

Generative KI an Schulen

Eine Studie über die Nutzung generativer KI aus Sicht von Schülerinnen und Schülern unter Berücksichtigung handlungsleitender Eigenschaften und ausgewählter sozialer Kontextfaktoren

Süße, Thomas & Kobert, Maria

Hochschule Bielefeld (HSBI)

Version 1 – Deutsch*

November 2023

English Version: <https://zenodo.org/doi/10.5281/zenodo.10210254>

Besonderer Dank

Diese Studie über die Nutzung generativer KI an Schulen wäre ohne die besondere Unterstützung von Schulleitungen, Lehrkräften und SchülerInnen der teilnehmenden Schulen nicht möglich gewesen.

Vielen Dank an Sie alle!

Für Fragen, Feedback und Empfehlungen wenden Sie sich bitte direkt an das Autorenteam.

Prof. Dr. Thomas Süße
Hochschule für angewandte Wissenschaft und Kunst
Bielefeld
Langer Weg 9a, 33332 Gütersloh, Deutschland
E-Mail: thomas.suesse@hsbi.de

Dr. Maria Kobert
Hochschule für angewandte Wissenschaft und Kunst
Bielefeld
Langer Weg 9a, 33332 Gütersloh, Deutschland
E-Mail: maria.kobert@hsbi.de

Zusammenfassung

Die vorliegende Studie gibt Aufschluss über die Nutzung generativer KI (z.B. ChatGPT) an ausgewählten Schulen in Deutschland. Hierbei steht die Perspektive von SchülerInnen im Vordergrund. Ziel der Studie ist es, zu untersuchen, ob, in welchem Umfang und für welche Aufgaben SchülerInnen generative KI einsetzen. Darüber hinaus wird untersucht, inwiefern Zusammenhänge zwischen der Intensität der Nutzung generativer KI und spezifischen handlungsleitenden Eigenschaften der SchülerInnen bestehen. Dabei konzentrieren wir uns auf eine Auswahl handlungsleitender Eigenschaften, die in digitalisierten Bildungsumgebungen als relevant angesehen werden. Dies sind z.B. soziale Wahrnehmung von KI, kognitives Engagement, KI-bezogene Selbstwirksamkeitserwartung, Technostress und Technikbereitschaft. Wir analysieren auch, inwiefern kontextuelle Faktoren wie die soziale Unterstützung seitens der Schule und Erziehungsberechtigter sowie Facetten des sozialen Status der Erziehungsberechtigten bei der KI-Nutzung durch SchülerInnen eine Rolle spielen. Zudem untersuchen wir den Zusammenhang zwischen KI-Nutzung und selbst eingeschätztem KI-bezogenem Lernerfolg der SchülerInnen. Abschließend analysieren wir wie sich die KI-Nutzung, die wir anhand von 31 Fragen erheben, in Nutzungsarten bzw. Typen der Nutzung auf einer höheren Ebene gruppieren und konsolidieren lässt.

Die quantitative Analyse wird auf Basis eines Datensatzes mit N=226 durchgeführt, der durch eine Online-Umfrage zwischen März 2023 und Juli 2023 unter SchülerInnen im Alter von 15 bis 19 Jahren erhoben wurde. Die Ergebnisse zeigen, dass generative KI nicht für alle Aufgaben in der Schule gleich häufig eingesetzt wird. Es ist davon auszugehen, dass generative KI bei der Bearbeitung von Hausaufgaben, beim Schreiben und bei der Unterstützung kreativer Prozesse wie Brainstorming und Recherche zunehmend an Bedeutung gewinnen wird. Die KI-bezogene Selbstwirksamkeitserwartung der SchülerInnen, z.B. mit dieser Technologie etwas Sinnvolles tun zu können, scheint hier eine wichtige Rolle zu spielen. Gleichzeitig zeigt sich, dass die Wahrnehmung von technologiebedingtem Stress durch die SchülerInnen bei der Verwendung von generativer KI ebenfalls von Bedeutung ist. Die Ergebnisse zeigen auch, dass die soziale Unterstützung durch Bildungseinrichtungen und seitens Erziehungsberechtigter eine wichtige Rolle bei der Nutzung von generativer KI spielen. Ein höheres Bildungsniveau und ein höherer Akademisierungsgrad im familiären Umfeld gehen teilweise mit einer geringer ausgeprägten Nutzung von generativer KI einher. Die Studie zeigt dies insbesondere in weniger schulbezogenen Bereichen der Nutzung generativer KI. Die Art der sozialen Wahrnehmung generativer KI Tools durch SchülerInnen, insbesondere die Wahrnehmung als teilweise menschenähnlich (anthropomorph), scheint bei der Nutzung dieser Tools relevant zu sein. Die Nutzungshäufigkeit generativer KI korreliert negativ mit dem kognitiven Engagement (z.B. der Freude am Nachdenken) und der Technikkompetenzüberzeugung der SchülerInnen. Ein höheres Maß an selbst wahrgenommenem KI-bezogenem Lernerfolg der SchülerInnen ist mit einer häufigeren Nutzung generativer KI verbunden. Angestrebter Lernerfolg scheint daher eine wichtige Rolle bei der Nutzung generativer KI zu spielen. Die von uns erhobenen 31 Arten der Nutzung generativer KI lassen sich in vier übergeordnete Konzepte

eingruppiert, die wir als „Erledigung von Standardaufgaben“, „Erkundung neuer Möglichkeiten“, „Verbesserung der eigenen Arbeitsergebnisse“ und „Anregung kreativen Denkens“ bezeichnen.

Mit dieser Studie über die Nutzung generativer KI aus der Sicht von SchülerInnen möchten wir mit dazu beitragen, ein besseres Verständnis darüber zu erlangen, wie generativer KI heute und in Zukunft das Lernen junger Menschen verändern kann. Im abschließenden Diskussionsteil dieser Studie argumentieren wir, wie die Ergebnisse unserer Analyse einen Beitrag dazu leisten könnten, nachhaltige Ansätze für das Bildungssystem zu entwickeln und junge Menschen vor dem Hintergrund technologischer und gesellschaftlicher Entwicklungen zukunftsorientiert zu befähigen und auszubilden.

Inhaltsverzeichnis

Zusammenfassung	I
Inhaltsverzeichnis	III
1. Einleitung	1
2. Methodik	3
2.1 Nutzung generativer KI	3
2.2 Handlungsleitende Eigenschaften der SchülerInnen	4
2.3 Kontextuelle Faktoren	6
3. Stichprobenmerkmale und deskriptive Statistiken	7
4. Ergebnisse	8
4.1 Häufigkeiten in der Nutzung generativer KI durch SchülerInnen (RQ1)	8
4.2 Zusammenhang: Nutzung generative KI & handlungsleitende Eigenschaften (RQ2)	9
4.3 Zusammenhang: Nutzung generative KI & kontextuelle Faktoren (RQ3)	14
4.4 Zusammenhang: Nutzung generative KI & wahrgenommener Lernerfolg (RQ4)	17
4.5 Gruppierung der Nutzung generativer KI in übergeordnete Konzepte (RQ5)	18
5. Diskussion	20
6. Limitationen	24
Literaturverzeichnis	25

1. Einleitung

Generative KI-Tools wie ChatGPT und andere üben zunehmend Einfluss auf den Bildungssektor aus (Su & Weipeng, 2023). Während diese neuartigen Tools und Technologien als starker Treiber für die Transformation des Bildungssystems angesehen werden können, ist über ihren tatsächlichen Einsatz im Bildungsbereich wenig bekannt. Jüngste Studien haben bereits gezeigt, dass sowohl Lernende als auch Lehrende zunehmend GPT-basierte Tools für Lehr- und Lernprozesse einsetzen (Müller et al., 2023). Allerdings fehlt es an systematischem Wissen darüber, wie Schülerinnen und Schüler an deutschen Schulen generative KI in der Schule oder für die Schule nutzen und wie unterschiedliche Nutzungsweisen generativer KI mit handlungsleitenden Eigenschaften der Schülerinnen und Schüler, z.B. hinsichtlich ihrer Bereitschaft diese Technologien zu nutzen, zusammenhängen könnten. Darüber hinaus bedarf es weiterer empirisch fundierter Forschung zu den Zusammenhängen zwischen der Nutzung generativer KI und anderen Kontextfaktoren, wie z.B. dem sozialen Status oder der sozialen Unterstützung seitens Bildungseinrichtung und familiärem Umfeld.

Die Forschung im Bereich der digitalen Transformation hat bereits gezeigt, dass Technikbereitschaft von NutzerInnen, die aus den drei Dimensionen Technikakzeptanz-, Technikkompetenz- und Technikkontrollüberzeugung bestehen kann, für die Art und Weise der Nutzung bzw. Ablehnung neuer Technologien besonders relevant sein kann (Davis, 1989; Neyer & Felber, 2012).

Zudem wird davon ausgegangen, dass die Art und Weise wie KI-Tools wahrgenommen werden, z.B. als sozial, kooperativ und anthropomorph, im Zusammenhang mit der Nutzung dieser Technologie steht (Mandl et al., 2022, 2023). Da generative KI oft als ein Werkzeug wahrgenommen wird, welches das kognitive Engagement der NutzerInnen, wie z. B. kritisches Denken oder Freude am Nachdenken, beeinflussen kann (z.B. Hahn & Lee, 2017; Yoon & Lee, 2021), finden wir es aufschlussreich zu untersuchen, wie das Bedürfnis der SchülerInnen nach Kognition mit der Nutzung von generativer KI einhergeht. Es wurde bereits vielfach dargelegt, dass auch Selbstwirksamkeitserwartung eine Rolle bei der Nutzung von Technologien spielt (z.B. Ellen et al., 1991; Kulviwat et al., 2014). Daher erscheint es uns wichtig zu untersuchen, welche Rolle KI-bezogene Selbstwirksamkeitserwartung bei der Nutzung generativer KI durch SchülerInnen spielt. Ein weiterer wichtiger Aspekt ist Technostress, der durch die Anforderungen bei der Nutzung digitaler Werkzeuge verursacht werden kann (z.B. Tarafdar et al., 2011). In der vorliegenden Studie untersuchen wir, welchen Zusammenhang es zwischen Technostress und der Nutzung generativer KI durch SchülerInnen gibt.

Zusätzlich zu den oben aufgeführten Aspekten halten wir es ebenfalls für relevant, Aspekte der Bildungsgerechtigkeit zu untersuchen, da generative KI-Tools, ebenso wie andere digitale Technologien, auch zu Ungleichheit in Bezug auf Zukunftschancen junger Menschen beitragen können (OECD, 2018, 2023). In unserer Studie untersuchen wir genauer, wie Facetten des sozialen Status im familiären Umfeld von Schülerinnen und Schülern mit der Nutzung von generativer KI in Lern- und Bildungskontexten zusammenhängen. Darüber hinaus wollen wir ein besseres Verständnis von der Rolle der elterlichen und schulischen Unterstützung in Bezug auf die Nutzung generativer KI gewinnen. Wir sind davon

überzeugt, dass diese kontextuellen Faktoren ebenfalls als entscheidende Aspekte der Technologienutzung betrachtet werden sollten, insbesondere im Hinblick auf SchülerInnen. Schließlich untersuchen wir, inwiefern die Häufigkeit der Nutzung generativer KI einhergeht mit der subjektiven Einschätzung der SchülerInnen, den eigenen Lernprozess und Lernerfolg in der Schule zu verbessern. Ausgehend von unserer Forschungsmotivation und den konzeptionellen Vorüberlegungen ergeben sich fünf zentrale Forschungsfragen:

RQ1: Wie umfangreich ist die Nutzung generativer KI durch SchülerInnen in der Schule und darüber hinaus?

RQ2: Welche Zusammenhänge bestehen zwischen der Nutzung generativer KI und den handlungsleitenden Eigenschaften der SchülerInnen?

RQ3: Welche Zusammenhänge bestehen zwischen der Nutzung generativer KI und kontextuellen Faktoren?

RQ4: Welcher Zusammenhang besteht zwischen der Nutzung generativer KI und dem selbst wahrgenommenen Lernerfolg?

RQ5: Wie kann die Nutzung generativer KI weiter gruppiert oder konsolidiert werden?

In der vorliegenden Studie kombinieren wir ein spezifisches Set von Messinstrumenten und Skalen. Einige Instrumente wurden gemeinsam mit ExpertInnen entwickelt und kritisch reflektiert. Die Mehrzahl der Skalen basiert auf etablierten Messinstrumenten aus der wissenschaftlichen Literatur. Abschnitt 2 enthält detaillierte Informationen über unser methodisches Vorgehen. Die Stichprobenstruktur und die deskriptiven Statistiken werden in Abschnitt 3 vorgestellt. In Abschnitt 4 werden die statistischen Ergebnisse in Bezug auf die fünf zentralen Forschungsfragen einschließlich der Informationen über die Datenanalyseverfahren aufgezeigt. In Abschnitt 5 fassen wir die Ergebnisse zusammen, diskutieren die wichtigsten Resultate und geben einige allgemeine Empfehlungen. Schließlich halten wir es für wichtig, in Abschnitt 6 die Ergebnisse unserer kritischen Reflexion über die Limitationen dieser Studie und den Gesamtansatz darzulegen.

2. Methodik

Die vorliegende Studie ist quantitativ und beruht auf einem Datensatz, der zwischen März und Juli 2023 im Rahmen einer Online-Umfrage unter 15- bis 19-jährigen SchülerInnen an vier Gymnasien in Deutschland erhoben wurde. Die Stichprobengröße beträgt N=226. Um die Forschungsfragen RQ1 bis RQ5 zu beantworten, haben wir Korrelationsanalysen sowie eine explorative Faktorenanalyse (EFA) durchgeführt. Die Variablen und Skalen, die im Rahmen der Analyse verwendet wurden, werden im Folgenden im Detail beschrieben.

2.1 Nutzung generativer KI

Wir haben gemeinsam mit ExpertInnen ein Set von 31 Items erstellt (siehe Tabelle 1), um zu untersuchen, für welche Aufgaben oder mit welchen Zielen und in welcher Intensität SchülerInnen generative KI nutzen. Die 31 Items wurden in einem iterativen Prozess im Rahmen von Workshops und Diskussionen zusammen mit LehrerInnen entwickelt und finalisiert. Die Nutzungsintensität wurde gemessen, indem SchülerInnen gefragt wurden, wie oft sie generative KI für jede der 31 Aufgabentypen nutzen. Die Antworten wurden auf einer 5-stufigen Skala gemessen (1 = nie, 2 = selten, manchmal = 3, oft = 4, sehr oft = 5). Alle 31 Arten der Nutzung generativer KI sind in Tabelle 1 zusammengefasst.

Tabelle 1: Items zur Messung der Intensität der Nutzung generativer KI

Nr.	Nutzungsarten generativer KI	Nr.	Nutzungsarten generativer KI
1	<i>Hausaufgaben erledigen</i>	17	<i>Post schreiben</i>
2	<i>Texte schreiben</i>	18	<i>E-Mail korrigieren</i>
3	<i>Schnell einen Text erstellen</i>	19	<i>Gliederung für einen Text erstellen</i>
4	<i>Übersetzen</i>	20	<i>Sätze vervollständigen</i>
5	<i>Spielen</i>	21	<i>Song Texte erstellen</i>
6	<i>Chatten</i>	22	<i>Mathematikaufgaben lösen</i>
7	<i>Programmieren</i>	23	<i>Social Media Beitrag erstellen</i>
8	<i>Feedback zu einer Lösung anderer geben</i>	24	<i>Eigene Texte sprachlich verbessern</i>
9	<i>Feedback zu meiner eigenen Lösung einholen</i>	25	<i>Quellen recherchieren</i>
10	<i>Ideenfindung unterstützen</i>	26	<i>Literatur finden</i>
11	<i>Texte für Blogs und Foren verfassen</i>	27	<i>Texte befragen</i>
12	<i>Rechtschreibung eigener Texte überprüfen</i>	28	<i>Als Inspirationsquelle nutzen</i>
13	<i>Eigene Texte inhaltlich überarbeiten</i>	29	<i>Als Ratgeber nutzen</i>
14	<i>Recherchieren</i>	30	<i>Chat mit literarischen Figuren oder historischen Personen</i>
15	<i>Fragen von LehrerInnen beantworten</i>	31	<i>Tabellen, Schaubilder und graphische Elemente erzeugen lassen</i>
16	<i>Blog schreiben</i>		

2.2 Handlungsleitende Eigenschaften der SchülerInnen

Soziale Wahrnehmung generativer KI (SOC, COO, ANT)

Die soziale Wahrnehmung generativer KI wird durch eine Anpassung der Skala zur sozialen Wahrnehmung von Robotern (SPRS) (Mandl et al., 2022) gemessen. Die SPRS-Skala ist ein 18 Punkte umfassender Fragebogen zur Bewertung der sozialen Wahrnehmung smarterer Technologien durch Menschen. Die Skala besteht aus den drei Dimensionen sozial (SOC), kooperativ (COO) und anthropomorph (ANT). In dieser Untersuchung messen wir mit diesem Instrument die Stärke der menschlichen Wahrnehmung hinsichtlich der sozialen, kooperativen und anthropomorphen Eigenschaften generative KI-Tools. Wir haben die ursprüngliche SPRS-Skala insofern an generative KI angepasst, dass wir die Frage „bewegt sich geschmeidig“ – „bewegt sich starr“ durch „eloquent“ – „nicht eloquent“ ersetzt haben. Außerdem haben wir die Skala zur besseren Interpretation unserer Ergebnisse umgepolt, sodass eine eins eine niedrige und eine fünf eine hohe Bewertung angibt.

Technostress (ADM, NSM)

Technostress bezeichnet denjenigen Stress, der durch die Unfähigkeit verursacht werden kann, mit den Anforderungen, die im Zusammenhang mit der Nutzung digitaler Technologien entstehen können, umzugehen (Tarafdar et al., 2011). Die Skala, die wir in dieser Studie verwenden, wurde entwickelt, um den Technostress von SchülerInnen zu messen (Wang et al., 2020). Demnach besteht Technostress aus zwei Komponenten. Einem subjektiv wahrgenommenen Defizit eigener Fähigkeiten in Bezug auf gegebene Anforderungen (ADM) sowie einem subjektiv wahrgenommenen Defizit des Angebots in Bezug auf individuelle Bedürfnisse (NSM). Präziser formuliert bezieht sich ADM auf die Diskrepanz zwischen den individuellen Fähigkeiten und Fertigkeiten, die für die Teilnahme an technologiegestütztem Lernen erforderlich sind, und der Zeit und dem Aufwand, die dafür zur Verfügung stehen. NSM bezieht sich auf die Diskrepanz zwischen den Bedürfnissen der SchülerInnen nach verbesserten Lernerfahrungen durch digitale Technologien und der Zufriedenheit mit technologiegestütztem Lernen (Wang et al., 2020). Die Skala besteht aus acht Items, die auf einer 5-stufigen Skala mit 1= „stimme überhaupt nicht zu“, 2= „stimme nicht zu“, 3= „unentschieden“, 4= „stimme zu“, 5= „stimme völlig zu“ gemessen werden.

KI-bezogene Selbstwirksamkeitserwartung (ASE)

Selbstwirksamkeitserwartung ist definiert als die Kompetenzerwartung einer Person, mit Schwierigkeiten und Hindernissen im täglichen Leben umzugehen (Bandura, 1986). Mit KI-bezogener Selbstwirksamkeitserwartung (ASE) kann somit die Kompetenzerwartung der SchülerInnen beschrieben werden, generative KI Tools ziel- und ergebnisorientiert nutzen zu können. Wir messen die KI-bezogene Selbstwirksamkeitserwartung mit einer Adaption der deutschen ASKU (Beierlein et al.,

2014). Bei der Skala handelt es sich um eine 5-stufige Selbsteinschätzungsskala von 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu).

Kognitives Engagement (NFC)

Kognitives Engagement (NFC) beschreibt die individuelle Fähigkeit einer Person, sich mit anspruchsvollen geistigen Aufgaben zu beschäftigen oder diese zu genießen (Cacioppo et al., 1984). Daher soll NFC das Engagement von SchülerInnen für kognitiv anspruchsvolle Aufgaben widerspiegeln. Die NFC-Teens-Skala (Preckel, 2016), die in der vorliegenden Studie verwendet wurde, wurde entwickelt, um NFC für ältere Kinder und Jugendliche (älter als 10 Jahre) zu messen. Sie besteht aus 19 Items und wird auf einer 5-stufigen Likert-Skala mit den Antwortmöglichkeiten 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu) gemessen. Für die Auswertung werden die Punkte aller Antworten addiert. Das Ergebnis kann Werte im Bereich zwischen 19 und 95 annehmen.

Technikbereitschaft (TAC, TCP, TCO)

Diese Skala (Neyer et al., 2016) basiert auf dem Konzept der Technikakzeptanz (Davis, 1989). Die erfolgreiche Nutzung von Technologie hängt von Einstellungen sowie von Kompetenz- und Kontrollüberzeugungen ab (Neyer & Felber, 2012). Die Skala besteht aus zwölf Items, die sich auf die drei Faktoren Technikakzeptanz- (TAC), Technikkompetenz- (TCP) und Technikkontrollüberzeugung (TCO) verteilen. Die Items adressieren die persönliche Einstellung der Nutzerinnen und Nutzer zum Umgang mit moderner Technologie. Damit soll der Erfolg der Nutzung neuer Technologien vorhergesagt werden. Bei dem Instrument handelt es sich um eine 5-stufige Likert-Skala mit den Optionen 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu).

Selbstwahrgenommener KI-bezogener Lernerfolg der SchülerInnen

Bei dieser Skala handelt es sich um ein selbst entwickeltes Bewertungsinstrument, das Aufschluss darüber geben soll, wie SchülerInnen ihren eigenen Lernerfolg in der Schule einschätzen, wenn sie KI einsetzen. Das Instrument besteht aus den 5 Fragen „Ich nutze KI-basierte Textgeneratoren sehr erfolgreich“, „KI-basierte Textgeneratoren helfen mir, meine schulischen Leistungen zu verbessern“, „Das Lernen fällt mir leichter, wenn ich KI-basierte Textgeneratoren nutzen kann“, „Meine Schulnoten haben sich verbessert, seitdem ich KI-basierte Textgeneratoren nutze“ und „Mit Hilfe von KI-Textgeneratoren kann ich meine Schularbeiten schneller und effizienter erledigen“. Die Antworten wurden auf einer 5-stufigen Likert-Skala von 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu) gegeben.

2.3 Kontextuelle Faktoren

Als Kontextfaktoren messen wir die soziale Unterstützung bei der Nutzung generativer KI, die von der Bildungseinrichtung und/oder den Eltern* erbracht wird, sowie Facetten des sozialen Status, die wir durch die Fragen nach der Schulbildung und dem Beruf der Eltern sowie der finanziellen Situation der Eltern erfassen. In dieser Studie verwenden wir ausschließlich aus Gründen einer besseren Lesbarkeit die Begriffe Eltern und Erziehungsberechtigte synonym.

Soziale Unterstützung bei der Nutzung von KI seitens der Bildungseinrichtung (SOE)

Da Unterstützung in Form von Ermutigung durch Lehrkräfte zur Nutzung generativer KI oder die Bereitstellung geeigneter IT-Infrastruktur SchülerInnen dazu ermutigen könnte, generative KI in oder für die Schule zu nutzen, entwickeln wir eine Variable, um zu bewerten, wie hoch SchülerInnen die soziale Unterstützung ihrer Schule bei der Nutzung generativer KI einschätzen. Das Item lautet wie folgt: „In der Schule werde ich bei der Nutzung von generativen KI-Werkzeugen unterstützt.“ Die Antworten werden auf einer 5-stufigen Likert-Skala von 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu) gegeben.

Soziale Unterstützung bei der Nutzung von KI seitens der Eltern (SOP)

Die soziale Unterstützung durch die Eltern zu Hause könnte ebenfalls eine entscheidende Rolle für die Nutzungsintensität generativer KI in der Schule spielen. Wir messen die Unterstützung durch das folgende Item: „Meine Eltern unterstützen mich bei der Nutzung von generativer KI.“ Das Item wird auf einer 5-stufigen Likert-Skala von 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu) gemessen.

Schulisches Bildungsniveau der Eltern (PAS)

Das Bildungsniveau der Eltern wird gemessen, indem die Befragten nach dem höchsten Schulabschluss von Elternteil A und Elternteil B gefragt werden. Die Skala ist 4-stufig mit den Antwortmöglichkeiten 1 = „nicht bekannt“, 2= „kein Schulabschluss“, 3 = „Haupt-/Realschulabschluss“ und 4 = „Fach-, Hochschulreife“. Der Wert für PAS durch Bildung der Summe über das Bildungsniveau von Elternteil A und Elternteil B berechnet, wobei Antwortmöglichkeit 1 vernachlässigt wird, da sie keinen Aufschluss zum Bildungsgrad der Eltern bietet.

Berufliches Niveau der Eltern (PAP)

Das höchste berufliche Niveau der Eltern wird durch die Variable „Berufliche Qualifikation der Eltern“ (PAP) gemessen. Auch hier werden die SchülerInnen nach der höchsten beruflichen Qualifikation von Elternteil A bzw. Elternteil B gefragt. Mögliche Antworten sind 1 = „nicht bekannt“, 2= „kein Berufsabschluss“, 3 = „Berufsausbildung“, 4= „Meister / Techniker“, 5 = „Studium“. Schließlich wird der PAP analog zu PAS berechnet.

Finanzielle Situation der Eltern (PAF)

Die finanzielle Situation im Elternhaus wird durch die Variable „Finanzielle Situation der Eltern“ (PAF) gemessen. Die Frage hier zu lautet: „Wie beurteilen Sie die finanzielle Situation in Ihrem Elternhaus?“ Die Antworten werden auf einer 5-stufigen Skala von 1 (deutlich unterdurchschnittlich) bis 5 (deutlich überdurchschnittlich) erfasst.

3. Stichprobenmerkmale und deskriptive Statistiken

Die Daten wurden im Rahmen einer Online-Umfrage erhoben, die von März bis Juli 2023 in Deutschland durchgeführt wurde. SchülerInnen aus vier Schulen in Deutschland nahmen an der Umfrage teil. Bei drei der Schulen handelte es sich um Gymnasien und bei einer um eine Gesamtschule. Insgesamt nahmen 226 Schülerinnen und Schüler im Alter von 15 bis 19 Jahren aus der Oberstufe (Klassenstufen 10 bis 13) an der Umfrage teil. Deskriptive Statistiken für die Stichprobe finden sich in Tabelle 2.

Tabelle 2: Deskriptive Statistik der Stichprobe

Geschlecht	männlich		weiblich		divers		Keine Angabe		
Anzahl /%	98 / 43.4		117 / 51.8		6 / 2.7 %		5 / 2.1 %		
Alter (Jahre)	15	16	17	18	19	Keine Angabe			
Anzahl /%	54 / 23.9 %	95 / 42.0 %	57 / 25.2 %	19 / 8.4 %	1 / 0.5 %	--			
Klassenstufe	10		11		12		13		Keine Angabe
Anzahl /%	151 / 66.8 %		63 / 27.9 %		10 / 4.4 %		2 / 0.9 %		--
Schulisches Bildungsniveau Eltern	Kein Schulabschluss		Haupt-/ Realschule		Abitur		Keine Angabe		
Anzahl /%	11 / 4.9 %		53 / 23.5 %		150 / 66.4 %		12 / 5.2 %		
Berufliches Niveau Eltern	Kein Berufsabschluss		Ausbildung		Techniker / Meister		Studium		Keine Angabe
Anzahl /%	9 / 4.0 %		56 / 24.8 %		55 / 24.3 %		88 / 38.9 %		18 / 8.0 %
Finanzielle Situation Eltern	deutlich unterdurchschnittlich		unterdurchschnittlich		durchschnittlich		überdurchschnittlich		deutlich überdurchschnittlich
Anzahl /%	3 / 1.3%		18 / 8.0%		104 / 46.0%		81 / 35.8 %		20 / 8.9 %

4. Ergebnisse

4.1 Häufigkeiten in der Nutzung generativer KI durch SchülerInnen (RQ1)

Die SchülerInnen wurden nach der Häufigkeit ihrer individuellen Nutzung generativer KI auf einer 5-stufigen Skala von „nie“ bis „sehr oft“ für alle 31 Items (Tabelle 1) gefragt. Die Nutzungshäufigkeiten sind in Abbildung 1 in Prozent dargestellt. Die Ergebnisse zeigen einige tiefgreifende Unterschiede zwischen den verschiedenen Arten der Nutzung generativer KI auf. In unserer Stichprobe verwenden beispielsweise mehr als 70 % der SchülerInnen generative KI nie für die Programmierung, das Schreiben von Texten für Blogs oder Foren, das Verfassen eines Posts oder eines Beitrags in sozialen Medien, das Erstellen von Song Texten oder für die Literaturrecherche. 82% der SchülerInnen gaben an, dass sie generative KI nie für das Schreiben eines Blogs verwenden würden. Für das Erledigen von Hausaufgaben, das eindeutig eine schulbezogene Aufgabe darstellt, verwenden 37% der Befragten nie generative KI, 23% selten, 28% manchmal, 10% oft und 2% sehr oft. Die Aufgaben Texte schreiben, schnell einen Text erstellen, übersetzen sowie Ideenfindung unterstützen, zeigen eine ähnliche Verteilung auf wie die Erledigung von Hausaufgaben. Zum Recherchieren scheinen sogar 21% der

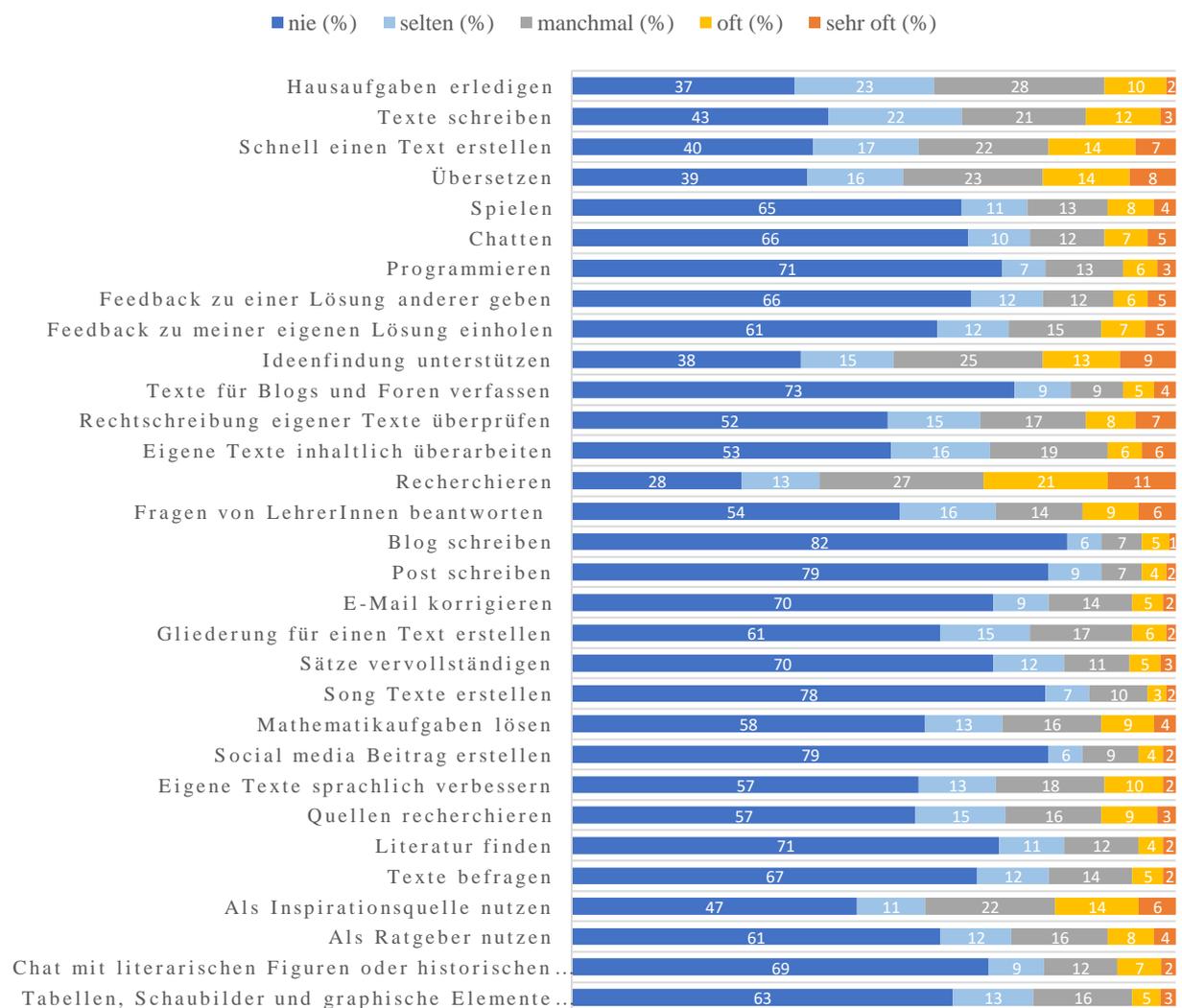


Abbildung 1: Häufigkeiten generativer KI-Nutzung von SchülerInnen

SchülerInnen generative KI häufig und 11% sehr häufig zu nutzen. Zusammenfassend lässt sich feststellen, dass der schulbezogene Einsatz von generativer KI potenziell an Bedeutung zunimmt. Zudem gibt es deutliche Unterschiede in der Häufigkeit zwischen den verschiedenen Arten der Nutzung generativer KI.

4.2 Zusammenhang: Nutzung generative KI & handlungsleitende Eigenschaften (RQ2)

Forschungsfrage 2 konzentriert sich auf die Beziehungen zwischen der Nutzung generativer KI durch SchülerInnen und einer Reihe von handlungsleitenden Eigenschaften. Die individuellen Eigenschaften, die wir in unserer Studie berücksichtigen, sind soziale Wahrnehmung generativer KI, Technikbereitschaft, Technostress, KI-bezogene Selbstwirksamkeitserwartung und kognitives Engagement. Die folgende Tabelle 3 gibt einen Überblick über die Metriken der Skalen, einschließlich Cronbachs Alpha und Mittelwerte. Cronbachs Alpha (Cronbach, 1951) ist ein weit verbreitetes Maß für die Reliabilität der jeweiligen verwendeten Skala. In der Regel wird ein Wert von .7 oder höher als Indikator für eine gute Reliabilität einer Skala angesehen (Taber, 2018).

Tabelle 3: Metriken der Skalen für handlungsleitende Eigenschaften der SchülerInnen

Handlungsl. Eigensch.	Facette	Abkürzung	Min	Max	Mittelwert	Standardabweichung	Cronb. Alpha
Soziale Wahrnehmung generativer KI	sozial	SOC	0.0	5.0	3.24	.75	.84
	kooperativ	COO	0.0	5.0	3.56	.83	.78
	anthropomorph	ANT	0.0	5.0	2.55	.72	.83
Technostress	Anforderungen übersteigen Fähigkeiten	ADM	1.0	4.0	2.27	.76	.84
	Bedürfnisse größer als Angebot	NSM	1.0	4.43	2.37	.74	.80
KI-bezogene Selbstwirksamkeitserwartung		ASE	1.0	5.0	3.10	1.06	.87
Kognitives Engagement		NFC	38.0	72.0	57.81	5.00	.84
Technikbereitschaft	Technikakzeptanz	TAC	4.0	20.0	12.6	3.41	.76
	Technikkompetenz	TCP	6.0	20.0	16.35	3.43	.85
	Technikkontrolle	TCO	5.0	20.0	13.43	3.01	.72

Um die Beziehungen zwischen den handlungsleitenden Eigenschaften der SchülerInnen und der Nutzung generativer KI zu untersuchen, testen wir auf bidirektionale Korrelationen zwischen diesen Variablen (Spearman-Korrelationen). Die Korrelationswerte sind in Tabelle 4 dargestellt. Die Ergebnisse zeigen einige positive Korrelationen zwischen der sozialen Wahrnehmung generativer KI (SOC), d.h. zum Beispiel der Wahrnehmung von KI als gesellig, höflich oder zuvorkommend (Mandl et al., 2023) und der Nutzung für das Erledigen von Hausaufgaben, das Schreiben von Texten und zur Unterstützung der Ideenfindung. Gleichzeitig gibt es negative Korrelationen zum Schreiben eines Posts, dem Erstellen eines Beitrags in sozialen Medien und dem Chatten mit literarischen Figuren.

Für die kooperative Wahrnehmung (COO) von generativer KI, d.h. die Wahrnehmung von KI als z.B. fleißig, aktivierend, selbstlos (Mandl et al., 2023) gibt es ebenfalls positive Korrelationen mit dem

Erledigen von Hausaufgaben und dem Schreiben von Texten. Es gibt negative Korrelationen zwischen COO und der Nutzung zum Chatten, zum Schreiben von Blogs oder Posts, zum Korrigieren von E-Mails, zum Erstellen von Songtexten, zum Erstellen von Social-Media-Posts, zur Literaturrecherche und zum Chatten mit literarischen oder historischen Figuren.

Hinsichtlich der Wahrnehmung generativer KI durch die SchülerInnen als anthropomorph (ANT), z. B. als real, warm oder organisch (Mandl et al., 2023), ist aus Tabelle 4 ersichtlich, dass 30 der 31 verschiedenen Verwendungszwecke generativer KI eine positive Korrelation aufweisen, wobei die meisten Korrelationswerte hoch signifikant sind. Die höchsten positiven Korrelationen zeigen sich bei der Nutzung für die Erledigung von Hausaufgaben, die Einholung von Ratschlägen, das Schreiben von Texten oder die Erstellung von Tabellen, Diagrammen und Grafiken. Das Übersetzen weist keine signifikante Korrelation auf und es gibt keine negativen Korrelationen für die Variable ANT.

Die erste Facette von Technostress, ausgelöst durch das subjektive Empfinden, dass die (technischen) Anforderungen die eigenen Fähigkeiten übersteigen (ADM), bezieht sich auf die Wahrnehmung oder das Gefühl eines Schülers/einer Schülerin, dass er/sie den Anforderungen des technologiegestützten Lernens auf der Grundlage seiner/ihrer eigenen Fähigkeiten, Fertigkeiten und seines/ihrer Zeit- und Arbeitsaufwands nicht gerecht werden kann (Wang et al., 2020, S. 99) zeigt eine Reihe von positiven Korrelationen mit 23 der 31 Möglichkeiten, generative KI zu nutzen. Die stärksten und hochsignifikanten Korrelationen bestehen zwischen ADM und dem Erstellen eines Beitrags in sozialen Medien, dem Schreiben eines Blogs, dem Verfassen eines Posts und dem Schreiben von Texten für Blogs oder Foren. Es gibt keine signifikanten negativen Korrelationen.

Die zweite Subdimension von Technostress, die berücksichtigt, dass individuelle Bedürfnisse nicht adäquat durch das technische Angebot adressiert werden, spiegelt unter anderem wider, dass technologiegestütztes Lernen nicht den Bedürfnissen und Präferenzen der SchülerInnen entspricht (Wang et al., 2020, S. 99). Wie aus Tabelle 4 hervorgeht, gibt es positive Korrelationen zwischen NSM und 12 von 31 Arten der Nutzung generativer KI. Die stärksten positiven Korrelationen zeigen sich zwischen dem Schreiben eines Blogs, dem Schreiben eines Posts, dem Schreiben von Texten für Blogs und Foren und dem Chatten. Wie für ADM gibt es auch für NSM keine signifikanten negativen Korrelationen.

Hinsichtlich der KI-bezogenen Selbstwirksamkeitserwartung (ASE) in unserer Stichprobe gibt es starke und hochsignifikante Korrelationen mit dem Erledigen von Hausaufgaben, dem schnellen Erstellen eines Textes, dem Schreiben von Texten und der Recherche. Signifikante und hochsignifikante Korrelationen finden sich zwischen 21 der 31 verschiedenen Arten, wie SchülerInnen generative KI nutzen, und der ASE. In unserer Stichprobe wurden keine signifikanten negativen Korrelationen gefunden. Insgesamt besteht die stärkste positive und hochsignifikante Korrelation von .63** zwischen ASE und dem Erledigen von Hausaufgaben.

Das kognitive Engagement der SchülerInnen (NFC), das z.B. Freude an anstrengenden kognitiven Bemühungen darstellt (Cacioppo et al., 1984, S. 306), korreliert signifikant negativ mit einer Reihe von

Nutzungsarten generativer KI. Es lassen sich einige Korrelationen aufzeigen, bei denen sich die Nutzung generativer KI nicht nur auf schulische Aufgaben, sondern auch auf außerschulische Handlungsfelder,

Tabelle 4: Korrelation zwischen den handlungsleitenden Eigenschaften der SchülerInnen und der Nutzung generativer KI

Nutzung generativer KI	Handlungsleitende Eigenschaften der SchülerInnen									
	SOC	COO	ANT	ADM	NSM	ASE	NFC	TAC	TCP	TCO
Hausaufgaben erledigen	.19**	.15*	.38**	.04	.00	.63**	-.22**	.08	-.02	-.01
Texte schreiben	.17*	.15*	.35**	.09	.03	.55**	-.22**	.16*	-.06	.03
Schnell einen Text erstellen	.11	.09	.29**	.07	.06	.58**	-.25**	.24**	-.05	.03
Übersetzen	-.11	-.06	.03	.16*	.09	.12	-.23**	-.02	-.21**	-.19**
Spielen	-.11	-.07	.16*	.26**	.17*	-.02	-.19**	.05	-.33**	-.09
Chatten	-.11	-.16*	.20**	.37**	.30**	.08	-.26**	.013	-.35**	-.12
Programmieren	.04	.00	.31**	.15*	.09	.19**	-.04	.20**	-.21**	.02
Feedback zu einer Lösung anderer geben	-.04	-.14	.31**	.35**	.24**	.19**	-.17*	.14	-.34**	-.12
Feedback zu meiner eigenen Lösung einholen	.05	-.06	.32**	.23**	.12	.23**	-.07	.23**	-.29**	-.01
Ideenfindung unterstützen	.18*	.12	.29**	-.01	-.10	.50**	-.06	.28**	-.04	.10
Texte für Blogs und Foren verfassen	-.06	-.10	.28**	.38**	.31**	.15*	-.22**	.04	-.39**	-.14*
Rechtschreibung eigener Texte überprüfen	.08	-.02	.26**	.19**	.06	.18*	-.05	.05	-.22**	-.03
Eigene Texte inhaltlich überarbeiten	.13	.06	.27**	.13	.04	.34**	-.09	.19**	-.18*	-.04
Recherchieren	.12	.11	.18*	.03	-.02	.52**	-.13	.11	-.06	.00
Fragen von LehrerInnen beantworten	.000	-.03	.27**	.12	.06	.34**	-.23**	.18*	-.15*	-.03
Blog schreiben	-.09	-.28**	.25**	.43**	.37**	.03	-.28**	-.04	-.39**	-.19**
Post schreiben	-.16*	-.29**	.23**	.42**	.32**	.02	-.23**	.02	-.41**	-.18*
E-Mail korrigieren	-.10	-.19*	.24**	.26**	.18*	.09	-.14*	.02	-.35**	-.16*
Gliederung für einen Text erstellen	.00	-.08	.28**	.19**	.12	.34**	-.19**	.14	-.22**	-.12
Sätze vervollständigen	-.07	-.10	.30**	.31**	.20**	.18*	-.23**	.07	-.30**	-.12
Song Texte erstellen	-.13	-.23**	.32**	.36**	.29**	.03	-.14*	-.01	-.36**	-.16*
Mathematikaufgaben lösen	.00	-.06	.29**	.26**	.17*	.29**	-.24**	.09	-.29**	-.06
Social Media Beitrag erstellen	-.18*	-.28**	.25**	.47**	.28**	.02	-.24**	-.04	-.35**	-.21**
Eigene Texte sprachlich verbessern	.08	.07	.34**	.17*	.11	.36**	-.10	.16*	-.22**	-.01
Quellen recherchieren	-.06	-.10	.22**	.18*	.11	.26**	-.20**	-.02	-.17*	-.26**
Literatur finden	-.011	-.26**	.20**	.27**	.17*	.11	-.15*	.03	-.32**	-.22**
Texte befragen	-.03	-.14	.29**	.19**	.14	.26**	-.13	.20**	-.22**	-.06
Als Inspirationsquelle nutzen	-.02	-.06	.24**	.07	.04	.39**	-.11	.15*	-.03	.05
Als Ratgeber nutzen	.01	-.07	.37**	.15*	.12	.32**	-.16*	.22**	-.16*	-.07
Chat mit literarischen Figuren oder historischen Personen	-.17*	-.20**	.27**	.31**	.24**	.07	-.19**	.06	-.32**	-.08
Tab., Schaub. und graphische Elemente erzeugen lassen	.04	-.04	.35**	.18*	.11	.23**	-.10	.19**	-.18*	-.04

SOC = sozial, COO = kooperativ, ANT = anthropomorph, ADM = Anforderungen übersteigen Fähigkeiten, NSM = Bedürfnisse größer als Angebot, ASE = KI-bezogene Selbstwirksamkeitserwartung, NFC = Kognitives Engagement, TAC = Technikakzeptanz, TCP = Technikkompetenz, TCO = Technikkontrollüberzeugung *: Signifikanzniveau $p < 0.05$ **: Signifikanzniveau $p < 0.001$

wie dem Schreiben eines Blogs oder eines Posts, beziehen. Gleichwohl existieren auch negative Korrelationen, speziell für schulische Aufgaben, wie z. B. schnelles Schreiben von Texten, Übersetzen, Erledigen von Hausaufgaben oder Lösen von Matheaufgaben. Es gibt keine signifikanten positiven Korrelationen zwischen NFC und der Nutzung von generativer KI in unserer Stichprobe.

Für die Technikbereitschaft der SchülerInnen, die mit den drei Komponenten Technologieakzeptanz (TAC), Technologiekompetenz (TCP) und Technologiekontrolle (TCO) zusammenhängt, gibt es in unserer Stichprobe sowohl signifikante positive (TAC) als auch negative (TCP & TCO) Korrelationen. TAC ist signifikant positiv korreliert mit dem Schreiben von Texten, dem schnellen Erstellen eines Textes, dem Programmieren, dem Einholen von Feedback zur eigenen Lösung, dem Unterstützen der Ideenfindung, dem Überarbeiten des Inhalts eigener Texte, dem Beantworten von Fragen von LehrerInnen, dem Verbessern der Sprache eigener Texte, der Befragung von Texten, dem Einholen von Inspiration und Ratschlägen und dem Erstellen von Tabellen, Diagrammen oder Abbildungen. Die Korrelationswerte reichen von .15 bis .28, was im Vergleich zu anderen in der Studie untersuchten Variablen nicht sehr hoch ist. Es gibt eine Reihe signifikanter negativer Korrelationen zwischen TCP und den meisten Arten der generativen KI-Nutzung, mit Ausnahme von Hausaufgaben erledigen, dem Schreiben von Texten, der schnellen Erstellung eines Textes, der Unterstützung der Ideenfindung, der Durchführung von Recherchen und der Einholung von Inspiration. Die höchsten absoluten Werte, die gleich oder größer als .39 sind, bestehen zwischen TCP und dem Schreiben von Texten für Blogs und Foren sowie dem Schreiben eines Blogs und eines Posts. Es gibt auch einige negative signifikante Korrelationen zwischen TCO und den Aufgaben Übersetzen, Schreiben für Blogs und Foren, Korrigieren einer E-Mail, Erstellen von Songtexten, Erstellen eines Beitrags für soziale Medien, der Quellenrecherche und dem Finden von Literatur. Die absoluten Korrelationen sind jedoch deutlich geringer als die für TCP. Von den drei Komponenten der Technikbereitschaft der SchülerInnen korreliert TCP mit den meisten Arten der Nutzung generativer KI mit den höchsten absoluten Werten.

Basierend auf diesen Ergebnissen zu Frage 2 über die Beziehung zwischen dem Einsatz generativer KI und einer Reihe von handlungsleitenden Eigenschaften, die als relevant für einige der Schlüsselaspekte der digitalen Transformation und der zukünftigen Bildung diskutiert werden, finden wir eine starke Unterstützung für unsere Annahme, dass es hochsignifikante Korrelationen zu den verschiedenen Arten der Nutzung generativer KI geben könnte. Ein Aspekt, den wir bemerkenswert finden, ist, dass die Wahrnehmung von generativer KI als anthropomorph eine wichtige Rolle für oder während der Nutzung dieser Technologien spielen kann. Zudem erachten wir es als relevant für Forschung und Praxis, dass KI-bezogene Selbstwirksamkeitserwartung als ein sehr wichtiger Aspekt für die Nutzung generativer KI angenommen werden kann. Darüber hinaus kann auch der von den SchülerInnen wahrgenommene Technostress für bestimmte Nutzungsszenarien eine wichtige Rolle spielen. Zudem kann die Tatsache, dass das kognitive Engagement der SchülerInnen negativ mit einer Reihe von Einsatzmöglichkeiten generativer KI korreliert, besorgniserregend sein, da bestimmte Einsatzmöglichkeiten generativer KI für kurzfristige Verbesserungen die zukünftigen Kompetenzen der SchülerInnen in den Bereichen Lernen,

Anpassungsfähigkeit oder Improvisation beeinträchtigen können. Wir argumentieren, dass z.B. in einer zunehmenden eher unreflektierten Nutzung dieser und ähnlicher KI-basierter Werkzeuge ein gewisses Risiko zu der Verringerung des kognitiven Engagements und der Freude am Lernen liegen kann. Die Ergebnisse in Tabelle 4 für die Variable NFC deuten dies an. Zur Untermauerung dieser Annahme sollten jedoch weitere empirische Untersuchungen, z. B. in Form einer Längsschnittstudie, durchgeführt werden.

Darüber hinaus finden wir es sehr interessant, dass es positive Korrelationen für die Technikakzeptanz der SchülerInnen als eine Variable der Technikbereitschaft gibt, aber insbesondere für die Technikkompetenzüberzeugung gibt es viele hochsignifikante negative Korrelationen. Ein Grund dafür kann in den Unterschieden zwischen Technikakzeptanz und Technikkompetenzüberzeugung gesehen werden. TAC befasst sich eher mit der Entwicklung einer Intention für die Nutzung einer Technologie auf der Basis von Benutzerfreundlichkeit und Nützlichkeit, während TCP die subjektive Erwartung eines Individuums über mögliche Handlungsoptionen in technologiebezogenen Situationen darstellt (vgl. Neyer et al., 2016, S. 88). Letztere ist stärker kognitiv ausgerichtet, was die Unterschiede in unseren Ergebnissen erklären könnte.

4.3 Zusammenhang: Nutzung generative KI & kontextuelle Faktoren (RQ3)

Zur Beantwortung unserer dritten Forschungsfrage über mögliche Zusammenhänge zwischen der Nutzung generativer KI durch SchülerInnen und kontextuellen Faktoren haben wir fünf Kontextvariablen betrachtet, die die soziale Unterstützung in Bezug auf die Nutzung von KI seitens der Bildungseinrichtung sowie der Eltern, sowie den sozialen Status, gemessen anhand eines Rankings des Bildungsniveaus der Eltern, der beruflichen Qualifikationen der Eltern und der allgemeinen finanziellen Situation der Eltern, darstellen (Tabelle 5). Die in Tabelle 5 dargestellten Werte zeigen, dass die SchülerInnen, im Vergleich zum Mittelwert von 2.5 der 5-stufigen Skalen, die Unterstützung seitens der Bildungseinrichtung sowie der Eltern im Durchschnitt etwas niedriger bewerten. Der Mittelwert für den sozialen Status zeigt, dass die SchülerInnen den sozialen Status ihrer Eltern höher einschätzen als den jeweiligen Skalenmittelwert (Tabelle 5).

Tabelle 5: Deskriptive Metriken der kontextuellen Faktoren

Kontextbezogene Variablen		Abkürzung	Deskriptive Merkmale		
			Min	Mittlere	Max
Soziale Unterstützung bei der Nutzung von KI seitens der Bildungseinrichtung		SOE	1.0	2.3	5.0
Soziale Unterstützung bei der Nutzung von KI seitens der Eltern		SOP	1.0	2.0	5.0
Sozialer Status	Schulisches Bildungsniveau der Eltern (Rangfolge)	PAS	4.0	7.2	8.0
	Berufliches Niveau der Eltern (Rangfolge)	PAP	4.0	8.2	10.0
	Finanzielle Situation der Eltern (Rangfolge)	PAF	1.0	3.4	5.0

Um die Beziehungen zwischen den fünf kontextuellen Faktoren und der Nutzung generativer KI zu untersuchen, testen wir auf signifikante bidirektionale Spearman-Korrelationen zwischen diesen Variablen. Die Ergebnisse sind in Tabelle 6 dargestellt. Die Korrelationsanalyse zwischen der sozialen Unterstützung seitens der Bildungseinrichtung (SOE) und der Nutzung generativer KI zeigt signifikante und in den meisten Fällen hochsignifikante positive Korrelationen für alle 31 Arten der Nutzung. Die stärksten und hochsignifikanten Korrelationen zeigen sich zwischen der Verbesserung der Sprache eigener Texte, dem Verfassen von Texten, dem Schreiben von Texten für Blogs und Foren, der inhaltlichen Überarbeitung eigener Texte, der Beantwortung von Fragen von LehrerInnen und der Quellenrecherche (Tabelle 6).

Für die Kontextvariable der sozialen Unterstützung seitens der Eltern (SOP) zeigen sich etwas stärkere Korrelationen mit den Arten der Nutzung generativer KI. Die stärksten Korrelationen mit Werten über .40 zeigen sich zwischen SOP und der Nutzung für das Schreiben von Texten, das Geben von Feedback zu einer Lösung anderer Personen, das Schreiben von Texten für Blogs oder Foren, das Überprüfen der Rechtschreibung eigener Texte und das Überarbeiten des Inhalts eigener Texte, das Beantworten von Fragen von LehrerInnen, das Schreiben eines Blogs, das Erstellen eines Beitrags in sozialen Medien,

das Verbessern der Sprache eigener Texte und das Erstellen von Tabellen, Diagrammen oder Abbildungen.

Tabelle 6: Korrelation zwischen kontextuellen Faktoren und der Nutzung generativer KI

Nutzungsarten generativer KI	Kontextuelle Faktoren				
	SOE	SOP	PAS	PAP	PAF
Hausaufgaben erledigen	.34**	.39**	-.07	-.06	.05
Texte schreiben	.38**	.41**	.03	-.01	.09
Schnell einen Text erstellen	.31**	.37**	-.03	-.02	.14
Übersetzen	.21**	.24**	.02	-.05	-.08
Spielen	.21**	.39**	-.06	-.05	.01
Chatten	.34**	.32**	-.24**	-.22**	-.06
Programmieren	.29**	.38**	-.26**	-.24**	-.03
Feedback zu einer Lösung anderer geben	.36**	.44**	-.24**	-.16*	-.09
Feedback zu meiner eigenen Lösung einholen	.29**	.34**	-.19*	-.14	.04
Ideenfindung unterstützen	.31**	.33**	-.08	.04	.12
Texte für Blogs und Foren verfassen	.37**	.49**	-.25**	-.18*	-.07
Rechtschreibung eigener Texte überprüfen	.32**	.43**	-.05	-.04	.02
Eigene Texte inhaltlich überarbeiten	.37**	.44**	-.10	-.17*	-.03
Recherchieren	.27**	.33**	-.08	-.05	.04
Fragen von LehrerInnen beantworten	.37**	.42**	-.11	-.07	.02
Blog schreiben	.33**	.41**	-.29**	-.18*	-.13
Post schreiben	.35**	.38**	-.25**	-.18*	-.14
E-Mail korrigieren	.28**	.35**	-.07	-.05	.03
Gliederung für einen Text erstellen	.26**	.38**	-.14	-.07	-.09
Sätze vervollständigen	.26**	.37**	-.13	-.09	-.13
Song Texte erstellen	.25**	.39**	-.12	-.09	-.05
Mathematikaufgaben lösen	.22**	.34**	-.11	-.04	.01
Social Media Beitrag erstellen	.31**	.44**	-.25**	-.20*	-.08
Eigene Texte sprachlich verbessern	.42**	.43**	-.16*	-.15*	.02
Quellen recherchieren	.37**	.37**	-.16*	-.18*	-.03
Literatur finden	.35**	.28**	-.13	-.10	-.09
Texte befragen	.30**	.32**	-.17*	-.14	-.06
Als Inspirationsquelle nutzen	.30**	.24**	-.06	.04	.05
Als Ratgeber nutzen	.28**	.38**	-.11	-.08	-.03
Chat mit literarischen Figuren oder historischen Personen	.29**	.36**	-.09	-.10	-.07
Tab., Schaub. und graph. Elem. erzeugen lassen	.26**	.41**	-.17*	-.13	-.07

SOE = Soziale Unterstützung bei der Nutzung von KI seitens der Bildungseinrichtung, SOP = Soziale Unterstützung bei der Nutzung von KI seitens der Eltern, PAS = Schulisches Bildungsniveau der Eltern, PAP = Berufliches Niveau der Eltern, PAF = Finanzielle Situation der Eltern

*: Signifikanzniveau $p < 0.05$

** : Signifikanzniveau $p < 0.001$

Die Analyse für den Kontextfaktor Schulabschluss der Eltern (PAS) zeigt, dass es in unserer Datenstichprobe einige negative Korrelationen mit der Nutzung generativer KI gibt. Die Nutzung von generativer KI zum Chatten, Programmieren, Feedback zur Lösung anderer geben, Feedback zur eigenen Lösung erhalten, Texte für Blogs und Foren schreiben, einen Blog schreiben, einen Post schreiben, einen Social-Media-Beitrag erstellen, die Sprache eigener Texte verbessern, Quellen recherchieren und Tabellen, Schaubilder und Grafiken erstellen lassen, korreliert signifikant negativ mit PAS.

Die Analyse des Zusammenhangs zwischen der beruflichen Qualifikation der Eltern (PAP) und der Nutzung generativer KI zeigt einige Ähnlichkeiten im Vergleich zur Schulbildung der Eltern (PAS). Es gibt einige Arten der KI-Nutzung, die nicht mit PAP, sondern mit PAS korrelieren und umgekehrt. Zudem sind ein paar Korrelationswerte für PAS leicht höher und signifikanter als die für PAP.

Es gibt keine signifikanten Korrelationen zwischen der finanziellen Situation der Eltern (PAF) und der Nutzung generativer KI. Im Vergleich zu den anderen Variablen, die wir als Indikatoren für den sozialen Status betrachtet haben, PAS und PAP, scheint der Zusammenhang zwischen PAF und der Nutzung generativer KI in unserer Stichprobe nicht relevant zu sein.

Die Ergebnisse in Tabelle 6 zeigen, dass insbesondere die soziale Unterstützung seitens der Schule und der Eltern als sehr wichtig für die Nutzung generativer KI angesehen werden kann. Die Ergebnisse auf der Basis unserer Stichprobe zeigen auch, dass der soziale Status eine wichtige Rolle für oder während der Nutzung generativer KI spielen kann. Allerdings scheinen die Beziehungen zwischen dem sozialen Status und den verschiedenen Arten der Nutzung generativer KI sehr viel fragmentierter zu sein, und insbesondere die finanzielle Situation der Eltern scheint weniger wichtig zu sein.

4.4 Zusammenhang: Nutzung generative KI & wahrgenommener Lernerfolg (RQ4)

Zur Beantwortung der Forschungsfrage 4 berechnen wir den Spearman-Korrelationskoeffizienten zwischen allen 31 Arten der generativen KI-Nutzung und dem von den SchülerInnen wahrgenommenen KI-bezogenen Lernerfolg anhand einer Teilstichprobe von N = 111. Der Grund für die Berechnung an einer Teilstichprobe ist, dass die Lernerfolgssitems im Laufe der Datenerhebung anhand von Rückmeldungen und Diskussionen mit Experten auf dem Gebiet hinzugefügt wurden. Die Skalensreliabilität des selbst wahrgenommenen KI-bezogenen Lernerfolgs (LS) der SchülerInnen hat einen Wert von Cronbachs Alpha = .912.

Tabelle 7: Korrelationskoeffizienten zwischen der Nutzung generativer KI und dem selbst wahrgenommenen KI-bezogenen Lernerfolg (LS)

Nutzungsarten generativer KI		Nutzungsarten generativer KI	
Hausaufgaben erledigen	.74**	Post schreiben	.19
Texte schreiben	.69**	E-Mail korrigieren	.26**
Schnell einen Text erstellen	.76**	Erstellen einer Gliederung für einen Text	.55**
Übersetzen	.24**	Vervollständigung von Sätzen	.31**
Spielen	.12	Erstellen von Liedtexten	.23**
Chatten	.22*	Lösen von Matheaufgaben	.46**
Programmieren	.29**	Erstellen eines Beitrags in sozialen Medien	.23**
Feedback zu einer Lösung anderer geben	.43**	Post schreiben	.57**
Feedback zu meiner eigenen Lösung einholen	.44**	E-Mail korrigieren	.39**
Ideenfindung unterstützen	.58**	Gliederung für einen Text erstellen	.24*
Texte für Blogs und Foren verfassen	.42**	Sätze vervollständigen	.45**
Rechtschreibung eigener Texte überprüfen	.33**	Song Texte erstellen	.48**
Eigene Texte inhaltlich überarbeiten	.58**	Mathematikaufgaben lösen	.40**
Recherchieren	.63**	Social media Beitrag erstellen	.33**
Fragen von LehrerInnen beantworten	.51**	Eigene Texte sprachlich verbessern	.47**
Blog schreiben	.29**		

*: Signifikanzniveau $p < 0.05$

** : Signifikanzniveau $p < 0.001$

Aus den Ergebnissen in Tabelle 7 lässt sich schließen, dass der von den SchülerInnen selbst wahrgenommene Lernerfolg signifikant positiv mit 29 von 31 Items der Nutzungshäufigkeit generativer KI korreliert. Für neun der Arten, nämlich für das Erledigen von Hausaufgaben, das Schreiben von Texten, das schnelle Erstellen eines Textes, die Unterstützung der Ideenfindung, das Überarbeiten des Inhalts eigener Texte, das Recherchieren, das Beantworten von Lehrerfragen, das Erstellen einer Gliederung für einen Text und das Verbessern der Sprache eigener Texte, sind die Korrelationskoeffizienten sogar größer als .50. Die Korrelationskoeffizienten für die Erledigung von Hausaufgaben und die schnelle Erstellung eines Textes liegen sogar bei über .70. Die Wahrnehmung, mit Hilfe von KI in der Schule erfolgreicher zu sein, scheint also in hohem Maße positiv mit der

tatsächlichen Nutzung generativer KI zusammenzuhängen. Zu bemerken ist auch, dass die Aufgaben, die hoch mit dem selbst wahrgenommenen KI-bezogenen Lernerfolg korrelieren, Aufgaben sind, die einen starken Bezug zur Schule aufzeigen.

4.5 Gruppierung der Nutzung generativer KI in übergeordnete Konzepte (RQ5)

Zur Beantwortung von Frage 5 führen wir eine explorative Faktorenanalyse (EFA) für die 31 Items durch. Zunächst testen wir, ob unsere Daten für die Durchführung einer EFA geeignet sind, indem wir den Bartlett-Test auf Sphärizität durchführen und uns das Kaiser-Meyer-Olkin-Maß (KMO) anschauen (Fabrigar und Wegener, 2011).

Tabelle 8: Gruppen der generativen KI-Nutzung

Items und Faktoren		Faktorladungen			
		1	2	3	4
Faktor 1: Erledigung von Standardaufgaben					
1	Hausaufgaben erledigen	.681			
2	Texte schreiben	.906			
3	Schnell einen Text erstellen	.746			
Faktor 2: Erkundung neuer Möglichkeiten					
5	Chatten		.551		
6	Spielen		.686		
11	Erstellen von Texten für Blogs und Foren		.681		
16	Blogs erstellen		.832		
17	Posts erstellen		.942		
20	Sätze vervollständigen		.607		
21	Song Texte erstellen		.773		
23	Social Media Beitrag erstellen		.839		
Faktor 3: Verbesserung der eigenen Arbeitsergebnisse					
9	Feedback zur eigenen Lösung einholen			.639	
12	Rechtschreibung eigener Texte überprüfen			.706	
13	Eigene Texte inhaltlich überarbeiten			.709	
Faktor 4: Anregung kreativen Denkens					
10	Ideenfindung unterstützen				.644
28	Als Inspirationsquelle nutzen				.614
29	Als Ratgeber nutzen				.571
Prozentsatz der erklärten Varianz (Summe: 67.7 %)		15.03%	29.68%	12,14%	10.83%
Cronbachs Alpha		.890	.927	.849	.808

Der Bartlett-Test prüft die Nullhypothese, ob die Korrelationsmatrix der Items eine Identitätsmatrix ist, um sicherzustellen, dass es Beziehungen zwischen den Items oder Gruppen von Items gibt. Daher sollte der Bartlett-Test signifikant sein. Das Kaiser-Meyer-Olkin-Maß ist ein Maß für den Anteil der Varianz zwischen den Items, der in der Varianz geteilt sein kann. KMO-Werte zwischen .8 und 1 bedeuten, dass die Stichprobe für die EFA geeignet ist. In unserem Fall ist der Bartlett-Test mit $\chi^2(300)=2468.09$ und $p < .001$ signifikant und der KMO-Wert beträgt .915. Da beide Tests zeigen, dass die Stichprobe für die EFA geeignet ist, fahren wir mit der Analyse fort. Wir führen eine Varimax-Rotation durch und

verwenden den „Weighted-Least Squares“ Schätzer. Nur Items, die mit mindestens zwei anderen Items signifikant korrelieren und Faktorladungen von mehr als .33 aufweisen, werden in die Analyse einbezogen. Somit werden die Items mit den Nummern 1, 2, 3, 5, 6, 9, 10, 11, 12, 13, 16, 17, 20, 21, 23, 28, 29 in die Analyse einbezogen.

Das Endergebnis der EFA (siehe Tabelle 8) besteht aus den vier Faktoren, die wir „Erledigung von Standardaufgaben“, „Erkundung neuer Möglichkeiten“, „Verbesserung der eigenen Arbeitsergebnisse“ und „Anregung kreativen Denkens“ genannt haben. Die erklärte Gesamtvarianz beträgt 67.7 %. In Tabelle 8 ist auf der Grundlage der statistischen Analyse dargelegt, wie die 31 Arten der Nutzung generativer KI in die vier übergeordneten Konzepte und Faktoren gruppiert werden können. Diese Ergebnisse können beispielsweise verdeutlichen, dass ein/e Schüler/in, die generative KI für das Erledigen von Hausaufgaben nutzt, diese auch zum Schreiben von Texten und zur schnellen Texterstellung verwendet. Die vier übergeordneten Konzepte der Nutzung generativer KI können daher als eine erste Konsolidierung der Nutzung generativer KI in vier komplementäre Bereiche interpretiert werden. Diese kann dazu beitragen, die Routinen der Nutzung generativer KI von SchülerInnen besser zu verstehen und diese bei der adäquaten Nutzung zu unterstützen.

5. Diskussion

In der vorliegenden Studie wurden die Ergebnisse einer quantitativen Erhebung vorgestellt, die anhand einer Befragung unter 15- bis 19-jährigen SchülerInnen im Zeitraum zwischen März und Juli 2023 durchgeführt wurde. Ziel der Studie war es, zu untersuchen, ob, für welche Aufgaben und wie häufig SchülerInnen generative KI in oder für die Schule nutzen. Darüber hinaus wurde untersucht, ob und welche Zusammenhänge zwischen der Intensität der Nutzung generativer KI und individuellen handlungsleitenden Eigenschaften bestehen, die als relevant für die Technologienutzung gelten, wie z.B. die soziale Wahrnehmung von KI, kognitives Engagement, KI-bezogene Selbstwirksamkeitserwartung, Technostress und die Technikbereitschaft. Es wurde auch analysiert, wie kontextuelle Faktoren, wie z.B. die soziale Unterstützung bei der KI-Nutzung durch Schule und Eltern und der soziale Status der Eltern, mit der Intensität der KI-Nutzung einhergehen. Es wurde auch untersucht, ob ein Zusammenhang zwischen der Intensität der KI-Nutzung und dem wahrgenommenen Lernerfolg der SchülerInnen besteht und ob sich die 31 Arten der Nutzung generativer KI zu übergeordneten Konzepten, die auch als Typen der Nutzung generativer KI bezeichnet werden können, zusammenfassen lassen.

Die Ergebnisse unserer Stichprobe zeigen, dass 50 % bis 80 % der befragten SchülerInnen generative KI für viele der angegebenen Zwecke eher nie verwenden. Dies könnte darauf zurückzuführen sein, dass generative KI noch ein sehr neues Werkzeug ist und in Relation zu vielen anderen Lernwerkzeugen und -methoden gesetzt werden sollte, die bereits in der Schule eingesetzt werden und weitgehend etabliert sind. Auf der Grundlage unserer Ergebnisse lässt sich dennoch argumentieren, dass der Einsatz dieser neuartigen Technologie im Zusammenhang mit Schule immer wichtiger wird, insbesondere für spezifische schulbezogene Aufgaben wie Hausaufgaben erledigen, Texte schreiben, übersetzen und neue Ideen einholen, z.B. bei der Recherche.

In Bezug auf die Nutzung generativer KI und die handlungsleitenden Eigenschaften der SchülerInnen fanden wir einige signifikante Zusammenhänge mit verschiedenen Arten der Nutzung generativer KI. Ein sehr interessanter Aspekt ist zum Beispiel, dass die Wahrnehmung generativer KI als menschenähnlich, z.B. als „warm“, „organisch“ oder „real“, eine wichtige Rolle bei der Nutzung solcher Technologien spielen könnte, da wir für die meisten Nutzungsarten eine signifikante positive Korrelation mit der Wahrnehmung der KI als anthropomorph fanden. Wir fanden auch signifikante positive Korrelationen zwischen der wahrgenommenen KI-bezogenen Selbstwirksamkeitserwartung und der Nutzungsintensität generativer KI für die meisten der 31 generativen KI-Nutzungstypen. Dies könnte bedeuten, dass die KI-bezogene Selbstwirksamkeitserwartung als entscheidend für die schulbezogene Nutzung generativer KI angenommen werden kann. Darüber hinaus scheinen die SchülerInnen der Studie in gewissem Maße ein Gefühl von Technostress im Zusammenhang mit dem Einsatz generativer KI zu empfinden. Dieses Ergebnis steht im Einklang mit anderen aktuellen Forschungsergebnissen (Ragu-Nathan et al., 2008; Tarafdar et al., 2011). In unseren Daten sind die Korrelationswerte für die meisten Nutzungsarten auf einem moderaten Niveau. Interessanterweise scheinen die SchülerInnen das ausgeprägteste Stresserleben zu verspüren, wenn sie gleichzeitig

generative KI nutzen, um einen Blog oder einen Post zu schreiben oder einen Beitrag in den sozialen Medien zu erstellen. Für diese Aufgaben ist die Korrelation größer als .40 und hoch signifikant. Die SchülerInnen schätzen hierbei insbesondere ihre Fähigkeiten im Verhältnis zu den Anforderungen während der Technologienutzung als zu gering ein. Es kann vermutet werden, dass sich die SchülerInnen als zu wenig produktiv oder effektiv in der Nutzung solcher und ähnlicher digitaler Tools empfinden. Gleichwohl können diese Arten der Nutzung generativer KI nicht unbedingt mit der Schule in Verbindung gebracht werden.

Demgegenüber sind das kognitive Engagement im Sinne allgemeiner Freude am Nachdenken, negativ mit einer Reihe verschiedener eindeutig schulbezogener Nutzungsweisen generativer KI korreliert, z. B. dem Lösen von Mathematikaufgaben oder dem Beantworten von Fragen der LehrerInnen. Daraus könnte man schließen, dass eine häufigere Nutzung von KI mit einer Abnahme des kognitiven Engagements einhergehen kann. Dies könnte als zukünftige Herausforderung interpretiert werden, da die Verfügbarkeit und Intensität der Nutzung generativer KI voraussichtlich zunehmen werden. Eine besondere Herausforderung und gleichzeitig auch eine Chance, die sich aus diesen Ergebnissen ergeben kann, ist, dass sich das Bildungssystem in Deutschland noch stärker auf eine nachhaltige Entwicklung von Kompetenzen der SchülerInnen wie kritisches Denken, Selbstreflexion, Improvisation oder Achtsamkeit fokussieren sollte.

Die Ergebnisse der Studie zeigen, dass es positive Korrelationen zwischen der Technikakzeptanz der SchülerInnen und den verschiedenen Arten der Nutzung von KI gibt, während es bei der Technikkompetenzüberzeugung viele hoch signifikante negative Korrelationen gibt. Die Technikakzeptanz konzentriert sich tendenziell auf die Entwicklung einer Absicht, eine Technik auf der Grundlage von Benutzerfreundlichkeit und Nützlichkeit zu verwenden, während Technikkompetenzüberzeugung eher die subjektive Erwartung einer Person hinsichtlich möglicher Handlungsweisen in technologiebezogenen Situationen darstellt (vgl. Neyer et al., 2016, S. 88), d.h. z.B. eher darauf fokussiert ist, wie sich Techniknutzung sinnvoll in individuelle Problembearbeitung integrieren lässt. Letztere orientiert sich demnach expliziter an den kognitiven Prozessen der Schülerinnen und Schüler in Bezug darauf, wie und zu welchem Zweck generative KI im Einzelnen eingesetzt werden könnte, um beispielsweise individuelle Ziele zu erreichen. Wir vermuten, dass der Grund für diesen Befund darin liegen könnte, dass sich die SchülerInnen einerseits noch nicht vollständig der stetig wachsenden Möglichkeiten bewusst sind, die generative KI-Werkzeuge ihnen bieten können, und andererseits dieser neuen und vielleicht ungewohnten Technik in gewissem Maße kritisch gegenüberstehen.

Die Analyse der Rolle von Kontextfaktoren bei der Nutzung generativer KI zeigt, dass die soziale Unterstützung durch Eltern und Bildungseinrichtungen sowie der soziale Status der Eltern eine wichtige Rolle bei der Nutzung generativer KI spielen können. Insbesondere die soziale Unterstützung durch Schulen und Eltern kann als sehr relevant bei der Nutzung generativer KI angesehen werden. Wir interpretieren das so, dass SchülerInnen ein gewisses Maß an Befähigung für die Nutzung generativer

KI wünschen, was in Form indirekter Unterstützung durch Eltern oder LehrerInnen gesehen werden kann, z.B. die situative Verfügbarkeit generativer KI durch die Bereitstellung des Zugangs zu dieser Technologie. Gleichzeitig könnten SchülerInnen aber auch eine direktere Unterstützung bei der Nutzung generativer KI benötigen, wie z.B. eine gemeinsame kritische Reflexion über die Chancen, aber auch die Risiken generativer KI. Aus unserer Sicht ist auch in diesem speziellen Bereich der sozialen Unterstützung umfassende Forschung erforderlich, um die direkten und indirekten Unterstützungsmöglichkeiten von Eltern, Erziehungsberechtigten und LehrerInnen genauer zu erfassen. Darüber hinaus zeigen die Ergebnisse, dass der soziale Status eine wichtige Rolle im Zusammenhang mit der Nutzung generativer KI an Schulen spielen kann. Allerdings scheinen die Beziehungen zwischen dem sozialen Status und den verschiedenen Arten der Nutzung generativer KI eher fragmentiert zu sein, und insbesondere die finanzielle Situation der Eltern scheint weniger bedeutend zu sein. Dennoch ist zu berücksichtigen, dass die Rolle der Bildungs- und Finanzsituation des familiären Umfelds im Hinblick auf die Bildungsgerechtigkeit nicht zu unterschätzen ist. Aus unserer Perspektive kann generative KI unter gewissen Voraussetzungen sowohl als eine neue Triebkraft für mehr Bildungsgerechtigkeit in Deutschland gesehen werden und gleichzeitig auch ein neuartiges Risiko für eine wachsende Ungleichheit in unserem Bildungssystem darstellen. Letzteres lässt sich vermuten, wenn z.B. der Zugang zu generativer KI für SchülerInnen und deren Nutzung besonders zu Beginn wenig oder gar nicht durch direkte und indirekte Unterstützungsangebote flankiert wird oder kein offener Dialog stattfindet.

Im Hinblick auf eine Konsolidierung der insgesamt 31 Nutzungsmöglichkeiten generativer KI zu übergeordneten Gruppen im Sinne von Typen der Nutzung können die betrachteten Nutzungsmöglichkeiten auf der Grundlage unserer statistischen Ergebnisse in vier Bereiche gruppiert werden. Diese lassen sich anhand unserer Interpretation zusammenfassen als das *Erledigen von Standardaufgaben*, die *Erkundung neuer Möglichkeiten*, die *Verbesserung eigener Arbeitsergebnisse* und die *Anregung kreativen Denkens*. Wir gehen davon aus, dass diese vier Bereiche bzw. Typen der Nutzung noch kein endgültiges Konzept darstellen. Vielmehr werden sich diese Bereiche aufgrund des zu erwartenden weiteren Ausbaus der Fähigkeiten generativer KI, der allgemeinen technologischen Entwicklungen sowie der Kreativität von SchülerInnen und anderen Nutzergruppen bei der Verwendung von KI weiterentwickeln. Wir verstehen diese Systematisierung eher als wichtigen Ausgangspunkt für weitere empirische Forschung und statistische Analysen in diesem Bereich, um darauf aufbauend Gestaltungsansätze im bzw. für das Bildungssystem zu entwickeln.

Schließlich zeigt die Studie, dass der selbst wahrgenommene KI-bezogene Lernerfolg der SchülerInnen signifikant positiv mit fast allen Arten der generativen KI-Nutzung korreliert. Wir finden es bemerkenswert, dass der Korrelationskoeffizient für generative KI-Nutzungen wie Texte schreiben, Ideenfindung unterstützen, den Inhalt des eigenen Textes überarbeiten, recherchieren, Fragen von LehrerInnen beantworten, eine Gliederung für einen Text erstellen sowie die Sprache eigener Texte zu verbessern sogar höher als .50 ist, während die Korrelationskoeffizienten zwischen KI-bezogenem

Lernerfolg und Hausaufgaben erledigen und schnell einen Text erstellen sogar höher als .70 sind. Gleichzeitig ist das Spielen nicht signifikant mit dem von den SchülerInnen selbst wahrgenommenen Lernerfolg korreliert.

Wir sind uns bewusst, dass die vorliegende Stichprobe nicht repräsentativ ist. Die Studie liefert jedoch erste Anhaltspunkte dafür, dass sich SchülerInnen heute und wahrscheinlich in Zukunft in größerem Umfang mit generativer KI als stets verfügbarem und sehr hilfreichen intelligenten Assistenten beschäftigen. Auf der anderen Seite sehen wir das Risiko, dass SchülerInnen dazu geneigt sein könnten, ihre individuellen Lernergebnisse, die in der Schule traditionell durch entsprechende Noten repräsentiert werden, durch umfangreiche Nutzung generativer KI versuchen zu optimieren. Dieses Spannungsverhältnis schafft aus unserer Sicht eine völlig neue Dynamik im Kontext moderner Bildung und kann perspektivisch den Druck erhöhen, das Bildungssystem noch schneller und explizierter an einem kompetenzorientierten Gesamtansatz zu orientieren. Generell folgern wir, dass die Frage danach, wie der individuelle Lernzuwachs oder die Kompetenz- und Persönlichkeitsentwicklung von SchülerInnen fair und transparent bewertet und anerkannt werden können, seit der umfangreichen Verfügbarkeit generativer KI nicht mehr mit den bisherigen Ansätzen beantwortet und gelöst werden kann. Natürlich sind die Ansätze hin zu einer stärkeren Kompetenzorientierung keineswegs neu und vieles ist in den letzten Jahren auch bewegt worden, jedoch gehen wir davon aus, dass auch bereits etablierte Lösungsansätze vor dem Hintergrund der neuen Chancen und Risiken von generativer KI kritisch hinterfragt werden sollten. Aus unserer Sicht können KI-basierte Technologien hierbei auch als starker Impulsgeber für innovative Lösungen in Richtung eines kompetenzorientierten Bildungssystems in Deutschland und international in Betracht gezogen werden.

Die vorliegende Studie als solches, einschließlich der Analyse und der berichteten Ergebnisse, ist eine der wenigen empirischen Untersuchungen unter SchülerInnen, die nicht zwingenderweise volljährig sind und sich auf Selbstauskünften über die Nutzung generativer KI stützt. Wir möchten WissenschaftlerInnen, BildungsexpertInnen, politische EntscheidungsträgerInnen, PraktikerInnen und alle weiteren Interessierten dazu einladen, die Ergebnisse dieser Arbeit als einen ersten Ausgangspunkt für die weitere kritische und gleichermaßen kreative Auseinandersetzung mit den aktuellen Dynamiken zu verstehen, um die Potenziale generativer KI für unsere Bildung nutzbar zu machen und hierbei gleichwohl die Risiken im Blick zu behalten.

Der wesentliche Beitrag dieser Studie besteht aus unserer Sicht darin, erste Belege für verschiedene Annahmen und Initiativen zu liefern, die derzeit in der Bildungspraxis und -forschung diskutiert werden. Anzumerken ist, dass weitaus mehr empirische Forschung in diesem Bereich notwendig ist. Wir regen an, insbesondere Längsschnittstudien durchzuführen, um die tatsächlichen Ursachen und Wirkungen, z.B. verschiedener Arten der Nutzung generativer KI, die daraus resultierenden Ergebnisse, wie den tatsächlichen Lernerfolg und eine kritische Reflexion darüber, wie der tatsächliche Kompetenzzuwachs von SchülerInnen in der Bildung bewertet, erkannt und gewürdigt werden kann, besser zu verstehen. Wir laden alle an diesem Thema Interessierten ein, sich an weiteren empirischen Studien oder

Untersuchungen zu beteiligen und gemeinsam im offenen Diskurs zu bleiben. Mit unserer Forschung wollen wir die kontinuierliche Entwicklung eines zukunftsfähigen Bildungssystems in Deutschland und darüber hinaus unterstützen.

6. Limitationen

Hinsichtlich der Ergebnisse der Faktorenanalyse und der Entwicklung übergreifender Typen der Nutzung generativer KI sind weitere kritische Reflexionen und empirische Untersuchungen auf Basis einer repräsentativen Stichprobe mit mehr Datenpunkten sowie eine konfirmatorische Faktorenanalyse entscheidend, um die Struktur und Benennung der vier Bereiche zur Nutzung generativer KI weiter zu fundieren. Weitere ausführliche Diskussionen mit Experten auf diesem Gebiet könnten dabei sehr hilfreich sein. Die in dieser Studie verwendete Skala zur Technologiebereitschaft wurde noch nicht mit SchülerInnen getestet. Die Reliabilitäten der Skala zeigten jedoch gute Ergebnisse, so dass wir uns entschieden haben, sie für unsere Analyse zu verwenden, da sie eine interessante Erweiterung des klassischen Technikakzeptanzmodells darstellt. Da diese Studie auf einer Korrelationsanalyse mit nur einem Messzeitpunkt basiert, ist es außerdem wichtig zu erwähnen, dass die Korrelationen nicht so interpretiert werden sollten, dass es kausale Beziehungen zwischen den gemessenen Variablen gibt.

Unsere Stichprobenstruktur zeigt einen relativ hohen Anteil von SchülerInnen (49 %), die in einem eher akademischen Haushalt leben. Dieser Wert liegt über dem deutschen Durchschnitt. Darüber hinaus wird die finanzielle Situation von etwa 45 % der SchülerInnen in unserer Stichprobe als überdurchschnittlich gut eingeschätzt. Ein Grund für diese Ergebnisse ist darin zu sehen, dass drei der vier Schulen, aus denen die SchülerInnen unserer Stichprobe stammen, Gymnasien sind. Zudem basieren diese Daten lediglich auf Selbsteinschätzungen der SchülerInnen. Diese Aspekte und die Tatsache, dass die Stichprobengröße noch limitiert ist, deuten darauf hin, dass unsere Studie noch nicht repräsentativ ist. Dennoch ist dies eine der ersten Studien, die sich mit den tatsächlichen Sichtweisen von SchülerInnen bei der Nutzung von generativer KI für verschiedene Aufgaben im Schulalltag beschäftigt. Aus unserer Sicht sind weitere empirische Untersuchungen und insbesondere Längsschnittstudien sehr wichtig, um besser zu verstehen, wie generative KI die Lernprozesse in Bildungseinrichtungen beeinflusst und wie Bildung zukünftig gestaltet werden sollte.

Literaturverzeichnis

- Bandura, A. (1986). Soziale Grundlagen des Denkens und Handelns: Eine sozial-kognitive Theorie. Prentice-Hall.
- Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2014). Allgemeine Selbstwirksamkeit Kurzsкала (ASKU). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). <https://doi.org/10.6102/ZIS35>
- Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). Die effiziente Bewertung des Kognitionsbedarfs. *Journal of Personality Assessment*, 48(3), 306-307. https://doi.org/10.1207/s15327752jpa4803_13
- Cronbach, L. J. (1951). Koeffizient Alpha und die interne Struktur von Tests. 38.
- Davis, F. D. (1989). Wahrgenommene Nützlichkeit, wahrgenommene Benutzerfreundlichkeit und Benutzerakzeptanz von Informationstechnologie. *MIS Quarterly*, 13, 319-340.
- Ellen, P. S., Bearden, W. O., & Sharma, S. (1991). Widerstand gegen technologische Innovationen: Eine Untersuchung der Rolle von Selbstwirksamkeit und Leistungszufriedenheit. *Journal of the Academy of Marketing Science*, 19(4), 297-307. <https://doi.org/10.1007/BF02726504>
- Hahn, M. H., & Lee, K. C. (2017). Exploring the Role of Self-Confidence, Need-for-Cognition, and the Degree of IT Support on Individual Creativity: Multilevel Analysis Approach. *Current Psychology*, 36(3), 565-576. <https://doi.org/10.1007/s12144-016-9445-z>
- Kulviwat, S., C. Bruner II, G., & P. Neelankavil, J. (2014). Self-efficacy as antecedent of cognition and affect in technology acceptance. *Journal of Consumer Marketing*, 31(3), 190-199. <https://doi.org/10.1108/JCM-10-2013-0727>
- Mandl, S., Bretschneider, M., Asbrock, F., Meyer, B., & Strobel, A. (2022). Die Skala zur sozialen Wahrnehmung von Robotern (SPRS): Entwicklung und Test einer Skala für erfolgreiche Interaktion zwischen Menschen und Robotern. In L. M. Camarinha-Matos, A. Ortiz, X. Boucher, & A. L. Osório (Eds.), *Collaborative Networks in Digitalization and Society 5.0* (Vol. 662, pp. 321-334). Springer International Publishing. https://doi.org/10.1007/978-3-031-14844-6_26
- Mandl, S., Kobert, M., Bretschneider, M., Asbrock, F., Meyer, B., Strobel, A., & Süße, T. (2023). Die Erforschung von Schlüsselkategorien der sozialen Wahrnehmung und moralischen Verantwortung von KI-basierten Agenten am Arbeitsplatz: Findings from a case study in an industrial setting. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1-6. <https://doi.org/10.1145/3544549.3585906>
- Müller, P., Sommer, S., & Eitner, A. (2023). Aufbruch ins Unbekannte - Schule in Zeiten von künstlicher Intelligenz und ChatGPT. Vodafone Stiftung Deutschland gGmbH. https://www.vodafone-stiftung.de/wp-content/uploads/2023/04/Aufbruch-ins-Unbekannte_Studie-zu-KI-im-Schulkontext.pdf
- Neyer, F. J., & Felber, C., J. & Gebhardt. (2012). Entwicklung und Validierung einer Kurzsкала zur Erfassung von Technikbereitschaft (technology commitment). *Diagnostica*, 58, 87-99.
- Neyer, F. J., Felber, J., & Gebhardt, C. (2016). Kurzsкала Technikbereitschaft (TB, technology commitment). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). <https://doi.org/10.6102/ZIS244>
- OECD. (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. Organisation für wirtschaftliche Zusammenarbeit und Entwicklung. https://www.oecd-ilibrary.org/education/equity-in-education_9789264073234-en

- OECD. (2023). Shaping Digital Education. <https://doi.org/10.1787/bac4dc9f-en>
- Preckel, F. (2016). NFC-Teens: Eine deutsche Need for Cognition Skala für ältere Kinder und Jugendliche. Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). <https://doi.org/10.6102/ZIS241>
- Ragu-Nathan, T. S., Tarafdar, M., Ragu-Nathan, B. S., & Tu, Q. (2008). Die Folgen von Technostress für Endbenutzer in Organisationen: Conceptual Development and Empirical Validation. *Information Systems Research*, 19(4), 417-433. <https://doi.org/10.1287/isre.1070.0165>
- Su, J., & Weipeng, Y. (2023). Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education. *ECNU Review of Education*, 6(3), 355-366. <https://doi.org/10.1177/20965311231168423>
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273-1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tarafdar, M., Tu, Q., & Ragu-Nathan, T. (2011). Impact of Technostress on End-User Satisfaction and Performance. *J. of Management Information Systems*, 27, 303-334. <https://doi.org/10.2307/29780194>
- Wang, X., Tan, S. C., & Li, L. (2020). Messung des Technostresses von Universitätsstudenten beim technologiegestützten Lernen: Scale development and validation. *Australasian Journal of Educational Technology*, 36(4), Artikel 4. <https://doi.org/10.14742/ajet.5329>
- Yoon, N., & Lee, H.-K. (2021). AI Recommendation Service Acceptance: Assessing the Effects of Perceived Empathy and Need for Cognition. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(5), Artikel 5. <https://doi.org/10.3390/jtaer16050107>