliche und didaktische Aufgabe.

Einige Länder haben hierzu bereits konkrete curriculare Antworten entwickelt. Singapur integriert AI Literacy seit mehreren Jahren systematisch in nationale Bildungsprogramme und verbindet technische Nutzungskompetenz mit kritischer Reflexion algorithmischer Systeme. Finnland verfolgt mit dem Programm *Elements of AI* einen breit angelegten Ansatz zur Grundbildung in künstlicher Intelligenz, der sich an die gesamte Bevölkerung richtet. Diese Beispiele zeigen, dass eine reflexive Integration generativer KI in Bildungssysteme möglich ist — sie erfordert jedoch gezielte politische Entscheidungen und didaktische Entwicklungsarbeit, die über den bloßen Einsatz von Technologie hinausgehen.

Gleichzeitig ist der empirische Forschungsstand in zentralen Fragen noch erheblich lückenhaft. Belastbare Langzeitstudien zu den neurokognitiven Auswirkungen dauerhafter KI-Nutzung fehlen bislang weitgehend. Kaum untersucht sind zudem differentielle Effekte nach Alter, Vorwissen und Nutzungskontext sowie die Frage, wie sich kollektive Wissensproduktion in Wissenschaft und Öffentlichkeit unter Bedingungen massenhafter KI-Nutzung verändert. Die gegenwärtige Debatte bewegt sich damit zwangsläufig in einem Raum zwischen plausiblen theoretischen Mechanismen und noch ausstehender empirischer Klärung — ein Umstand, den sowohl öffentliche Diskussion als auch bildungspolitische Entscheidungen stärker berücksichtigen sollten.

Vor diesem Hintergrund — regulatorischer Ansätze, internationaler Beispiele und noch offener Forschungsfragen — lässt sich die gesellschaftliche Herausforderung der kommenden Jahre präziser bestimmen.

Die gesellschaftliche Herausforderung der kommenden Jahre wird daher nicht primär technischer Natur sein. Entscheidend ist vielmehr die Frage, welche Formen menschlicher Kognition weiterhin aktiv kultiviert werden. Demokratien, Wissenschaftssysteme und Bildungseinrichtungen beruhen auf Menschen, die Informationen eigenständig prüfen, Unsicherheit tolerieren und komplexe Zusammenhänge kritisch reflektieren können. Genau diese Fähigkeiten entstehen nicht durch passive Rezeption, sondern durch aktive geistige Beteiligung.

ChangeCockpit Coaching & Consulting

Die vorliegenden wissenschaftlichen Befunde deuten darauf hin, dass generative künstliche Intelligenz weniger das Gehirn selbst verändert als die Bedingungen, unter denen menschliches Denken trainiert wird. KI kann geistige Leistungsfähigkeit erweitern, wenn sie analytische Prozesse unterstützt. Sie kann jedoch ebenso kognitive Passivität fördern, wenn sie eigenständige Denkarbeit dauerhaft ersetzt.

Ob künstliche Intelligenz langfristig zu einer Erweiterung menschlicher Kognition oder zu wachsender geistiger Abhängigkeit führt, ist daher keine ausschließlich technische Frage. Entscheidend bleibt, welche Rolle der Mensch innerhalb dieser Systeme weiterhin selbst einnimmt - als aktiver Denkender oder als bloßer Empfänger maschinell erzeugter Antworten.

ChangeCockpit
Coaching & Consulting