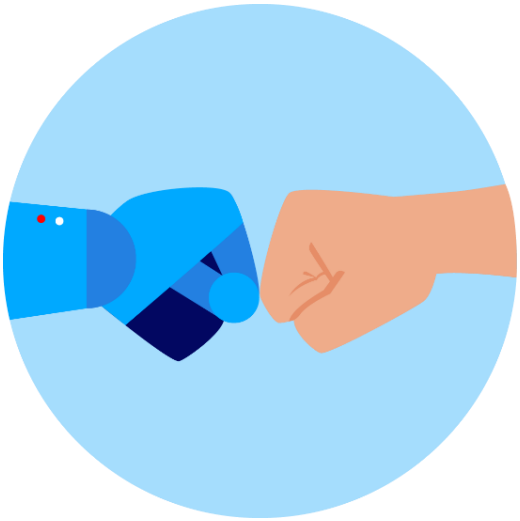
Fachhochschule Nordwestschweiz
Hochschule für Angewandte Psychologie

Veränderung der Arbeitswelt durch ChatGPT – Einfluss von ChatGPT auf die Erinnerungsleistung

MASTER-ARBEIT

2024



Autorin / Autor
Genovese, Rosina Martina

Begleitperson
Dr. prof. Schulze, Hartmut
Mateescu, Magdalena

Praxispartnerin
Swisscom Schweiz AG
Dr. Dobos, Daniel

Abstract (Englisch)

In this master's thesis, the influence of ChatGPT use on the memory performance of Swisscom employees was investigated. Accordingly, the question was: «How does ChatGPT use affect the memory performance of Swisscom employees?».
To collect the data, a memory test was conducted with two measurement points. An experiment was carried out at the first measurement point and a questionnaire survey was conducted at the second measurement point, with the questionnaire being sent out two to three weeks after the experiment. The subjects (N = 45) from the treatment groups had the task of answering trivia questions using ChatGPT, while the control groups were asked to use Google. The results show no significant differences in memory performance between the ChatGPT and Google groups. Only slight, insignificant trends are recognizable in that the Google groups were able to demonstrate better memory performance. These are positive results for Swisscom in view of all digitalization projects with AI, as the risk of impairing employees' memory performance can be classified as low. A comparison with other AI tools or an analysis of the influences on other cognitive abilities would be possible further research approaches for the future.

Keywords: ChatGPT and memory performance, ChatGPT use and impacts, internet use and memory performance / cognitive performance

Abstract (Deutsch)

In der vorliegenden Masterarbeit wurde der Einfluss von der ChatGPT-Nutzung auf die Erinnerungsleistung von Swisscom-Mitarbeitenden erforscht. Die Fragestellung lautete dementsprechend: «*Wie wirkt sich die ChatGPT-Nutzung auf die Erinnerungsleistung von Swisscom-Mitarbeitenden aus?*».

Zur Datenerhebung wurde hierfür ein Erinnerungstest mit zwei Messzeitpunkten durchgeführt. Am ersten Messzeitpunkt wurde ein Experiment durchgeführt und am zweiten Messzeitpunkt erfolgte eine Fragebogenerhebung, wobei der Bogen zwei bis drei Wochen nach dem Experiment verschickt wurde. Die Probanden (N = 45) aus den Treatmentgruppen hatten die Aufgabe, mittels ChatGPT Trivia-Fragen zu beantworten, während die Kontrollgruppen Google verwenden sollten. Die Ergebnisse zeigen keine signifikanten Unterschiede in der Erinnerungsleistung zwischen den ChatGPT- und den Google-Gruppen. Es sind nur leichte, insignifikante Trends dahingehend erkennbar, dass die Google-Gruppen eine bessere Erinnerungsleistung zeigen konnten. Für die Swisscom sind das in Anbetracht aller Digitalisierungs-Projekte mit AI positive Ergebnisse, da sich das Risiko einer Beeinträchtigung der Erinnerungsleistung der Mitarbeitenden als gering einstufen lässt. Ein Vergleich mit anderen AI-Tools oder eine Analyse der Einflüsse auf andere kognitive Fähigkeiten wären mögliche weitere Forschungsansätze für die Zukunft.

Schlüsselwörter: ChatGPT und Gedächtnisleistung, ChatGPT-Nutzung und Auswirkungen, Internetnutzung und Gedächtnisleistung / kognitive Leistung

Inhaltsverzeichnis

1 Einleitung 6
1.1 Forschungsfrage und Zielsetzung 6
1.2 Relevanz der Studie und Praxispartnerin Swisscom Schweiz AG 8
1.3 Abgrenzungen 9
1.4 Aufbau der Arbeit 9
2 Theoretische Grundlagen 11
2.1 Die unterschiedlichen Gedächtnisspeicher 11
2.1.1 Das sensorische Gedächtnis 11
2.1.2 Das Kurzzeitgedächtnis 11
2.1.3 Das Langzeitgedächtnis 11
2.1.4 Das transaktive Gedächtnis 12
2.2 Gedächtnis und Erinnerungsvermögen 12
2.2.1 Messung der Erinnerungsleistung 15
2.3 Artificial Intelligence oder künstliche Intelligenz 15
2.3.1 Chatbots und ChatGPT 16
2.3.2 Prompting 17
2.4 Entwicklung von Generative Artificial Intelligence in der Arbeitswelt 18
2.5 Erinnerungsleistung bei Internetnutzung 19
2.5.1 Einflussfaktoren auf die Erinnerungsleistung 21
2.5.2 Informationsquellen als Gedächtnishilfe 22
3 Methodik 23
3.1 Forschungsdesign und Untersuchungsart 23
3.1.1 Variablen 25
3.1.2 Material 27
3.1.3 Versuchsplan 28
3.1.4 Trivia-Fragen 30
3.2 Pretest 31

3.3 Versuchsteilnehmende 31
3.4 Datenerhebung 32
3.4.1 Erinnerungstests 32
3.4.2 Fragebogen 33
Datenanalyse 35
3.4.3 Ethik 36
3.4.4 Interne Validität 37
3.4.5 Externe Validität 38
4 Ergebnisse 39
4.1 Beschreibung der Stichprobe 39
4.2 Deskriptive Statistiken aus Erinnerungstest 40
4.2.1 Hypothesen 45
4.2.2 Posthoc Tests 46
4.3 Deskriptive Statistiken und qualitative Ergebnisse des Fragebogens 47
4.3.1 ChatGPT und Erinnerung 47
4.3.2 Erinnerungsmethodik 48
4.3.3 Einschätzung der Wiedererkennung der Antworten bei Aufgabe 2 48
4.3.4 ChatGPT Nutzung 48
4.3.5 Speicherort Aussagen 49
4.3.6 Einstellungsveränderung 49
4.3.7 Veränderung der Erinnerung seit ChatGPT 49
5 Diskussion 51
5.1 Implikationen der Ergebnisse 54
5.2 Reflektion und Limitationen 55
5.3 Handlungsempfehlungen zur Nutzung von ChatGPT 56
5.4 Ausblick auf zukünftige Forschung 57
6 Literaturverzeichnis 60
7 Abbildungsverzeichnis 64
8 Anhang 65

1 Einleitung

Artificial Intelligence (AI) revolutioniert die Arbeitswelt, indem sie nicht nur repetitive Aufgaben übernimmt, sondern auch komplexe kognitive Tätigkeiten, deren Ausführung bisher Menschen vorbehalten war (Delcker, 2023). Mit fortschrittlichen Systemen wie ChatGPT erweitert sich das Automatisierungsspektrum deutlich. ChatGPT ermöglicht die Automatisierung von routinemässigen Aufgaben in grossem Ausmass, weil der intelligente Chatbot innert Sekunden z. B. gut formulierte Texte oder Bilder durch Prompts generieren kann. Damit dem Chatbot das gelingt, verwendet dieser Large Language Models (grosse Sprachmodelle), anhand deren Technologie ChatGPT lernt, Sprache zu «verstehen» und Sätze zu formulieren. Hierzu mehr in den theoretischen Grundlagen.
Studien zeigen, dass ChatGPT die Effizienz der Arbeitsleistung um 5–12 % steigern kann, und Experten erwarten eine weitere Zunahme dieser Zahl in den nächsten fünf Jahren (McKinsey, 2023; Statista, 2024). Die Integration von AI in den Arbeitsalltag stellt sich als unausweichlicher Prozess dar, doch ihre Auswirkungen auf menschliche Fähigkeiten sind noch ungewiss. Während AI in vielen Bereichen unterstützen kann, gibt es noch deutliche Grenzen, was menschliche und soziale Interaktion (z. B. Empathie) und ethische Entscheidungen betrifft.

Die rasante Entwicklung der AI-Technologie verspricht einen umfassenden Wandel, der zunächst einige Berufsfelder und Qualifikationsstufen mit gut automatisierbaren Tätigkeiten betrifft. Unterschiedliche Zeitschriften und Studien diskutieren den möglichen Einfluss von AI-Technologien auf die menschliche Intelligenz (Delcker, 2023).
Die Beziehung zwischen menschlicher Intelligenz und menschlichem Lernen ist vielschichtig und umfasst genetische, kognitive und emotionale Komponenten, die die Bildungsergebnisse erheblich beeinflussen (Plomin & Deary, 2015). Wie gut Informationen gespeichert werden können, beeinflusst daher stark den Lernerfolg und somit das psychologische Konstrukt des Lernens (Sternberg, 1997). Die Informationsverarbeitung im Gedächtnis bzw. die Erinnerungsleistung wird deshalb in dieser Masterarbeit fokussiert. Es soll untersucht werden, ob die Nutzung von ChatGPT die Verarbeitung der generierten Inhalte und Informationen sowie die spätere Enkodierung im Gedächtnis beeinflusst.

1.1 Forschungsfrage und Zielsetzung

Die vorliegende Masterarbeit soll Antworten darauf liefern, wie sich die ChatGPT-Nutzung auf die Erinnerungsleistung von Mitarbeitenden der Firma Swisscom auswirkt.

Die Swisscom Schweiz AG ist die Praxispartnerin dieser Masterarbeit. Sie investiert ein grosses Ausmass an Zeit und viele Ressourcen in unterschiedliche Digitalisierungsprojekte, in denen AI eingesetzt wird, um die Produktivität und Effizienz zu steigern. Die Erkenntnisse dieser Studie sollen der Swisscom mehr Handlungsspielraum bei zukünftigen AI-Integrations-Projekten bieten und sie dabei unterstützen, den Einsatz von AI-Tools im Sinne einer human-gerechten Arbeit zu optimieren.

Die Forschungsfrage der vorliegenden Masterarbeit lautet daher wie folgt:

> *Wie wirkt sich die ChatGPT-Nutzung auf die Erinnerungsleistung von Swisscom-Mitarbeitenden aus?*

Wie bereits in der Forschungsfrage angedeutet, wird davon ausgegangen, dass die Nutzung von ChatGPT einen möglichen Einfluss auf die Erinnerungsleistung hat, da sich die Nutzungsart von ChatGPT von bisherigen Mitteln der Informationsbeschaffung wie der Internetrecherche unterscheidet. Es gilt also zu untersuchen, ob die ChatGPT-Nutzung einen Einfluss auf die Informationsverarbeitung und Verarbeitungstiefe und somit auf die Erinnerungsleistung hat. Mit ChatGPT erhält die Person Antwort auf den Prompt, was im Vergleich zur Internetrecherche ein gezielteres und kompaktes Resultat ermöglicht. Bei einer Internetrecherche muss die Person in der Regel mehrere Suchergebnisse prüfen, was somit einen anderen Umgang mit den Informationen zur Folge hat.

In Hinblick darauf wurden die folgenden Arbeitshypothesen aufgestellt:

> *H0: Die Erarbeitung von Inhalten mit ChatGPT hat keinen Einfluss auf die Erinnerungsleistung der erarbeiteten Inhalte nach zwei bis drei Wochen.*
>
> *H1: Die Erinnerungsleistung in Bezug auf die mit ChatGPT erarbeiteten Inhalte ist nach zwei bis drei Wochen signifikant schlechter gegenüber den Inhalten, die ohne ChatGPT erarbeitet wurden.*

Die Literaturrecherche ergab, dass die meisten Studien zu diesem Thema mittels Interviews durchgeführt wurden. Diese basierten auf Vermutungen der jeweiligen Experten wie beispielsweise bei Bai, Liu und Su (2023). Es konnten aber keine empirischen Studien gefunden werden, die diesen Einfluss auch tatsächlich getestet haben. Deshalb wurde ein Experiment als Forschungsdesign gewählt.

Um eine Aussage zu den Hypothesen über diese zwei bis drei Wochen machen zu können, wurden an zwei Zeitpunkten eine Messung durchgeführt. Am ersten Messzeitpunkt wurde ein Experiment durchgeführt, bei dem ein Erinnerungstest stattfand, und zwei bis drei Wochen nach dem Experiment wurde ein Fragebogen versendet. Dieser beinhaltete ebenfalls einen Erinnerungstest und stellt somit den zweiten Messzeitpunkt dar. Es wurden nach der Theorierecherche noch weitere detailliertere Hypothesen zu wichtigen Komponenten erstellt, welche im Kapitel Methodik ausgeführt werden.

1.2 Relevanz der Studie und Praxispartnerin Swisscom Schweiz AG

Durch den aktuellen und schnellen Wandel der Arbeitswelt ist die Effizienz und die Produktivität, die AI verspricht, offensichtlich und bereits in Studien unter Beweis gestellt worden (Statista, 2023). Die neue Generation von AI unter Verwendung von Large-Language-Modellen hielt nun auch Einzug in den wissensbasierten Sektor. Diese Form der AI, häufig als ChatGPT bezeichnet, wird die Arbeitswelt immer stärker verändern, indem sie mehr Aufgaben übernimmt und Berufe, wie wir sie heute kennen, ablöst oder zumindest stark verändert. In erster Linie davon betroffen sind Berufe mit mehrheitlich routinemässigen oder automatisierbaren Aufgaben (beispielsweise bei repetitiven Tätigkeiten als kaufmännische angestellte Person). Es werden auch neue Berufe geschaffen, die stärker auf AI aufsetzen (Arntz, Gregory & Zierah (2016). Wie sich die stärkere Nutzung von AI auf die Kompetenzen und das Wissen des Menschen längerfristig auswirkt, ist unklar und bisher nur wenig erforscht (Bai et al., 2023; Nagam, 2023). Es zeigt sich, dass Menschen vermehrt die Speicherung von Informationen an das Internet abgeben (Nagam, 2023). Nebst Nagam (2023), der die Internetnutzung und Erinnerungsleistung experimentell erforscht hat, nahmen dieses Thema auch zwei weitere Studien in den Fokus (Sparrow et al., 2011; Storm et al., 2017). Diese insgesamt drei Studien stellen daher die Grundlage für das Forschungsdesign dieser Masterarbeit dar. Auch aus diesem Grund soll diese Masterarbeit experimentell untersuchen, ob die Nutzung von ChatGPT einen Einfluss auf die Erinnerungsleistung bezüglich der mit ChatGPT erarbeiteten Inhalte hat. Denn die Erinnerungsleistung stellt eine wesentliche Komponente der Kompetenz- und Wissensaneignung dar.

Die Praxispartnerin, die Swisscom Schweiz AG, ist eines der grössten Schweizer Unternehmen für Mobilfunk- und Festnetztelefonie sowie Information und Technologie. Das Unternehmen investiert stark in neue innovative Technologien, um wettbewerbsfähig zu bleiben. Die Swisscom konzentriert sich hauptsächlich auf den Schweizer Markt. Dort bietet das Unternehmen eine Palette von Telekommunikationsdiensten an, darunter Festnetz- und Mobilfunktelefonie, Breitbandinternet, digitales Fernsehen und IT-Dienstleistungen für Geschäfts- sowie Privatkunden. Das Unternehmen hebt sich von der Konkurrenz nach eigener Aussage durch die hohe Qualität und Innovationskraft der Produkte ab (Swisscom, 2024).

Gerade die Integration von AI-Lösungen stellt Teil zahlreicher Projektvorhaben bei Swisscom dar, um die eigene Effizienz und Produktivität zu steigern. Dabei ist es der Unternehmung wichtig, dass die Mitarbeitenden weiterhin einer würdigen und sinnvollen

Tätigkeit nachgehen können, bei der ihre Kompetenzen gefragt, genutzt und ausgebaut werden können.
Bei Swisscom spielen AI-Tools also eine zentrale Rolle, um innovativ zu sein und die Arbeitsprozesse effizienter zu gestalten. Die Mitarbeitenden nutzen diese Tools tendenziell also vermehrter als zuvor und die längerfristigen Auswirkungen sind noch zu erforschen. Gerade auch die Einflüsse auf die Erinnerungsleistung sind dabei von Interesse, wenn die Nutzung von ChatGPT bei Swisscom immer weiter steigt. Denn wie bereits erwähnt, stellt die Erinnerungsleistung eine der wesentlichen Komponenten dar für die Kompetenz- und Wissensaneignung.

1.3 Abgrenzungen

Die Begriffe Behaltensleistung und Erinnerungsleistung unterscheiden sich für die Autorin bezüglich der Wiedergabe der Inhalte. Der Terminus Behaltensleistung nimmt dabei eher das Behalten von Informationen, also eine Reproduktion, in den Fokus. Der Begriff der Erinnerungsleistung ist hingegen offener und beinhaltet sowohl die mögliche Reproduktion als auch Wiedererkennung von Inhalten. Im Experiment könnte es sein, dass sowohl reproduziert als auch wiedererkannt wird. Deshalb soll im Rahmen dieser Arbeit die Erinnerungsleistung gemessen werden.
Nicht untersucht werden die über die Erinnerungsleistung hinausgehenden Prozesse der Kompetenz- und die Wissensaneignung der Menschen bei der Nutzung von ChatGPT.
Wie eingangs erwähnt, werden in dieser Studie keine Lernprozesse oder Intelligenzen gemessen, da sich diese psychologischen Konstrukte im Rahmen einer Masterarbeit als zu umfassend darstellen, um erforscht zu werden. Allerdings bildet die Erinnerungsleistung den Kern dieser genannten Konstrukte (Lernen, Kompetenz-/Wissensaneignung und Intelligenz), weshalb nur dieses Konstrukt genauer untersucht werden soll.

1.4 Aufbau der Arbeit

In der vorliegenden Masterarbeit werden im Kapitel «Theoretische Grundlagen» als Erstes wichtige Begriffe erklärt, deren Verständnis im Rahmen dieser Studie wichtig ist. Dabei handelt es sich einerseits um technische und andererseits um psychologische Begriffe. Nach der Darstellung der Begrifflichkeiten folgt jeweils die detailliertere Vorstellung der Theorie. Diese umfasst die relevanten Gedächtnismodelle und eine Einführung in das Feld der künstlichen Intelligenz in der Arbeitswelt. Zum Schluss folgt ein Unterkapitel zu relevanten Studien im Bereich der Erforschung der Erinnerungsleistung und der Internetnutzung.

Des Weiteren wird das methodische Vorgehen der Studie unter Einbezug der Literatur vorgestellt. Nach der Erläuterung der Datenerhebung werden die Ergebnisse dieser Masterarbeit präsentiert. Zu guter Letzt werden diese Ergebnisse in der Diskussion erneut zur Theorie und Beobachtung der Erhebung in Bezug gesetzt, interpretiert und reflektiert.

2 Theoretische Grundlagen

In diesem Kapitel werden zuerst wichtige Begriffe und Konstrukte zu unterschiedlichen Gedächtnisspeichern beschrieben und im Anschluss daran wird detailliert auf Gedächtnismodelle sowie das Erinnerungsvermögen eingegangen. Daraufhin werden wichtige Begriffe für den Einstieg in die Entwicklung von Generative Artificial Intelligence in der Arbeitswelt vorgestellt. Der Hauptfokus des Theoriekapitels liegt darauf, für die Studie relevante Erkenntnisse vorzustellen, auf denen das Forschungsdesign basiert.

2.1 Die unterschiedlichen Gedächtnisspeicher

Da die Themen Gedächtnis, Informationsspeicherung und der Abruf von Informationen (Erinnerung) in dieser Arbeit eine zentrale Rolle spielen, werden nachfolgend kurz das sensorische Gedächtnis, das Kurz- und Langzeitgedächtnis sowie das Konzept des impliziten Wissens vorgestellt. Als Gedächtnis wird dabei die Fähigkeit eines Organismus bezeichnet, Informationen aufnehmen, speichern und bei Bedarf wieder abrufen zu können (Raab, Unger & Unger, 2022).

2.1.1 Das sensorische Gedächtnis

Laut Gruber (2018) speichert das sensorische Gedächtnis kurzzeitig (0.1–1.0 Sek.) eingehende Informationen, was als Ultrakurzzeitgedächtnis bezeichnet wird. Die perzeptuellen Systeme, wie unser Hör- und Sehsystem, transformieren sensorische Informationen, die von verschiedenen Rezeptoren erfasst werden, in perzeptuelle Darstellungen. Nur die Informationen, die Aufmerksamkeit erregen, werden weiterverarbeitet und können zu langfristigen Erinnerungen werden.

2.1.2 Das Kurzzeitgedächtnis

Das Kurzzeitgedächtnis, auch Arbeitsgedächtnis genannt, speichert Informationen aus dem sensorischen Gedächtnis bis zu einer Minute lang (Raab et al., 2022). Laut Raab et al. (2022) selektiert und verarbeitet es diese Eindrücke, bevor sie ins Langzeitgedächtnis übergehen, und greift in diesem Prozess auch auf bereits gespeicherte Langzeiterinnerungen zurück. Dabei ist auch die Verarbeitungstiefe wichtig, sodass Informationen vertieft enkodiert werden können (Elaboration) (Gruber, 2018). Das Kurzzeitgedächtnis fungiert als Bindeglied zwischen sensorischem und Langzeitgedächtnis.

2.1.3 Das Langzeitgedächtnis

Laut Raab et al. (2022) gelangen vom Kurzzeitgedächtnis nur ausgewählte Informationen ins Langzeitgedächtnis, wo sie dauerhaft gespeichert werden. Das Langzeitgedächtnis hat

eine unbegrenzte Kapazität und repräsentiert das eigentliche menschliche Gedächtnis. Es wird zwischen explizitem (deklarativem) und implizitem (non-deklarativem) Wissen unterschieden (Hobmair, 2013).

2.1.3.1 Implizites oder non-deklaratives Wissen

Implizites Wissen bezieht sich auf unbewusste Fähigkeiten und Abläufe, wie das Fahren oder Skifahren (Raab et al., 2022). Es umfasst drei Hauptbereiche: Bewegungs- und Handlungsabläufe, Priming und verschiedene Konditionierungsformen (Raab et al., 2022).

2.1.3.2 Explizites oder deklaratives Wissen

Raab et al. (2022) definieren das deklarative bzw. explizite Wissen als das gesamte Faktenwissen eines Menschen, das in der Regel unmittelbar zugänglich ist, wie zum Beispiel Schulwissen. Hierbei wird zwischen einem episodischen Gedächtnis, das alle persönlichen Erlebnisse oder Erfahrungen enthält, wie zum Beispiel einen Urlaub mit Freunden, und einem semantischen Gedächtnis, auch als Wissenssystem bezeichnet, das das Faktenwissen enthält, unterschieden.

2.1.4 Das transaktive Gedächtnis

Laut Nagam (2023) haben die Menschen im Laufe der Geschichte und der Zivilisation gelernt, sich aufeinander zu verlassen, um kognitive Aufgaben, wie das Erinnern komplexer Informationen, zu teilen. Das hat dazu geführt, dass sich die Gesellschaft auf Beziehungen spezialisiert hat. Das transaktive Gedächtnis bezieht sich auf das Konzept, dass Menschen ein «System der Kodierung» entwickeln, was die Speicherung und den Abruf von Informationen aus verschiedenen Wissensbereichen betrifft. Diese Art des Gedächtnisses bezieht einerseits die Quelle der Informationen als auch das Wissen darüber mit ein, wie man auf diese Informationen zugreifen oder sie abrufen kann, wenn sie benötigt werden. Folglich werden eben diese «Transaktionen» zwischen Menschen oder zwischen Wissensdatenbanken als transaktives Gedächtnissystem verstanden.

2.2 Gedächtnis und Erinnerungsvermögen

Das Gedächtnis konzentriert sich auf den Informationsfluss in und aus den Gedächtnissystemen (Raab et al., 2022). In der Psychologie werden zwei Hauptmodelle zur Darstellung dieser Prozesse unterschieden: das Mehr-Speicher-Modell und das Ein-Speicher-Modell (Ebbinghaus, 1885).
Laut Stangl (2024) wurde das Ein-Speicher-Modell von der Gedächtnispsychologie in den 1960er Jahren entwickelt und stellt ein sehr vereinfachtes Modell dar. Es wurde von einem Gedächtnisspeicher ausgegangen, in dem Informationen undifferenziert abgelegt werden. In Ebbinghaus' Forschung standen vor allem die Leistungsfähigkeit und die Güte des

Speichers bei den Veränderungen des Lerninhalts im Vordergrund (Stangl, 2024). Das Mehr-Speicher-Modell wurde von Richard Atkinson und Richard Shiffrin entwickelt (Myers, Hoppe-Graff & Myers, 2014).

Atkinson und Shiffrin haben das Mehr-Speicher-Modell in dem Sinne weiterentwickelt, indem sie mehrere Gedächtnisspeicher unterschieden, als es im simplen Ein-Speicher-Modell der Fall war (Stangl, 2024). So werden nunmehr die verschiedenen Gedächtnisspeicher unterschieden, die für die Kodierung und den Abruf der aufgenommenen Inhalte oder Informationen zuständig sind (Stangl, 2024). In der nachfolgenden *Abbildung 1* wird der Prozess der Informationsverarbeitung im Gedächtnis gut verbildlicht. Gemäss diesem Prozess durchlaufen die entsprechenden Informationen die verschiedenen Speicher im Verlauf der Zeit. Im Modell geht es nicht primär darum, die verschiedenen Gedächtnisspeicher zu trennen, sondern darum, die unterschiedliche Gedächtnisleistung von Menschen mit den verschiedenen Stufen der Informationsverarbeitung zu begründen (Stangl, 2024).

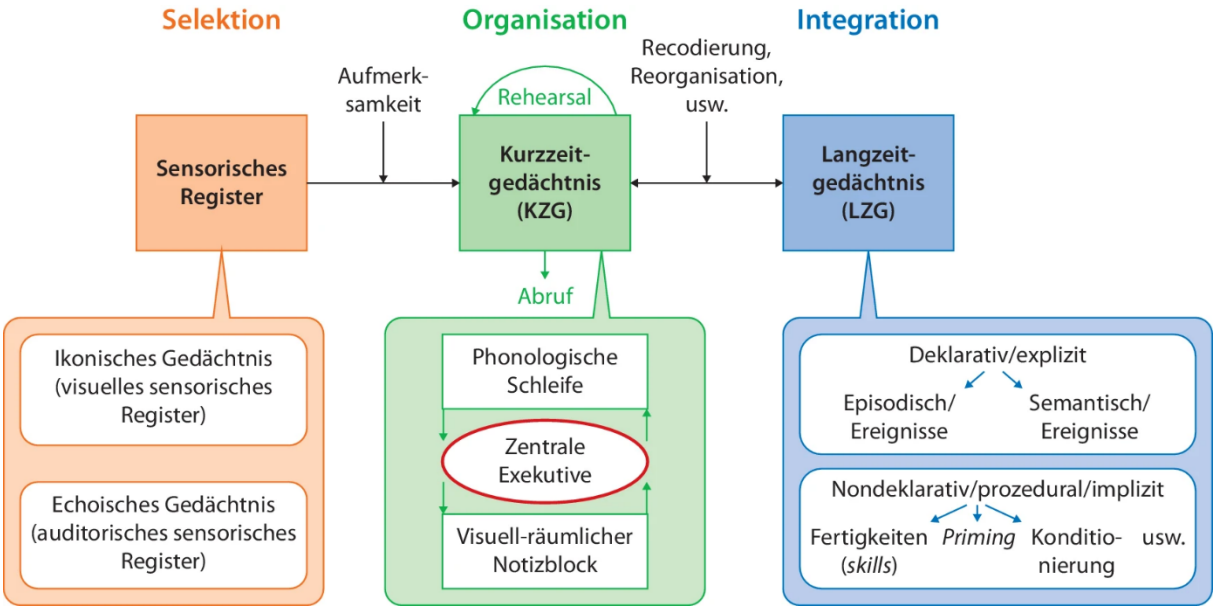


Abbildung 1: Mehr-Speicher-Modell von Atkinson und Shiffrin (Zoelch et al., 2019)

Zunächst unterschied man bei den Mehr-Speicher-Modellen nur zwei Gedächtnisspeicher, das Kurz- und das Langzeitgedächtnis (Stangl, 2024). Es wurde festgestellt, dass Informationen durch mehrmaliges Wiederholen (Rehearsal) aus dem Kurzzeitgedächtnis in das Langzeitgedächtnis übermittelt werden können (Stangl, 2024). Das Kurzzeitgedächtnis hat jedoch nur eine begrenzte Kapazität, weshalb nur ein Teil der Information übertragen werden kann (Myers et al., 2014).

Atkinson und Shiffrin haben das im Jahr 1968 entwickelte Zweispeichermodell um einen dritten Speicher ergänzt (deshalb Mehr-Speicher-Modell) (Stangl, 2024). Dieser dritte Speicher ist das sensorische Register, das vor das Kurzzeitgedächtnis geschaltet ist. Die drei Speicher unterscheiden sich in Bezug auf die Menge der Informationen, die verarbeitet werden kann, der Verfügbarkeitsdauer und der Verschlüsslung der Information (Kodierung) (Stangl, 2024).

Baddeley, Eysneck, Anderson & Anderson (2010) kritisierten das Mehr-Speicher-Modell jedoch, weil das Kurzzeitgedächtnis mehr ist als nur ein vorübergehender Speicherort für Informationen, es stellt ihrer Meinung nach eine «aktive Arbeitsfläche» dar. Baddeley et al. (2010) zeigen, dass auf dieser Ebene neue Informationen intensiv verarbeitet und mit bestehenden Langzeiterinnerungen verknüpft werden. Alle drei Modelle versuchen, die Informationsverarbeitung aus verschiedenen Blickwinkeln zu beleuchten. Die aktuellste Theorie zur Informationsverarbeitung bezieht sich ausschliesslich auf die erwähnten Gedächtnismodelle (Raab et al., 2022).

Unabhängig vom Typ des Gedächtnissystems sind drei geistige Prozesse für die Nutzung von gespeicherten Informationen zu einem späteren Zeitpunkt erforderlich: die ***Enkodierung,*** die ***Speicherung*** und der **Abruf** (Raab et al., 2022).

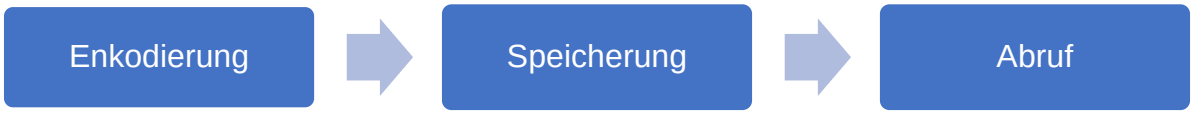


Abbildung 2: eigene Darstellung (2024)

Die Enkodierung ist der erste Schritt in der Informationsverarbeitung und resultiert in einer Gedächtnisrepräsentation (Myers et al., 2014). Die Speicherung beinhaltet das Bewahren von enkodierten Informationen über eine bestimmte Zeitspanne hinweg (Myers et al., 2014). Der Abruf bezieht sich auf das Wiedererlangen gespeicherter Informationen zu einem späteren Zeitpunkt (Myers et al., 2014).

Laut der Interferenztheorie wird die spätere Wiedergabe einer gelernten Information gehemmt, wenn sie von vorher und nachher gespeicherten Informationen überlagert wird, was als Interferenz bezeichnet wird (Gruber, 2018).

Die bewusste Anwendung von Gedächtnisstrategien spielt eine entscheidende Rolle für das Behalten und Wiedergeben von Informationen (Gruber, 2018). Diese Strategien dienen dazu, die Gedächtnisleistung zu verbessern. Die Fähigkeit, Informationen zu speichern und

bei Bedarf abzurufen, variiert von Person zu Person (Stuck, 2020). Einige Menschen können sich problemlos an Informationen erinnern, während andere Schwierigkeiten haben. Oft wird dies auf ein «gutes» oder «schlechtes» Gedächtnis zurückgeführt. Studien zeigen jedoch, dass das Problem nicht in der Kapazität des Gedächtnisses liegt, sondern darin, dass Menschen mit schlechtem Erinnerungsvermögen ihr Gedächtnis nicht optimal nutzen, da sie zu wenige Gedächtnisstrategien kennen (Gruber, 2018). Zwei dieser Strategien sind sogenannte Mnemotechniken, auch bekannt als «Eselsbrücken» und Chunking, das Zusammenfügen oder die Organisation von Informationen oder Lerninhalten in sinnhafte Gruppen (Gruber, 2018). Vergessen ist oft das Ergebnis von falsch kodierten Erinnerungen oder Problemen beim Abrufen von Informationen (Stuck, 2020).

2.2.1 Messung der Erinnerungsleistung

Für die Messung von Erinnerungsleistung werden die genaue Reproduktion oder Wiedererkennung von Inhalten einer Aufgabe gemessen (Raab et al., 2022).
Die ersten getesteten Methoden zur Messung gespeicherter Informationen lassen sich in direkte und indirekte Verfahren unterteilen.
Direkte Verfahren zielen darauf ab, gespeicherte Informationen zu extrahieren.
Indirekte Verfahren basieren auf der Beobachtung, dass der Aufwand, der benötigt wird, um dargestelltes Material – verbaler oder bildlicher Natur – wieder zu lernen, stark davon abhängt, wie viel von dem gelernten Material noch im Gedächtnis vorhanden ist.
Daher wird das Behalten oder Vergessen indirekt durch den für das Wiederlernen erforderlichen Aufwand gemessen. Dabei werden Methoden wie die freie Reproduktion (Free Recall), die unterstützte Reproduktion (Aided Recall) oder das Wiedererkennen (Recognition) angewendet.

2.2.1.1 Reproduktion und Wiedererkennung

Laut Raab et al. (2022) wird bei der Methode der freien Reproduktion von der Person erwartet, dass sie das Gelernte ohne jegliche Unterstützung wiedergibt. Im Gegensatz dazu erlaubt die Methode der unterstützten Reproduktion den Einsatz von Gedächtnishilfen, wie zum Beispiel Hinweisen auf den Kontext des gelernten Materials. Eine weitere Methode ist das Wiedererkennen, bei dem der Person das gelernte Material zusammen mit anderem Material präsentiert wird. Hierbei wird untersucht, an welches Material sich die Person erinnert.

2.3 Artificial Intelligence oder künstliche Intelligenz

In dieser Masterarbeit wird vorwiegend der englische Begriff Artificial Intelligence (AI) statt des deutschen Begriffs Künstliche Intelligenz (KI) verwendet.

Nach Krauss (2023) beschreibt AI die Fähigkeit von Maschinen, komplexe Aufgaben auszuführen, die gewöhnlich menschliche Intelligenz erfordern. Dazu gehören Mustererkennung, Erfahrungslernen, Entscheidungsfindung und Problemlösung. Innerhalb des AI-Feldes ist das maschinelle Lernen ein spezifischer Bereich, der Maschinen ermöglicht, aus Daten zu lernen und sich kontinuierlich zu verbessern und weiterzuentwickeln.

AI ist ein Überbegriff für die Intelligenzleistung von Maschinen, wie Computern oder Robotern. Das Ziel der Weiterentwicklung von AI ist es, die Maschine intelligent zu machen (Simplilearn.com, 2024).

Die AI ahmt die menschliche Intelligenz nach, indem sie Informationen durch Algorithmen und Datenanalyse verarbeitet. Dabei fehlt es der AI an emotionaler Intelligenz und Kreativität (Bhandari, 2022). Menschliche Intelligenz entwickelt sich durch Lebenserfahrung und soziale Interaktionen weiter, wobei die AI zur Weiterentwicklung andauerndes Datentraining oder Anpassungen im Programmcode benötigt (Bhandari, 2022).

Laut Krauss (2023) ist tiefes Lernen (Deep Learning) eine Unterdisziplin des maschinellen Lernens, die sich auf komplexe «neuronale» Netzwerke konzentriert. Diese Netzwerke sind in der Lage, tiefe Strukturen in grossen Datenmengen zu erkennen und zu analysieren. Obwohl Deep Learning erhebliche Fortschritte in Bereichen wie der Bild- und Spracherkennung sowie dem autonomen Fahren ermöglicht hat, bleiben dennoch Herausforderungen hinsichtlich der Interpretation und Erklärung der Funktionsweise dieser Modelle bestehen, wie zum Beispiel bei tiefen neuronalen Netzen, um Entscheidungen nachzuvollziehen oder Muster zu erkennen (Krauss, 2023).

2.3.1 Chatbots und ChatGPT

Ein Chatbot ist ein Computerprogramm, das Artificial Intelligence und natürliche Sprachverarbeitung (NLP) nutzt, um freitextliche Anfragen des Nutzenden zu verstehen und die Antworten darauf zu automatisieren, indem es menschlichen Datenaustausch simuliert (IBM, 2023). Häufig werden Chatbots mit NLP auf Webseiten für einfache Kundenanfragen verwendet (IBM, 2023).

Je nach der Anbindung an andere Systeme oder Datensätze kann der Chatbot ganze Prozesse automatisieren und Kundenanfragen kategorisieren und entsprechend verarbeiten oder weitergeben. Die Qualität der Anfrageverarbeitung hängt von der Datenqualität ab, mit der der Chatbot trainiert wird (IBM, 2023).

ChatGPT wurde von OpenAI entwickelt und ist ein GAI-Chatbot, GAI steht für Generative Artificial Intelligence (CHIP, 2023; Haase & Hanel, 2023). Die GAI ist ein innovativer Ansatz, der es Maschinen ermöglicht, aus zuvor gesammelten Daten zu lernen und sich an neue Situationen anzupassen (Haase & Hanel, 2023).

ChatGPT ist ein intelligenter Chatbot, mit dem Nutzende durch Text-Nachrichten oder Sprachaufzeichnungen kommunizieren können. ChatGPT verwendet eine moderne maschinelle Lerntechnologie, das Large-Language-Modell (LLM), übersetzt bedeutet das: grosses Sprachmodell. Mit diesem Modell gelingt es ChatGPT, natürlich klingende Gespräche zu führen und Antworten zu generieren, die für den Nutzenden relevant sind. Das «GPT» in ChatGPT steht für «Generative Pre-Trained Transformer». Die AI lernt nach dem Ansatz des Transformer-Netzwerks maschinell (ChatGPT, 2023). Das Netzwerk ist darauf spezialisiert, Text zu «verstehen» und zu generieren. ChatGPT ist also vortrainiert, da es grosse Mengen an Text aus dem Internet analysiert und so ein Sprachverständnis und menschliche Kommunikationsstrukturen entwickelt hat. Hierbei greift der Chatbot auf Auftretenswahrscheinlichkeiten zurück. Er erkennt Muster in der Sprache, wobei der Kontext eine wichtige Rolle spielt, damit dieser den Prompt auch einordnen kann (Hundertmark, 2023).
Zurzeit ist die ChatGPT-Version 3.5 für alle Menschen mit Internetzugriff kostenlos verfügbar. Die Daten von «ChatGPT 3.5» reichen bis zum Jahr 2022 zurück. Die Daten der neusten kostenpflichtigen Version «ChatGPT 4.5» reichen hingegen bis zum Jahr 2023 (ChatGPT, 2023). Der grösste Unterschied zwischen den zwei aktuell zugänglichen Versionen liegt in ihrer Leistungs- und Anwendungsfähigkeit. ChatGPT bietet nunmehr eine erhöhte Schlussfolgerungs- und Problemlösefähigkeit (ChatGPT, 2023). Im Sommer 2024 wird die aktuellste Version von ChatGPT lanciert (Edwards, 2024).

2.3.2 Prompting

Als Prompting werden «Befehle», «Anweisungen» oder «Anfragen» bezeichnet, die einer AI gestellt werden (KI-Cafe, 2023). Ein Prompt definiert die spezifische Aufgabe, die von der AI ausgeführt werden soll. Er kann wichtige Hintergrundinformationen, Beispiele und andere weiterführende Informationen enthalten, die dem Chatbot oder der AI helfen, die Aufgabe zu verstehen und sie effektiver zu bearbeiten (KI-Cafe, 2023).
Es ist empfehlenswert, die Antworten von AII-Chatbots wie ChatGPT zu prüfen, da ChatGPT Informationen generieren kann, die nicht auf Fakten basieren (Marr, 2023). Dieses Phänomen wird auch «AI-Halluzination» genannt (Marr, 2023). Es ist wichtig zu betonen, dass ChatGPT nicht bewusst *lügt* (Marr, 2023). Der GAI-Chatbot hat kein Bewusstsein oder Verständnis für die Bedeutung der Texte, die er generiert. Daher kann es vorkommen, dass ChatGPT Antworten als Tatsachen darstellt, obwohl diese nicht der Wahrheit entsprechen (Marr, 2023). Gründe für künstliche Halluzination können zum Beispiel Abweichungen in Quellen sein oder die Art und Weise, wie der Chatbot mit Daten trainiert wurde (Marr, 2023).

2.4 Entwicklung von Generative Artificial Intelligence in der Arbeitswelt

Bislang wurden hauptsächlich repetitive Aufgaben an die AI übergeben (die Ablösung von Wissensarbeit), unter der Kontrolle eines Menschen ausgeführt oder als Hilfsmittel eingesetzt (Delcker, 2023). Solche routinemässigen Aufgaben können z. B. Recherchen oder Tätigkeiten in Fertigungsberufen sein (Fregin & Stops, 2023).
Gleichzeitig hat der digitale Wandel in spezifischen Bereichen zu einem Beschäftigungsanstieg geführt (Bundesrat, 2017). Laut Bundesrat (2017) besonders in technologie- und wissensintensiven Berufen sowie in Berufen mit einem hohen Anteil an kreativen Aufgaben und unternehmerischem, konzeptionellem oder abstraktem Denken. Auch körperlich anstrengende oder gefährliche Tätigkeiten können vermehrt von Robotern oder Systemen übernommen werden (Bundesrat, 2017). Dadurch wird das Risiko von Arbeitsunfällen gesenkt und neue sinnstiftende Aufgabengebiete können eröffnet werden (Bundesrat, 2017).
AI-Lösungen sind daher kaum mehr aus der Arbeitswelt wegzudenken. Neue AI-Systeme wie ChatGPT zeigen jedoch, dass sich das Spektrum der Aufgaben, die automatisiert werden können, in Zukunft erheblich erweitern wird (Schaller et al., 2023). Während Computer- und Robotersysteme hauptsächlich bestimmte manuelle und Routineaufgaben automatisieren, können AI-Systeme nun auch kognitive und nicht routinemässige Aufgaben übernehmen, die bisher qualifizierte kognitive Arbeit von Menschen erforderten und deren Lösung zum Teil über die menschlichen Fähigkeiten hinausgeht. Im Falle von Large-Language-Modellen (LLM oder neuronalen Netzen, die menschliche Sprache auf der Grundlage sehr grosser Datenmengen verstehen), die hinter ChatGPT und anderen AI-Anwendungen stehen, geschieht dies beispielsweise in Bereichen, in denen die Arbeit auf dem Sammeln und (Wieder-)Zusammensetzen einer grossen Menge von Informationen, zum Beispiel Bildern, Texten oder Programmcodes, beruht (Fregin & Stops, 2023). So werden viele routinemässige Tätigkeiten automatisiert und Kompetenzen, die für solche Systeme nötig sind, werden Teil der heutigen Jobanforderungen (Bundesrat, 2017). Berufe, die sich durch ein hohes Ausmass an sozialer Interaktion und durch Nichtroutineaufgaben auszeichnen, sind laut Bundesrat weniger stark von der Automatisierung betroffen (2017).

ChatGPT kann nachweislich die Effizienz und Produktivität steigern (Statista, 2023). Es ist somit sehr wahrscheinlich, dass Menschen sich in Zukunft vermehrt mit AI auseinandersetzen müssen und nicht vermeiden können, sie für gewisse Aufgaben zu nutzen. Laut Schaller et al. (2023) werden AI-Technologien zurzeit in einer Geschwindigkeit (weiter-)entwickelt, die weitaus schneller ist als je zuvor, sodass in immer kürzerer Zeit neue Systeme verfügbar gemacht werden, die eine Vielzahl von Aufgaben automatisieren und

unterstützen können. Dies könnte zu grundlegenden Veränderungen in der Arbeitswelt führen, die nicht nur inkrementell erfolgen, sondern sich auch über verschiedene Branchen und Qualifikationsniveaus hinweg erstrecken. Allerdings gibt es Bereiche, die bislang noch nicht automatisiert werden können. Dies betrifft zum Beispiel die Bereiche der Intuition, der ethisch-moralischen Abwägungen, der menschlichen Sozialisierung oder auch der persönlichen Präferenzen.

Hünebeck (2024), Dozent und Speaker an der Fachhochschule Graubünden für künstliche Intelligenz, ist der Meinung, Menschen bei der Arbeit müssten lernen, mit AI umzugehen, wenn sie wettbewerbsfähig bleiben möchten. Die Stärken und Schwächen von AI oder spezifisch von ChatGPT zu kennen, kann längerfristig ein wichtiges Asset für Unternehmen sein. Es ist somit nicht zielführend, Angst vor dem Wandel und der Nutzung von AI zu haben, sondern eher gewinnbringend, sich damit zu konfrontieren und auseinanderzusetzen. Denn ChatGPT sowie andere effiziente AI-Tools werden unseren Arbeitskontext zunehmend stärker weiterentwickeln.

2.5 Erinnerungsleistung bei Internetnutzung

In Rahmen dieser Masterarbeit wurde nebst den allgemeinen Theorien zu den Gedächtnissystemen auch Literatur zur Erinnerungsleistung in Zusammenhang mit ChatGPT gesucht. Zum gegenwärtigen Zeitpunkt sind keine Studien vorhanden, die die Erinnerungsleistung mit Bezug auf die Nutzung von ChatGPT untersucht haben. Die einzigen relevanten Studien finden sich im Bereich von Erinnerungsleistung und dem Internet. Es wurden drei relevante Studien identifiziert, die darauf hindeuten, dass Menschen, wenn sie mit schwierigen Fragen konfrontiert sind, eher an Computer zur Internetrecherche denken (Nagam, 2023; Sparrow et al., 2011; Storm et al., 2017). Ausserdem wurde festgestellt, dass sie sich weniger gut an gesuchte Informationen erinnern können, wenn sie erwarten, dass sie in Zukunft Zugang zu diesen Informationen haben werden; stattdessen wissen sie eher, wo sie diese finden können (Nagam, 2023; Sparrow, Liu & Wegner, 2011; Storm, Stone & Benjamin, 2017). Das Internet ist zu einer primären Form des externen oder transaktiven Gedächtnisses geworden, in dem Informationen *kollektiv ausserhalb von uns* selbst gespeichert sind (Nagam, 2023). Deshalb ist dieses Gedächtnissystem nicht in der *Abbildung 1* zu verorten, welches den Gedächtnis- und Informationsspeicherprozess eines Individuums darstellt (Nagam, 2023).
Zwei Gedächtnissysteme, die aus den Literaturrecherchen zur Erinnerungsleistung bei der Nutzung eines Tools mit Zugriff auf unendlich viel Wissen, wie zum Beispiel des Internets,

wichtig erscheinen, sind das semantische Gedächtnis und das transaktive Gedächtnissystem.

Wegner, Erber und Raymond (1991) halten fest, dass die kognitiven Prozesse, die der mentalen Arbeitsteilung in transaktiven Gedächtnissystemen zugrunde liegen, nach effizienzbezogenen Auswahlkriterien funktionieren. Die Struktur dieser Systeme wird im Allgemeinen anhand von zwei Schlüsselprinzipien definiert: der relativen Kompetenz und dem Zugang zu Informationen.

Gemäss Nagam (2023) zeigt die Literatur, dass Menschen in der Lage sind, das Internet als transaktive Gedächtnisquelle zu nutzen, jedoch sind die Entstehungsmechanismen solcher transaktiven Gedächtnissysteme noch nicht umfassend erforscht. Ebenfalls sind die Auswirkungen des Internets auf das transaktive und das semantische Gedächtnis noch relativ unbekannt. Im Zusammenhang mit ChatGPT könnte dies bedeuten, dass auch hier die Nutzung des AI-Tools Auswirkungen auf das transaktive und das semantische Gedächtnis hat.

Nachfolgend werden die relevanten Aspekte dieser drei Studien kurz zusammengefasst: Sparrow et al. (2011) waren die ersten, die die Auswirkungen der Internetnutzung in einem Experiment erforscht haben. Die Probanden mussten dabei in der ersten Phase des Experiments Trivia-Aussagen erinnern. Diejenigen, die glaubten, dass die Trivia-Aussagen von einem Computer gespeichert wurden, erinnerten sich nicht an die Aussagen, selbst wenn sie ausdrücklich dazu angewiesen wurden, sie sich zu merken. Die Resultate der Studie deuten also darauf hin, dass Menschen intuitiv und automatisch die Verantwortung für das Erinnern von Informationen an das Internet abgeben. Diese Tendenz ist so stark, dass selbst die explizite Anweisung, sich etwas zu merken, unwirksam ist. Dieses spricht für die Existenz eines transaktionalen Gedächtnisses, bei dem die Verantwortung der Speicherung des Inhalts an das Internet abgegeben wird.

Die Studie von (Storm et al., 2017) untersuchte an 60 Studierenden der University of California, Santa Cruz, wie Menschen dazu gebracht werden, das Internet als eine Form kognitiver Entlastung zu nutzen. In drei Experimenten wurde gezeigt, dass die Nutzung des Internets zum Abrufen bestimmter Informationen die Neigung einer Person erhöht, das Internet auch zum Abrufen anderer Informationen zu nutzen. Insbesondere jene Teilnehmenden, die Google zur Beantwortung einer Reihe von schwierigen Trivia-Fragen genutzt hatten, entschieden sich in weiterer Folge eher dazu, Google auch zur

Beantwortung einer neuen Reihe von Trivia-Fragen heranzuziehen, als Teilnehmende, die die ersten Fragen aus dem Gedächtnis beantwortet hatten. Die Autoren kommen zu dem Schluss, dass die Nutzung des Internets zur Informationsbeschaffung die Wahrscheinlichkeit erhöht, dass die Menschen das Internet auch in Zukunft als bevorzugte Informationsquelle nutzen werden. Zusammengefasst: Wer einmal begonnen hat, das Internet als Informationsquelle zu nutzen, wird dies auch in Zukunft tun.
Die Studie von Nagam (2023) untersuchte mit 20 Probanden die zunehmende Nutzung des Internets im Zusammenhang mit Gedächtnis und Kognition. Diese Studie umfasste zwei experimentelle Erhebungsphasen mit Gedächtnisaufgaben. Die Studienergebnisse der ersten Phase zeigten auf, dass die Abrufraten aus dem Gedächtnis niedriger sind, wenn erwartet wird, dass die Informationen gespeichert und abrufbar sind. Dies zeigte sich, wie schon bei Sparrow, Lio und Wegner (2011), unabhängig von expliziten Gedächtnisanweisungen.

Zusammenfassend legen die berichteten Studienergebnisse die Erwartung nahe, dass auch die Arbeit mit ChatGPT zu einer Verantwortungsübertragung der erarbeiteten Inhalte an das AI-Tool führen wird. Entsprechend würden dann die Erinnerungsleistungen bezüglich der erarbeiteten Inhalte schlechter ausfallen. Andererseits sind die Menschen durch das Prompten stärker in die Erarbeitung der Inhalte involviert, als dies bei den Trivia-Aufgaben in den Studien von Nagam (2023) und Sparrow et al. (2011) der Fall war. Dies könnte einer Verantwortungsübertragung und damit auch einem negativen Einfluss auf die Erinnerungsleistung entgegenwirken.

2.5.1 Einflussfaktoren auf die Erinnerungsleistung

Durch eine weitere Literaturrecherche sollten vor der experimentellen Erhebung mögliche Einflüsse auf die Erinnerungsleistung ermittelt werden.

Wie bereits erwähnt, haben (Sparrow et al., 2011) herausgefunden, dass Informationen weniger gut erinnert werden können, wenn der Ablage- oder Speicherort von Informationen bekannt und zugänglich ist.
Nach Bernhardt, Maurer und Frölich (2002) ist es leichter, Informationen im Gedächtnis zu behalten, wenn eine emotionale Verknüpfung vorhanden ist. Positive Emotionen fördern dabei das Behalten einer Information mehr als negative Emotionen.
Zudem spielt auch der Gemütszustand laut (Bernhardt et al., 2002) eine Rolle für die Erinnerungsleistung. Etwas, was man in einem glücklichen Zustand gelernt hat, ist schwieriger zu erinnern, wenn man traurig ist, und umgekehrt. Weiter unterstützt ein guter körperlicher und seelischer Zustand das Behalten von Informationen.

Die intrinsische Motivation, gewisse Informationen zu erlangen, verbessert die Fähigkeit, diese auch zu behalten (Bernhardt et al., 2002).
Auch eine ruhige Arbeitsumgebung, eine Umgebung ohne Lärm, Geräusche, Musik oder Hektik, fördert das Behalten von Informationen (Bernhardt et al., 2002).
Es wurde festgestellt, dass durch den Kontext der Lernerfolg positiv beeinflusst werden kann (Müsseler & Rieger, 2017). Das nennt sich Kontextinterferenzeffekt. Der Lernerfolg wird durch Interferenzen im Kontext beeinflusst. Man unterscheidet in diesem Zusammenhang geblockte Übungen (niedrige Kontextinterferenz), bei den identischen Ausführungen einer Übung ohne Wechsel erfolgen, und randomisierte Übungen (hohe Kontextinterferenz), die sich durch ständige Aufgabenwechsel im Kontext der Ausführung auszeichnen. Eine erhöhte Kontextinterferenz beeinflusst das Lernen stärker als ein tiefe (Müsseler & Rieger, 2017).
Beim Behalten von Informationen kann auch der Interferenzeffekt einen Einfluss haben. Wie im Kapitel «Gedächtnis und Erinnerungsvermögen» besprochen, lassen sich Primacy-Effekte dem Langzeitgedächtnis und Recency-Effekte dem Kurzzeitgedächtnis zuschreiben (Gruber, 2018). So können sich vorher oder später gelernte Informationen gegenseitig beeinflussen.

2.5.2 Informationsquellen als Gedächtnishilfe

Wie bereits im Kapitel zum transaktiven Gedächtnis beschrieben, zeichnet sich das transaktive Gedächtnis dadurch aus, dass Informationen und Wissen an entsprechende Orte wie z. B. Bibliotheken oder an Personen mit entsprechendem Fach- oder Expertenwissen abgegeben werden (Nagam, 2023). Das geschieht, um die kognitiven Ressourcen für die eigene Spezialisierung zu benutzen. Das Teilen der jeweiligen Informationen wird dann als «Transaktion» bezeichnet (Nagam, 2023).
Wenn das transaktive Gedächtnis nun in Zusammenhang mit den besprochenen Studien (Nagam, 2023; Sparrow et al., 2011; Storm et al., 2017) gestellt wird, wird klar, dass dieses Konzept des transaktiven Gedächtnisses auf das Internet angewendet werden kann. Menschen wenden sich mit spezifischen Anfragen und Anliegen an das Internet. Die eigentlichen Informationen werden auch nicht mehr kognitiv abgespeichert, da das Internet der Ort der Ablage ist und dieses unlimitiert zugänglich bleibt (Storm et al., 2017). Dementsprechend merken sich Menschen wie auch in den oben beschriebenen Studien eher den genauen Speicherort als die Informationen selbst, insofern diese Informationen an dem Speicherort zugänglich bleiben. Es ist nun zu erforschen, ob sich dies mit der ChatGPT-Nutzung gleich verhält.

3 Methodik

Im nachfolgenden Kapitel wird das methodische Vorgehen dieser wissenschaftlichen Forschungsarbeit erläutert.

3.1 Forschungsdesign und Untersuchungsart

Für die Erhebung der Daten wurde ein quantitativer Forschungsansatz und als Forschungsdesign das Experiment gewählt. Wie bereits eingangs in der Theorie erwähnt, gibt es bis dato keine empirischen Untersuchungen zur Erinnerungsleistung in Zusammenhang mit der Nutzung von ChatGPT. Alle vorhandenen Studien basieren auf Interviews, die sich auf Vermutungen der jeweiligen Experten beziehen (Bai et al., 2023). Aus diesem Grund sollten neue experimentelle Erkenntnisse hierzu gefunden werden.

Es wurde ein Erinnerungstest an zwei Messzeitpunkten durchgeführt. Der erste Messzeitpunkt fand unmittelbar nach einer Rechercheaufgabe im Rahmen des Experiments statt, die später ausführlicher erklärt wird. Den zweiten Messzeitpunkt stellt eine nachfolgende Fragebogenerhebung dar, bei der der Bogen innerhalb von zwei bis drei Wochen nach dem Experiment ausgefüllt wurde. Diese zeitliche Trennung sollte sicherstellen, dass etwas langfristigere Erinnerungsleistungen erfasst werden konnten.

Zur Beantwortung der Forschungsfrage «*Wie wirkt sich die ChatGPT-Nutzung auf die Erinnerungsleistung von Swisscom-Mitarbeitenden aus?*» wurde eine intensive Literaturrecherche gestartet. Dabei wurden drei wichtige Studien von Sparrow et al. (2011), Storm et al. (2017) und Nagam (2023) identifiziert, die im Theorieabschnitt bereits näher erläutert wurden. Der Aufbau des Studiendesigns lehnt sich stark an den Aufbau dieser Studien an, um eine bessere Reliabilität und Validität der Studie zu gewährleisten.

Die Studien stellen fest, dass die Behaltensleistung vor allem dadurch beeinflusst wird, ob die Probanden glauben, dass die Informationen an einem Ort gespeichert werden und somit weiterhin für sie zugänglich bleiben. Aus diesem Grund wurde diese Komponente in die Experimente der vorliegenden Arbeit miteinbezogen. Aus den Studien geht hervor, dass sich die Probanden, die an eine externe Speicherung der Informationen glauben, trotz expliziter Aufforderung nicht an die recherchierten Informationen erinnern können. Deshalb wurden die Probanden der vorliegenden Studie in jeder Bedingungsgruppe dazu aufgefordert, sich die Informationen bzw. die richtigen Antworten zu merken.

Es ergibt sich somit die folgende Verteilung:

Tabelle 1: Gruppenverteilungen

	Keinen Zugriff auf ChatGPT / Google	**Zugriff auf ChatGPT / Google**
Google	Gruppe 1 (K1)	Gruppe 2 (K2)
(Kontrollgruppe)	Experiment N =12	Experiment N =12
ChatGPT	Gruppe 3 (T1)	Gruppe 4 (T2)
(Treatmentgruppe)	Experiment N =12	Experiment N =12

In dieser Studie sollte untersucht werden, ob die intensivere Erarbeitung von Antworten mit ChatGPT (durch die Unterhaltung mit dem Chatbot/durch das Prompten) im Vergleich zur Recherche mit Google einen stärkeren Einfluss auf die Erinnerungsleistung hat. Wie in der Tabelle zu erkennen ist, wurde der ChatGPT-Gruppe 3 und der Google-Gruppe 1 mitgeteilt, dass sie später ChatGPT oder Google nicht mehr benutzen dürften und es deshalb gut wäre, sich die richtigen Antworten zu merken.
Die Bedingung einer fehlenden expliziten Aufforderung, sich die Inhalte zu merken, wurde im Experiment nicht untersucht, da zu vermuten ist, dass die Inhalte nicht erinnert werden, wenn sie nicht erinnert werden müssen.

Es wurden folgende Haupt-Arbeitshypothesen aufgestellt:
H0: Die Erarbeitung von Inhalten mit ChatGPT hat keinen Einfluss auf die Erinnerungsleistung der erarbeiteten Inhalte nach zwei bis drei Wochen.
H1: Die Erinnerungsleistung in Bezug auf die mit ChatGPT erarbeiteten Inhalte ist nach zwei bis drei Wochen signifikant schlechter gegenüber den Inhalten, die ohne ChatGPT erarbeitet wurden.

Da wie oben erwähnt die Erinnerungsleistung vor allem dadurch beeinflusst wird, ob die Probanden glauben, dass die Informationen gespeichert werden und sie somit weiterhin «Zugriff» auf den Speicherort (in diesem Kontext: Google oder ChatGPT) haben, wurden noch weitere Hypothesen zu dieser Komponente formuliert.

H1.1 ➔ *Wenn Antworten mit ChatGPT erarbeitet werden, die Probanden überzeugt sind, dass ChatGPT* ***nicht zugänglich*** *bleibt, und man sie dazu auffordert, sich die Antworten zu merken, dann wird besser erinnert als in ChatGPT-Gruppe 4 und Google-Gruppe 2.*

H1.2 ➔ *Wenn Antworten mit ChatGPT erarbeitet werden, die Probanden überzeugt sind, dass ChatGPT* ***weiterhin zugänglich*** *bleibt, und dazu aufgefordert wird, sich an die Antworten zu erinnern, dann wird weniger gut erinnert als in der ChatGPT-Gruppe 3.*

H2.1 ➔ *Wenn die Probanden überzeugt sind, dass Google* ***nicht zugänglich*** *bleibt, und sie dazu aufgefordert werden, sich die richtigen Antworten zu merken, dann wird über alle Gruppen hinweg besser erinnert.*

H2.2 ➔ *Wenn die Probanden überzeugt sind, dass Google* ***weiterhin zugänglich*** *bleibt, und sie dazu aufgefordert werden, sich die richtigen Antworten zu merken, dann wird weniger gut als in Google-Gruppe 1 und besser als in den ChatGPT-Gruppen erinnert.*

Diese Hypothesen werden im Kapitel zu den Ergebnissen wieder aufgegriffen.

3.1.1 Variablen

In dieser Studie wurden als abhängige Variable (AV) die Erinnerungsleistung und als unabhängige Variable (UV) die Nutzung von ChatGPT sowie auch die Komponente des Zugriffs auf ChatGPT oder Google definiert.

Zusätzlich zu jenen aus der Theorie hat die Autorin folgende weitere Gründe ausgemacht, die die Erinnerungsleistung beeinflussen können:

- **Informationsverfügbarkeit:** Mitarbeitende merken sich die Informationen nicht, weil sie jederzeit auf die Informationen über ChatGPT oder das Internet zugreifen können.
- **Informationslokalisierung/Speicherort:** Mitarbeitende merken sich die Informationen nicht, da sie wissen, wo sie die Informationen finden können.
- **Unbewusste Erwartung:** Mitarbeitende merken sich die Informationen nicht, weil sie nicht wissen, dass sie sich diese merken sollen.
- **Abneigung:** Mitarbeitende haben eine Abneigung gegenüber ChatGPT und dieses Vorbehalten beeinflusst die Erinnerungsleistung.
- **Vorkenntnisse/Wissen:** Mögliche Vorkenntnisse zum Thema, das abgefragt wird, können die Erinnerungsleistung positiv beeinflussen.
- Die **Atmosphäre** kann die Erinnerungsleistung positiv oder negativ beeinflussen.
- **Eigenständige oder intensive Erarbeitung:** Wenn Informationen selbständig oder intensiver erarbeitet werden, kann sich das auf die Erinnerungsleistung auswirken.
- **Tageszeit**: Gewisse Personen sind zu bestimmten Zeiten am Tag aufnahmefähiger als andere. Um dies zu berücksichtigen, wird den Teilnehmenden die Möglichkeit

gegeben, die Zeit des Versuchs flexibel auszuwählen. Ebenso wurden die Teilnehmenden darauf aufmerksam gemacht, erholt zu den Terminen zu erscheinen. Bestenfalls sollten sie sogar eine kurze Pause machen.

Diese Einflüsse sind gemeinsam mit den in der Theorie genannten Einflüssen entsprechend in das Risikomanagement der Störvariablen eingeflossen.

3.1.1.1 Kontrolle der Störvariablen

Der Einfluss von Störvariablen auf die Ergebnisse des Experiments oder auf die Erinnerungsleistung wurde wie folgt minimiert.

Tabelle 2: Störvariablen (eigene Darstellung, 2024)

Risiko/ Störungsvariable	Risiko minimieren durch	Beispiel
Schlechte Miro-Board-Kenntnisse	Vor-Ort-Durchführung	Persönlicher Eingriff der Versuchsleitenden mit Computermaus als Unterstützung
Technische Störungen	Je nach Zeitpunkt im Experiment wurde entweder ein neuer Termin festgelegt oder ein Neustart durchgeführt	Bei Online-Durchführungen, bei denen es starke Bildschirms-Verzögerungen oder schlechte Audioqualität gab, wurde gleich zu Beginn des Experiments abgebrochen und ein neuer Termin festgelegt
ChatGPT-Störung	Neustart von ChatGPT, Verbindung zum VPN prüfen und trennen	ChatGPT funktioniert mit VPN von Swisscom nicht und wird unterdrückt
Internetausfall	Handy-Hotspot immer bereit	Bei schlechter Internetqualität wurde der Handy-Hotspot verwendet
Nervosität des Probanden	Kurze Pause, Einführung ins Thema, Snack anbieten	
Lärm und ähnliche Störungen während des Experiments	Bestimmte Raumauswahl, um Störung zu vermeiden	Kennzeichnung an den Räumen, dass eine Studie durchgeführt wird

Aufnahmefähigkeit je nach Tageszeit	Terminauswahl durch Probanden; a. o. Termine bei Bedarf	Es wurden Termine zu passenderen Zeiten angeboten
Versuchsleitenden-Effekt	Wenige Interaktionen während des Experiments nach den Instruktionen	Nach der Einleitung und Erklärung musste während des Experiments wenig bis gar nicht kommuniziert werden

Durch diese Massnahmen und Überlegungen sollte sichergestellt werden, dass die Ergebnisse der Studie valide sind und die Einflussfaktoren auf die Erinnerungsleistung umfassend berücksichtigt werden. ChatGPT-Störungen (2) wegen des eingeloggten Swisscom-VPN und schlechte Miro-Board-Kennnisse (3) kamen in den Experimenten vor.

3.1.2 Material

Für den Versuch wurde ein Miro-Board erstellt, das die Probanden Schritt für Schritt durch die verschiedenen Etappen im Experiment führte. Dies ermöglichte eine flexible Durchführung remote oder vor Ort in den Räumlichkeiten von Swisscom. Die Teilnehmenden benötigten hierfür nur einen Laptop und einen ChatGPT-3.5-Account. Je nach technischer Affinität und um technischen Störungen vorzubeugen, wurden einige Teilnehmenden vor Ort eingeladen. Wenn Teilnehmende keinen ChatGPT-Account hatten oder installieren wollten, wurden diejenigen auch vor Ort eingeladen, damit ihnen ein solcher Account zur Verfügung gestellt werden konnte.
Alle anderen Probanden aus den ChatGPT-Gruppen wurden vorab gebeten, sich einen Account für die Version 3.5 zu erstellen. ChatGPT 3.5 wurde verwendet, da dieser für alle Personen frei zugänglich ist.
Für die Kontrollgruppen wurde die Google-Suche in einem herkömmlichen Webbrowser wie Chrome oder Edge im Inkognito-Modus verwendet. Der Inkognito-Modus fand deshalb Anwendung, damit die Google-Suchen nach dem Experiment nicht mehr angezeigt wurden. Um sich gegen Datenverlust abzusichern, wurden alle Versuche auf MS Teams aufgezeichnet. Da im Rahmen der Experimente, abgesehen von den Instruktionen, kaum gesprochen wurde, wurden die Gesprächsverläufe nicht transkribiert.

Während des Experiments wurden den Probanden über das Miro-Board einfache und schwierige Trivia-Fragen gestellt, wie dies auch in den Studien von Nagam (2023), Storm et al. (2017) und Sparrow et al. (2011) erfolgt war. Es wurde bei der Wahl der Fragen darauf geachtet, dass sie in den Schweizer Kulturkreis passen. Die Fragen wurden nacheinander

aufgedeckt. Die Probanden hatten Zeit, die Fragen zu recherchieren und ihre Antworten einzutragen, bis die nächste Frage auf dem Miro-Board aufgedeckt wurde.
Zweieinhalb Wochen nach dem Experiment wurde der Fragebogen an die Teilnehmenden gesendet. Der Fragebogen wurde auf Tivian erstellt und an die Probanden per Mail verschickt. Um sicherzustellen, dass alle den Fragebogen 2.5 Wochen später erhalten, wurde der E-Mailversand vorprogrammiert und am entsprechenden Tag ausgelöst. Der Fragebogen diente als Messwiederholung der Erinnerungsleistung im Vergleich zu den Antworten aus dem Versuch. Im Fragebogen wurden dementsprechend dieselben zwei Aufgaben aus dem Versuch gestellt. Zusätzlich zum Erinnerungstest wurden die Probanden zu Einschätzungen zu ChatGPT in Zusammenhang mit Erinnerung und Informationsverarbeitung befragt. Der Fragebogen wird später im Kapitel vorgestellt.

3.1.3 Versuchsplan

Die Versuche wurden 1:1, also einzeln mit dem jeweiligen Probanden, durchgeführt. Da der gesamte Versuch auf einem Miro-Board aufgebaut wurde, erhielten die Probanden aus allen vier Gruppen zu dem jeweiligen Board Zugangsberechtigungen. Auf dem Board befanden sich zu Beginn die Instruktionen (Anhang A) zu dem Versuch, die Aufgabe mit den verdeckten Trivia-Fragen («Aufgabe» siehe Abbildung) sowie die anschliessenden Lösungen dieser Trivia-Fragen sowie der Erinnerungstest bestehend aus zwei Aufgabenblocks.

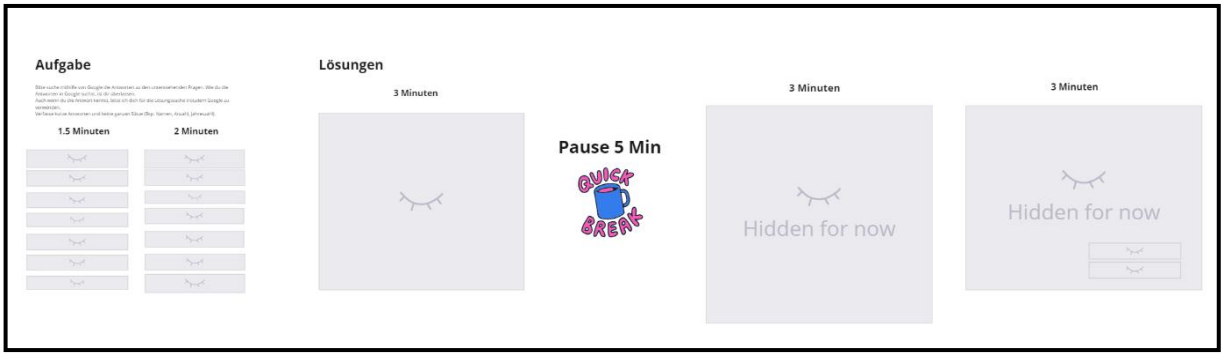


Abbildung 3: Miro-Board-Experimentaufbau zur Illustration des Experiments (eigene Darstellung, 2024)

Nach dem der jeweilige Proband die Instruktionen genau durchgelesen hatte und allfällige Fragen zum Ablauf geklärt wurden, wurde eine Einverständniserklärung (Beispiel hierfür im Anhang B) unterschrieben und die Probandennummer wurde erstellt. Bei remote durchgeführten Versuchen wurde die Einverständniserklärung im Nachgang per E-Mail vom Probanden retourniert. Das gesamte Experiment dauerte in der Regel 40 bis 45 Minuten.

Das Experiment setzte sich aus den unten abgebildeten Bestandteilen zusammen.

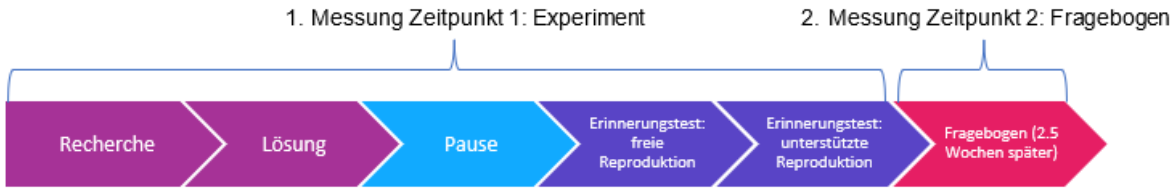


Abbildung 4: Schritt-für-Schritt-Versuchsaufbau (eigene Darstellung, 2024)

Recherche: Die Teilnehmenden durften frei auf ChatGPT oder Google die Antworten zu den Trivia-Fragen recherchieren. Für die einfachen Trivia-Fragen wurde den Probanden ein Zeitraum von 1.5 Minuten und für schwierige Fragen ein Zeitraum von zwei Minuten zur Verfügung gestellt. Die Fragen wurden nacheinander aufgedeckt. Die Probanden hatten Zeit, die Fragen zu recherchieren und ihre Antworten einzutragen, bis die nächste Frage auf dem Miro-Board aufgedeckt wurde. Sie durften die gegebene Zeit für die Recherche benutzen, bis sie das Gefühl hatten, die korrekte Antwort gefunden zu haben. Die Art und Weise, wie sie in ChatGPT oder Google suchen wollten, wurde den Probanden überlassen. Die Probanden sollten Antworten suchen, selbst wenn sie die Antwort bereits kannten. Die Antworten der Probanden sollten kurz sein (z. B. nur die Jahreszahlen, Namen oder Anzahl des Gegenstands umfassen). Unter «Aufgabe» auf dem Miro-Board (siehe Abbildung 3) gab es eine kurze Instruktion, wo das soeben Beschriebene kurz zusammengefasst wurde. Die Kontrollgruppen durften bei ihren Recherchen nur die erste Seite der Google-Resultate verwenden und nicht auf andere Seiten (wie beispielsweise Wikipedia etc.) ausweichen. Wenn keine passende Antwort zu finden war, musste die Suchanfrage angepasst werden. Bei den ChatGPT-Gruppen durfte in ChatGPT so lange anhand von unterschiedlichen Prompts recherchiert werden, bis die Probanden eine zufriedenstellende Antwort in der gegebenen Zeit finden konnten. Im Vergleich zu den ChatGPT-Gruppen sollten die Recherchen der Kontrollgruppe mithilfe der Google-Suche weniger intensiv ausfallen, da Google die Antwort direkt vorschlägt.

Lösung: Alle Gruppen erhielten anschliessend die Lösungen für drei Minuten eingeblendet und wurden gebeten, sich die richtigen Antworten zu merken.

Pause: Danach erfolgte eine kurze Fünf-Minuten-Pause, in der die Teilnehmenden durch Small-Talk von der Aufgabe abgelenkt wurden.

Erinnerungstest: Die zweite Phase des Experiments, der Erinnerungstest, dauerte ca. 6 Minuten.
Im ersten Teil sollten die Probanden die gemerkten Antworten der Trivia-Fragen wiedergeben und im Miro niederschrieben. Dabei sollte es sich um eine **freie Reproduktion** der Antworten ohne jegliche Gedächtnisstütze handeln.
Danach folgte der zweite Teil, bei dem die Probanden die vorherigen Trivia-Fragen als Gedächtnisstütze eingeblendet hatten. Die mögliche Wiedererkennung der Fragen sollte die Probanden bei der Erinnerung an die korrekten Antworten unterstützen (**unterstützte Reproduktion**).
Anschliessend durften die Teilnehmenden als kurzes Debriefing die korrekten Lösungen nochmals sehen. Ebenso wurde daran erinnert, dass sie in ca. drei Wochen einen Fragebogen erhalten würden.

Fragebogen: Der Fragebogen diente der Messwiederholung der Erinnerungsleistung bezüglich der Antworten aus dem Versuch. Im Fragebogen wurden dementsprechend dieselben zwei Aufgaben aus dem Versuch gestellt (Erinnerungstestaufgabe freie Reproduktion und Wiederkennung). Zusätzlich zum Erinnerungstest wurden die Probanden zu Einschätzungen zu ChatGPT in Zusammenhang mit Erinnerung und Informationsverarbeitung befragt. Auch die Kontrollgruppe wurde gebeten, die Fragen zu ChatGPT zu beantworten.
Ein richtiges Debriefing wurde den Teilnehmenden nach dem Fragebogen auf Wunsch angeboten.

3.1.4 Trivia-Fragen

Für das Experiment wurden bewusst Trivia-Fragen verwendet. Bei der Wahl wurden einige wichtige Faktoren bedacht.
Einerseits wurden in allen drei vorliegenden Studien zu Erinnerungsleistung und Internetnutzung Trivia-Fragen verwendet und die vorliegende Arbeit sollte diese Studien auf eine gewisse Weise replizieren. Andererseits wurden die Themen Vorwissen, Lösbarkeit sowie kulturelle Passung eingehend diskutiert.
Trivia-Fragen mit einer Google-/ChatGPT-Recherche zu lösen, benötigt kein Vorwissen. Alle Mitarbeiten sollten bei der Lösung der Fragen dieselben Chancen haben.
Die Trivia-Fragen sollten für alle Mitarbeitenden lösbar sein. Es wurden einfache sowie schwierige Fragen ausgesucht, damit die Aufgabe als relevant genug wahrgenommen werden würde.
Die meisten Fragen, die ausgesucht wurden, beziehen sich mit vereinzelten Ausnahmen auf den Raum Europa.

Da die Personen für das Experiment an den Swisscom-Standorten Bern und Zürich rekrutiert wurden, durften die Fragen oder die Aufgabe nicht auf einen spezifischen Bereich abzielen, da sonst einzelne Probanden bevorteilt würden. Des Weiteren ist der Datenschutz bei Swisscom ein relevanter Faktor, weshalb die Fragen und Aufgabenstellungen keine firmeninternen Informationen beinhalten sollten, die möglicherweise in ChatGPT eingegeben werden könnten.

Hier je ein Beispiel für eine einfache und eine schwierige Trivia-Frage.

Einfach: *«Wie viele Spieler gibt es in einer vollständigen Baseballmannschaft auf dem Spielfeld?»*

Schwierig: *«Welches chemische Element wurde nach einem berühmten Physiker benannt, der das Periodensystem erweitert hat?»*

Bei einer durchmischten Stichprobe von Lernenden bis hin zu Senior-Mitarbeitenden aus unterschiedlichen Bereichen mit diversen Anforderungen ist eine objektive Einschätzung des Schwierigkeitslevels und der Machbarkeit der Trivia-Fragen nicht eindeutig durchführbar.

3.2 Pretest

Insgesamt wurden fünf Pretests durchgeführt. Hierfür wurden jeweils Personen aus dem Umfeld zu einem kurzen Versuch eingeladen. Das Experiment wurde wie oben beschrieben durchgeführt und die Anmerkungen und Verbesserungsvorschläge der Personen wurden in die Überarbeitung des experimentellen Designs eingebracht. Zwei Personen testeten zu einem späteren Zeitpunkt auch noch den Fragebogen. Die Pretests verliefen gut und ohne jegliche Schwierigkeiten.

3.3 Versuchsteilnehmende

Die Studie wurde bei der Swisscom Schweiz AG, der Praxispartnerin dieser Masterarbeit, durchgeführt. Deshalb wurden für die Studie Swisscom-Mitarbeitende für das Experiment rekrutiert.

Für die Teilnahme an der Studie wurden keine finanziellen Anreize geboten. Lediglich für Snacks und Getränke wurde gesorgt. Das Interesse der Swisscom-Mitarbeitenden war trotzdem sehr hoch.

Da Vorkenntnisse im Thema ChatGPT sowie auch die Erinnerungsleistung der Probanden im Allgemeinen eine Rolle im Experiment spielen, wurde ein Fragebogen via Microsoft Forms erstellt. Der Fragebogen funktionierte als Screener sowie Rekrutierungstool. Es wurden darin die Vorkenntnisse zu ChatGPT erfragt (mehrmals pro Woche bis noch nie genutzt) sowie auch Einschätzungsfragen zum eigenen Kurzzeit- und Langzeitgedächtnis.

Mit diesem Screener sollten Extremwerte in den Daten vermieden werden. Ebenso wurden Alter und Geschlecht abgefragt.
Da die Experimente auf Hochdeutsch durchgeführt wurden, wurden nur Mitarbeitende aus der Deutschschweiz mit guten Deutschkenntnissen rekrutiert.
Bei der Rekrutierung wurde darauf geachtet, eine ausgeglichene Stichprobe zu bilden, sodass die einzelnen Gruppenbedingungen einander in ihren Ausprägungen ähnelten (ausgeglichene Vorkenntnisse von ChatGPT, Leute mit gutem oder weniger gutem Kurzzeit- oder Langzeitgedächtnis und gute Altersverteilung).
Ausgeschlossen von der Studie wurden Personen mit medizinischen Beeinträchtigungen bei der Gedächtnis- oder der Erinnerungsleistung.
Für die statistische Power der Studie und eine Signifikanz von $\alpha = .05$ wurde vor Beginn der Datenerhebung ein Umfang von ca. 45 Probanden berechnet. Für das Experiment konnten effektiv 48 Teilnehmende rekrutiert werden.

3.4 Datenerhebung

Im Folgenden wird zunächst das Vorgehen der Datenerhebung im Rahmen des Experiments und nachfolgend dasjenige des Fragebogens genauer vorgestellt.

3.4.1 Erinnerungstests

Im ganzen Studiendesign gibt es zwei Erhebungen zur Erinnerungsleistung: den Versuch als erste Messung sowie den Fragebogen als zweite Messung.
Beim Versuch wurde die Anzahl der korrekt gemerkten Antworten gewertet. Dabei musste festgelegt werden, ab wann die Antwort einer Frage als korrekt, teilweise richtig oder falsch galt.
Für die Bewertung der Antworten wurde daher das nachfolgende Bewertungsraster festgelegt. Die Bewertungen wurden von einer zweiten Person kontrolliert.

Tabelle 3: Bewertungsraster Erinnerungstests (eigene Darstellung, 2024)

Bewertungsraster	**Beispiel**
Richtige Antwort richtig geschrieben volle Punktezahl	-
Vorname und Nachname jeweils 1 Pkt	Albert Einstein 2 Pkte, Einstein 1 Pkt,
Verdeutschungen von Namen wurden akzeptiert	Edward statt Edvard, Tensing statt Tenzing
Tiernamen vollständig richtig (1 Pkt)	Zwergspitzmaus, Felsenklapperschlange
Tiernamen teilweise richtig (0.5 Pkte)	Spitzmaus, Klapperschlange
Name von Ortschaft (1 Pkt)	Aconcagua 1 Pkt
Namen von Ortschaften teilweise richtig (0.5 Pkte)	Acongua

Alle Experimente wurden zwecks Beobachtungen aufgezeichnet. Da die Beobachtungen für die Beantwortung der Forschungsfrage nicht relevant sind, wurden die Experimentsaufzeichnungen nicht transkribiert. Die Beobachtungen aus den Experimenten werden zu Interpretationszwecken in die Diskussion der Ergebnisse einfliessen.

3.4.2 Fragebogen

Der Fragebogen, den die Probanden ca. 2.5 Wochen nach dem Versuch erhielten, enthält einerseits eine Messwiederholung des Versuchs (Erinnerungstest) sowie Einschätzungsfragen der Probanden zu ChatGPT.
Die Einschätzungsfragen dienen als Hintergrundinformationen für die Ergebnisse aus dem Erinnerungstest. Die Ergebnisse aus dem Fragebogen sollten Kontextinformationen zur Interpretation in der Diskussion liefern.

Zur Messwiederholung wurden die Teilnehmenden nochmals gebeten alle Antworten wiederzugeben, an die sie sich erinnern konnten (freie Reproduktion der Antworten). Danach folgte wieder eine Aufgabe, bei der die Teilnehmenden mithilfe der Fragen die Antwort aus der Erinnerung (unterstützte Reproduktion) wiedergeben sollten.

Tabelle 4: Fragebogenaufbau (eigene Darstellung, 2024)

Fragen aus dem Fragebogen	**Antworten**
Wie hat sich deine Einstellung gegenüber ChatGPT im letzten Jahr verändert?	Offenes Antwortfeld
Inwieweit stimmst du den folgenden Aussagen zu? **6er Skala (1 stimme überhaupt nicht zu, 6 stimme voll und ganz zu)**	ChatGPT beeinflusst das Wissen von Menschen. ChatGPT beeinflusst die Erinnerungsleistung / Gedächtnis der Menschen. ChatGPT beeinflusst die Art wie Informationen erinnert werden.
Welche Methode hast du gewählt, um dir die Antworten der Trivia Fragen einzuprägen? (Plechatá et al., 2019) **(Auswahlmöglichkeit)**	- Auswendiglernen (Wiederholung im Kopf) - Sortieren der Gegenstände in Kategorien (Obst, Gemüse, Kosmetika ...) - Visualisierung der Gegenstände im Kopf - Assoziationen zwischen Gegenständen (z. B. Löffel gehört auf den Teller) - Visualisierung des Weges oder der Verteilung der Artikel im Supermarkt - Erfinden einer Geschichte aus den Gegenständen auf der Liste - andere Antwort
Freie Reproduktion: Bitte gib alle Antworten aus dem Experiment an, an die du dich noch erinnern kannst. Verwende hierfür bitte keine Hilfsmittel.	Erinnerungstest Aufgabe 1
Unterstützte Reproduktion durch die Trivia-Fragen: Die Fragen sollten dir helfen dich zu erinnern. Bitte gib alle Antworten aus dem Experiment an, an die du dich noch erinnern kannst. Verwende hierfür bitte keine Hilfsmittel.	Erinnerungstest Aufgabe 2
Inwieweit trifft die nachfolgende Aussage zu? **5er Skala (1 definitiv nein, 5 definitiv ja)** (Plechatá et al., 2019)	Ich konnte alle korrekten Antworten wiedererkennen.
Wie hat sich deine Erinnerungsleistung seit der Nutzung von ChatGPT verändert?	Offenes Antwortfeld
Nutzung von ChatGPT	- Ich nutze ChatGPT oft auch für kleine Aufgaben. Es nimmt mir Arbeit ab.

	- Ich nutze ChatGPT absichtlich nicht oft, um mich nicht zu stark daran zu gewöhnen und darauf angewiesen zu sein. - Ich nutze ChatGPT für Ideengenerierung. - Ich nutze ChatGPT, um meine bereits entwickelten Konzepte weiterzuentwickeln. - Ich nutze ChatGPT öfter als noch vor 6 Monaten. - Ich finde ChatGPTs Antworten vertrauenswürdig. - Ich hinterfrage ChatGPTs Antworten mit meinem eigenen Fachwissen. - Ich prüfe ChatGPTs Antworten mit anderen Quellen.
Inwieweit treffen die folgenden Aussagen auf dich zu? **5er Skala (1 trifft überhaupt nicht zu, 5 trifft voll und ganz zu)**	- Wenn ich Informationen jederzeit online abrufen kann, fällt es mir schwer, mich an diese zu erinnern. - Wenn ich mich intensiv mit einem Thema beschäftige, dann kann ich mich gut daran erinnern, obgleich Informationen abrufbar sind. - Wenn mich ein Thema weniger interessiert, dann habe ich Mühe, mich daran zu erinnern. - Indem ich mir gewisse Informationen nicht merke, schaffe ich mir mehr kognitive Ressourcen für andere, relevantere Informationen. - Wenn ich weiss, wo ich Informationen finde, erinnere ich mich weniger gut an die Information selbst. - Ich merke mir Informationen eher, wenn ich das Gefühl habe, diese nicht wiederfinden zu können. - Ich erinnere mich eher an Informationen, die ich interessant finde. - Wenn ein Thema unverständlich ist, dann habe ich Mühe, mich daran zu erinnern. - Ich merke mir Informationen, obgleich ich sie online abrufen kann.

Zwei Fragen und die dazugehörigen Items für den Fragenbogen wurden aus der Literatur abgeleitet. Die restlichen Fragen und Items mit entsprechenden Likert-Skalen wurden von der Autorin passend zum Forschungsthema erstellt.

Datenanalyse

Für die Auswertung wurden die Experimente aufgezeichnet. Da das Gesagte während des Experiments nicht ausgewertet wurde, sondern das Verhalten in der Recherchearbeit beobachtet wurde, wurde auf eine Transkription verzichtet. Die Daten aus den Erinnerungstests und Fragebögen wurden in SPSS-Version 29.0.2.0 eingegeben und aufbereitet.

Da Gruppenunterschiede bezüglich der Erinnerungsleistung bei einer unabhängigen Stichprobe festgestellt werden sollten, wurde eine ANOVA mit Messwiederholung durchgeführt, um die Erinnerungsleistung zwischen den Gruppen zu vergleichen, die entweder Google oder ChatGPT verwendet hatten. Die Entscheidung für eine ANOVA mit Messwiederholung basiert auf der Notwendigkeit, Unterschiede zwischen mehreren Messzeitpunkten innerhalb der gleichen Probanden zu untersuchen. Auch Post-hoc-Tests wie Bonferroni und Tukey-HSD wurden geprüft.

Die abhängige Variable war die Leistung in den Erinnerungstests, gemessen als Punktezahl, und die unabhängige Variable war die Art des Zugriffs auf Google oder ChatGPT (nominal kodiert). Die Punktezahl der Erinnerungstests wurde basierend auf der Anzahl der korrekt erinnerten Informationen berechnet.

Zur Kontrolle möglicher Störvariablen wie Alter, Geschlecht, Vorerfahrung mit AI-Technologien und Erinnerungsvermögen wurden diese mittels erwähntem Screener ausgewertet und gleichmässig auf die vier Gruppen verteilt, um deren Einfluss zu minimieren.

Für die restlichen quantitativen Antworten aus dem Fragebogen, insbesondere die Likert-Skalen, wurden Häufigkeits- und Mittelwertberechnungen durchgeführt. Zusätzlich wurden Korrelationsanalysen durchgeführt, um mögliche Zusammenhänge zwischen den Likert-Skalen und der Erinnerungsleistung zu identifizieren. Deshalb wurden im Nachhinein zwei Variablen aus dem Fragebogen in einer weiteren ANOVA als Kovariaten geprüft.

Alle Analysen basieren auf einem mittleren Signifikanzniveau von $p < 0{,}05$.

3.4.3 Ethik

Die Teilnehmenden wurden vor der Durchführung über die Aufgabe aufgeklärt, ohne vorab zu viele Informationen freizugeben. Es wurde also nicht weitergegeben, dass die Erinnerungsleistung gemessen wird. Den Teilnehmenden wurde lediglich mitgeteilt, dass durch die Versuche der Einfluss auf eine kognitive Leistung gemessen wird. Diese Formulierung liess Raum zur Interpretation, ohne etwas vorzutäuschen. In der Regel war den Teilnehmenden am Ende des Versuches klar, dass es sich um die Erinnerungsleistung oder das Gedächtnis handeln musste.

Die Probanden erhielten eine Einverständniserklärung, bei der sie über die Studie informiert wurden. Den Teilnehmenden wurde mitgeteilt, dass sie das Experiment jederzeit ohne

Angabe von Gründen abbrechen könnten. Ebenso wurde den Probanden versichert, dass ihre Daten mithilfe der Probandennummer anonymisiert wurden und die Daten nach Abschluss der Studie vollständig gelöscht werden.

Nach Abschluss des Experiments wurde den Teilnehmenden angeboten, nach Abgabe des Fragebogens ein umfassendes Debriefing durchzuführen. Den Probanden wurde dann erklärt, dass die Erinnerungsleistung tatsächlich der Hauptfokus der Studie war. Die ethischen Kriterien wurden vor Beginn der Studie mit der Ethikcheckliste der FHNW geprüft. Das Experiment ist gemäss dieser akzeptiert (Anhang C). Es wurde so gestaltet, dass für die Teilnehmenden keine Risiken entstanden.
Die Teilnahme an der Studie war für alle Probanden freiwillig.

3.4.4 Interne Validität

Damit die vorliegende Studie auch wirklich die Erinnerungsleistung misst, wurde Literatur zur Messung der Erinnerungsleistung recherchiert. Bezugnehmend darauf wurde mit dem gewählten Studiendesign das Erinnerungsvermögen anhand der freien und unterstützten Reproduktion gemessen. Der Einfluss von ChatGPT- oder Google-Recherche sollte ausschlaggebende Unterschiede aufzeigen.
Die Experimente wurden unter kontrollierten Bedingungen durchgeführt, um externe Einflüsse zu minimieren. Die Umgebung war für alle Teilnehmenden an den Standorten Bern und Zürich gleich, und Ablenkungen wurden weitgehend ausgeschlossen.

Allen Teilnehmenden wurden identische Instruktionen gegeben, um sicherzustellen, dass Unterschiede in der Leistung nicht auf unterschiedliche Instruktionen zurückzuführen sind. Um einen Bewertungsbias zu verhindern, konnte die Experimentleitung die Teilnehmenden aufgrund der anonymisierten Probandennummern den Antworten der Erinnerungstests nicht zuordnen.

Um die Objektivität der Auswertung zu gewährleisten, wurden die Ergebnisse der Erinnerungstests von der Experimentleitung sowie zwei weiteren unabhängigen Kodierern ausgewertet. Unterschiede in den Bewertungen wurden diskutiert und konsensual entschieden (siehe abschliessendes Bewertungsraster im Unterkapitel Erinnerungstests).

Durch diese Massnahmen wird die interne Validität der Studie gestärkt, sodass die Ergebnisse zuverlässig die Auswirkungen der Nutzung von ChatGPT und Google auf die Erinnerungsleistung der Teilnehmenden widerspiegeln.

3.4.5 Externe Validität

Die Teilnehmenden wurden aus verschiedenen Abteilungen und mit unterschiedlichen beruflichen Hintergründen rekrutiert, um eine möglichst breite Repräsentation innerhalb der Zielpopulation zu gewährleisten. Da diese Studie jedoch ausschliesslich mit deutschsprachigen Swisscom-Mitarbeitenden durchgeführt wurde, sind die Resultate nicht auf die gesamte Swisscom oder die Bevölkerung verallgemeinerbar.
Da die Aufgabenstellung keine typischen beruflichen Aktivitäten umfasste, ist anzunehmen, dass die externe Validität dieser Studie nicht hoch ist.
Eine grössere Vielfalt bezüglich der demografischen Zusammensetzung der Stichprobe, einschliesslich Altersgruppen, Geschlechtern, Bildungshintergründen und Berufserfahrungen, könnte dazu beitragen, die Generalisierbarkeit der Ergebnisse zu erhöhen.
Zusammenfassend lässt sich sagen, dass die Ergebnisse der Studie zwar nicht ohne Weiteres auf die gesamte Bevölkerung oder auf alle Unternehmen übertragbar sind, die erhobenen Daten jedoch wertvolle Einblicke bieten, die auf andere Aufgaben im Geschäftskontext anwendbar sein könnten. Weitere Forschung ist notwendig, um die externe Validität zu erhöhen und eine Generalisierbarkeit der Ergebnisse zu gewährleisten.

4 Ergebnisse

Im vorliegenden Kapitel werden die Ergebnisse der Studie präsentiert. Zunächst wird die Stichprobe genauer beschrieben und danach werden die Resultate der Erinnerungstests aus Experiment und Fragebogen sowie die Ergebnisse aus den Fragebogenitems vorgestellt.

4.1 Beschreibung der Stichprobe

Für das Experiment konnten 48 Mitarbeitende der Swisscom der Standorte Bern und Zürich rekrutiert werden. Es ist wichtig zu erwähnen, dass nur deutschsprachige Mitarbeitende für die Experimente rekrutiert wurden, da die Durchführung der Experimente nur auf Deutsch stattfand und Sprachbarrieren keine Störungsvariable kreieren sollten. Wie bereits erwähnt, wurden den Teilnehmenden nebst dem Erinnerungstest auch Fragen zu ChatGPT-Vorkenntnissen und zu ihrem Erinnerungsvermögen gestellt, damit alle Teilnehmenden in ausgeglichene Gruppen verteilt werden konnten. Ein grosser Anteil der Probanden war zwischen 25-34 Jahre alt. Es waren Probanden von allen Altersgruppen bis auf die Gruppe 65+ vorhanden.

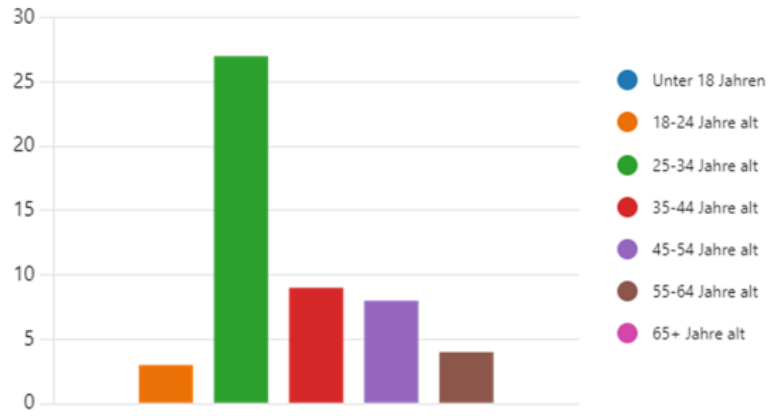


Abbildung 5: Altersverteilung der Stichprobe (Microsoft-Forms-Darstellung, 2024)

Ein grosser Anteil der Probanden hatte bislang ChatGPT mehrmals die Woche oder 1- bis 2-mal pro Woche genutzt. Ein bisschen mehr als ein Viertel der Probanden nutzte ChatGPT bisher weniger als 1-mal pro Monat, hatte es schon mal genutzt oder noch nicht genutzt (Anhang D: Kreisdiagramm zur Verteilung der Nutzungserfahrung).
Grundsätzlich nahmen etwas mehr Frauen als Männer an den Experimenten teil (siehe Anhang E für die genaue Geschlechtsverteilung).

Da trotz Erinnerungsmail nicht alle Teilnehmenden nach dem Experiment auch den Fragebogen retournierten, konnten für die nachfolgenden Ergebnisse nur 45 der 48 Ergebnisse aus dem Erinnerungstest des Experiments (1. Messzeitpunkt) mit dem Erinnerungstest aus dem Fragebogen (2. Messzeitpunkt) verglichen werden (N = 45). Drei Personen aus den Google-Gruppen hatten den Fragebogen nicht retourniert (Google-Gruppe 1: 11 Teilnehmende; Google-Gruppe 2: 10 Teilnehmende). Bei beiden ChatGPT-Gruppen füllten jeweils alle 12 Personen den Fragebogen aus (Anhang F Verteilung Teilnehmende in Gruppen).

4.2 Deskriptive Statistiken aus Erinnerungstest

Nachfolgend werden die Ergebnisse aus dem Erinnerungstest aus dem Experiment und dem Fragebogen vorgestellt und erläutert.
Wie in den Methoden beschrieben, wurde für den Erinnerungstest mit zwei Messzeitpunkten eine ANOVA mit Messwiederholung in SPSS gerechnet. Die Komponente Zugriff, also ob den Probanden mitgeteilt wurde, dass sie im späteren Verlauf des Experiments ChatGPT oder Google weiterhin nutzen dürften, wurde mit einer ANOVA mit Messwiederholung berechnet. Es wurde vorgängig geprüft, ob diese Variable einen Einfluss auf die Erinnerungsleistung der entsprechenden Gruppen hatte.
Laut Hypothese hätten die Gruppen mit keinem Zugriff eine höhere Punktezahl im Erinnerungstest aufweisen sollen, als es die Berechnungen angeben (siehe hierfür auch die Grafiken zum Messzeitpunkt 1 und 2 in *Abbildung 6*). Beide Zeitpunkte verhalten sich jedoch ähnlich.

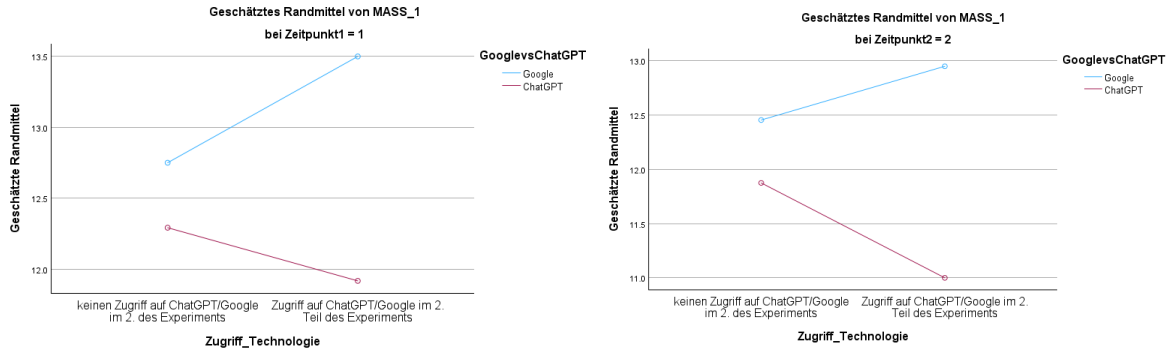


Abbildung 6: Strichdiagramme zur Variable Zugriff Messzeitpunkt 1 (links) Messzeitpunkt 2 (rechts) (SPSS, 2024)

Da diese Berechnung keine signifikanten Interaktionen oder Unterschiede anzeigte, wurde die ANOVA mit Messwiederholung später ohne diese Variable berechnet. Dies geschah einerseits, um die Interpretation der Daten zu erleichtern, und andererseits, weil es keine

signifikanten Unterschiede aus der vorherigen Berechnung anzeigte (alle Werte waren unter $p > 0{,}001$). Die genaue Tabelle mit den entsprechenden Werten befindet sich im Anhang G.

Reminder

Google-Gruppe 1 und **ChatGPT-Gruppe 3** wurde mitgeteilt, dass sie im späteren Verlauf **keinen Zugriff** auf Google oder ChatGPT haben werden.

Google-Gruppe 2 und **ChatGPT-Gruppe 4** wurde mitgeteilt, dass sie im späteren Verlauf weiterhin **Zugriff** auf Google oder ChatGPT haben werden.

Die Mittelwerte aus den Erinnerungstests freie Reproduktion (Aufgabe 1) sowie unterstützte Reproduktion (Aufgabe 2) zeigen im Allgemeinen ein ziemlich einheitliches Bild der Erinnerungsleistung. Die Mittelwertsunterschiede aus den Gruppen unterscheiden sich im Vergleich nur gering voneinander, wobei die Google-Gruppen einen leicht höheren Mittelwert aufweisen. In der nachfolgenden Tabelle werden die Gruppen mit und ohne Zugriff einander gegenübergestellt (siehe obigen Reminder zum Verständnis).

Tabelle 5: Ergebnisse Mittelwerte Erinnerungstests aus SPSS (eigene Darstellung, 2024)

	Gruppen	Mittelwert	Standardabweichung
Messzeitpunkt_1_freie_Reproduktion	Google-Gruppe 1	11.500	3.5071
	ChatGPT-Gruppe 3	11.333	3.1286
	Google-Gruppe 2	12.600	4.4272
	ChatGPT-Gruppe 4	11.042	2.7672
Messzeitpunkt_2_freie_Reproduktion	Google-Gruppe 1	7.000	4.1893
	ChatGPT-Gruppe 3	6.375	3.7847
	Google-Gruppe 2	5.900	3.3400
	ChatGPT-Gruppe 4	5.833	3.8337
Messzeitpunkt_1_unterstützte_Reproduktion	Google-Gruppe 1	14.000	3.3764
	ChatGPT-Gruppe 3	13.250	2.0394
	Google-Gruppe 2	14.400	2.8944
	ChatGPT-Gruppe 4	12.792	2.4445
Messzeitpunkt_2_unterstützte_Reproduktion	Google-Gruppe 1	10.909	3.6250
	ChatGPT-Gruppe 3	10.500	2.9310
	Google-Gruppe 2	11.500	3.3082
	ChatGPT-Gruppe 4	9.208	4.3976

Die Unterschiede der Mittelwerte sind klein und zeigen in der ANOVA mit Messwiederholung keine signifikanten Unterschiede der Gruppen in der Erinnerungsleistung an.
Die einzigen Signifikanzen zeigen sich beim Zeitpunkt ($F(1,41) = 103{,}026$; $p < 0{,}001$), in der Erinnerung ($F(1,41) = 186{,}455$; $p < 0{,}001$) sowie bei der Interaktion von Zeitpunkt und Erinnerung ($F(1,41) = 23{,}275$; $p < 0{,}001$). Weitere deskriptive Statistiken befinden sich im Anhang H.
Es folgen einige Grafiken zur Veranschauulichung dieser obigen Mittelwerte und der Signifikanz im Erinnerungszerfall der beiden Messzeitpunkte.

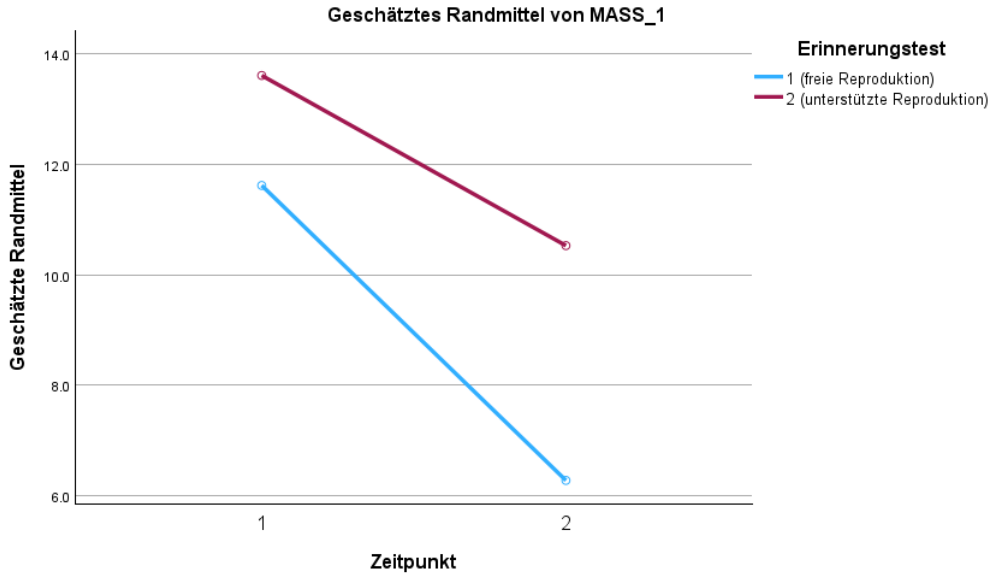


Abbildung 7: Strichdiagramm beide Erinnerungstest (SPSS, 2024)

In der Grafik (s. *Abbildung 7*) ist zu erkennen, dass zu beiden Messzeitpunkten die Aufgabe mit unterstützter Reproduktion (violett) besser gelöst wurde und dass im Vergleich zur Aufgabe mit freier Reproduktion (hellblau) mehr Punkte erzielt werden konnten.

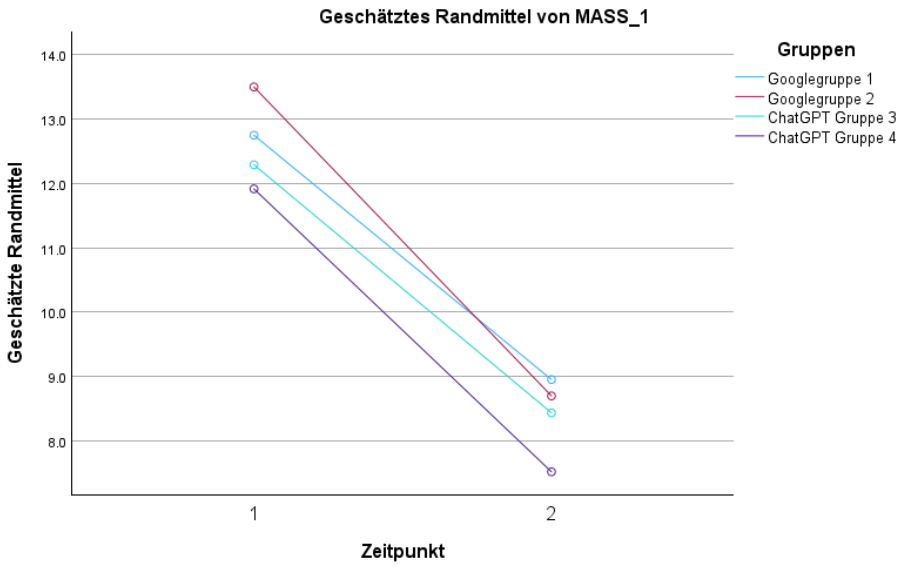


Abbildung 8: Strichdiagramm alle Gruppen Erinnerungsleistung über zwei Zeitpunkte (SPSS, 2024)

Über alle Gruppen hinweg ist die Erinnerungsleistung aus dem Erinnerungstest (freie und unterstützte Reproduktion) nach 2.5 Wochen signifikant schlechter. Es ist ein Abfall der Leistung sichtbar.

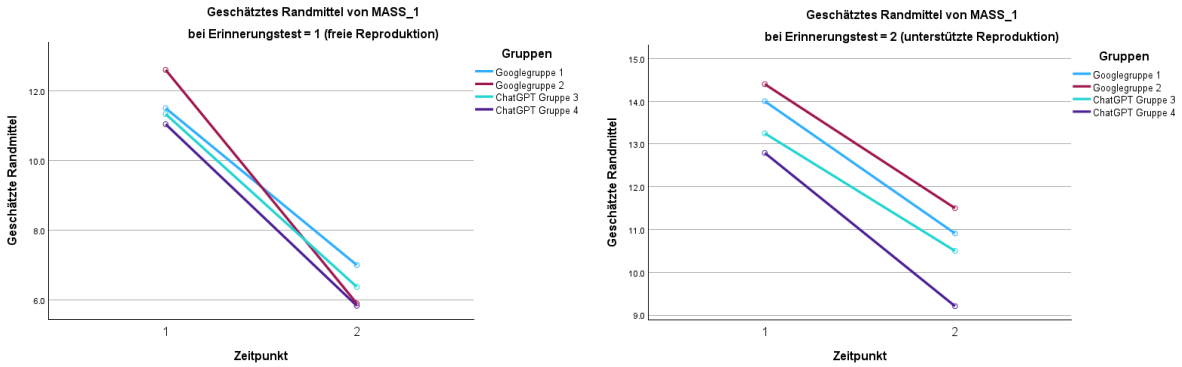


Abbildung 9: Detaillierte Strichdiagramme zu den einzelnen Erinnerungstests; Links: Erinnerungstests freie Reproduktion; Rechts: Erinnerungstest unterstützte Reproduktion (SPSS, 2024)

Dasselbe Bild zeigt sich in den Diagrammen beider Aufgaben zur freien Reproduktion (links) und unterstützten Reproduktion (rechts) separat ausgeführt. Die Gruppen verhalten sich gleich. Im Durchschnitt wurde die Aufgabe mit unterstützter Reproduktion besser gelöst.

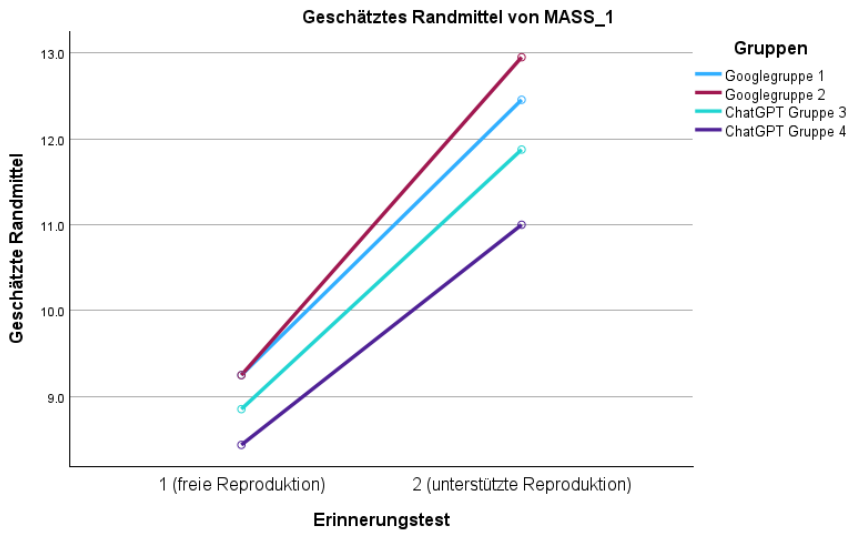


Abbildung 10: Strichdiagramm zu beiden Erinnerungsaufgaben (freie und unterstützte Reproduktion) Ergebnisse aller Gruppen (SPSS, 2024)

Wie in der Grafik (s. *Abbildung 10*) zu erkennen ist, ist der Unterschied in der Erinnerungsleistung bei den zwei Aufgaben, der freien vs. der unterstützten Reproduktion, sehr gross. Alle Gruppen weisen eine verbesserte Erinnerungsleistung auf, wenn sie mit den Fragen beim Erinnern unterstützt wurden (unterstützte Reproduktion). In einer weiteren ANOVA mit Messwiederholung konnte jeweils eine Signifikanz für den Unterschied der Erinnerungsleistung beider Aufgaben von $p > 0{,}001$ festgestellt werden ($F(1,41) = 23.275$; $p > 0{,}001$). Die detaillierte Tabelle befindet sich im Anhang I.

Um auszuschliessen, dass die einfachen Trivia-Fragen-Resultate den Erinnerungstest in der ANOVA verzerren könnten, wurden diese in einer weiteren ANOVA mit Messwiederholung ausgeschlossen.
Es wurde dementsprechend eine weitere ANOVA mit Messwiederholung nur für die Summe der Punkte aus den schwierigen Trivia-Fragen verwendet. Dabei zeigte sich kein verändertes Bild. Die Mittelwerte der Gruppen unterscheiden sich kaum und es wurden nur die Signifikanzen zu Zeitpunkt ($F(1,41) = 33{,}750$; $p < 0{,}001$) und Erinnerung ($F(1,41) = 55{,}089$; $p < 0{,}001$) ausgewiesen (Anhang J). Ausserdem war bei dieser Berechnung die Interaktion von Zeitpunkt und Erinnerung nicht mehr signifikant wie in der ANOVA mit allen Resultaten der Trivia-Fragen ($p = 0{,}003$).

Die Ergebnisse des Erinnerungstests zeigen somit keinen signifikanten Einfluss von ChatGPT auf die Erinnerungsleistung.

Überlegungen, woran das liegen könnte, werden später in der Diskussion dargelegt. Zunächst werden noch die Ergebnisse aus den restlichen Fragen im Fragebogen vorgestellt. Nachfolgend wird auf die einzelnen Hypothesen eingegangen.

4.2.1 Hypothesen

Die übergeordnete Arbeitshypothese lautet:

> *Die Erinnerungsleistung in Bezug auf die mit ChatGPT erarbeiteten Inhalte ist nach zwei bis drei Wochen signifikant schlechter gegenüber den Inhalten, die ohne ChatGPT erarbeitet wurden.*

Diese Hypothese konnte aufgrund der nicht signifikanten Ergebnisse nicht bestätigt werden. Ein Blick auf die Mittelwerte (z. B. Google-Gruppe 2 M = 12.600 und der ChatGPT-Gruppe 4 M = 11.042) lässt den Trend vermuten, dass die ChatGPT-Gruppen eine leicht schlechtere Erinnerungsleistung aufweisen als die Google-Gruppen. Diese Vermutung wird in der Diskussion wieder aufgegriffen.

Die untergeordneten Hypothesen gehen allesamt auf das Zusammenspiel von Erinnerungsleistung und der Variable des Zugriffs ein. Da die Komponente des Zugriffs schon in der ersten ANOVA als signifikante Variable ausgeschlossen werden konnte, gelten die Hypothesen 1.1, 1.2 und 2.2 als nicht bestätigt. Es hat sich ebenfalls nicht bestätigt, dass die Google-Gruppe 1 über alle Gruppen hinweg die beste Erinnerungsleistung aufweist (H2.2).

H1.1 ➔ Wenn Antworten mit ChatGPT erarbeitet werden, die Probanden überzeugt sind, dass ChatGPT nicht zugänglich bleibt, und man sie dazu auffordert, sich die Antworten zu merken, dann wird besser erinnert als in ChatGPT-Gruppe 4 und Google-Gruppe 2.

H1.2 ➔ Wenn Antworten mit ChatGPT erarbeitet werden, die Probanden überzeugt sind, dass ChatGPT weiterhin zugänglich bleibt, und dazu aufgefordert wird, sich an die Antworten zu erinnern, dann wird weniger gut erinnert als in der ChatGPT-Gruppe 3.

H2.1 ➔ Wenn die Probanden überzeugt sind, dass Google nicht zugänglich bleibt, und sie dazu aufgefordert werden, sich die richtigen Antworten zu merken, dann wird über alle Gruppen hinweg besser erinnert.

H2.2 ➔ Wenn die Probanden überzeugt sind, dass Google weiterhin zugänglich bleibt, und sie dazu aufgefordert werden, sich die richtigen Antworten zu merken, dann wird weniger gut als in Google-Gruppe 1 und besser als in den ChatGPT-Gruppen erinnert.

Die Google-Gruppen wiesen tendenziell leicht höhere Mittelwerte in der Erinnerungsleistung auf. Da das Wissen darüber, ob es Zugriff oder keinen Zugriff geben würde, die Ergebnisse nicht veränderte, sind alle diese Hypothesen nicht bestätigt.

4.2.2 Posthoc Tests

Im Rahmen der Interpretation der Daten wurde im Nachhinein noch eine weitere Variable eingefügt. Diese Variable beruht auf den Beobachtungen der Intensität der Nutzung der Probanden im Rahmen der Experimente.
Diese Variable definiert, welche der Probanden aus den ChatGPT-Gruppen sich intensiver mit den Inhalten durch Rückfragen oder Ähnliches beschäftigt haben. Die Berechnungen mit dieser Variable konnte jedoch keinen nachweislichen Einfluss auf die Erinnerungsleistung aufzeigen. Die Ergebnisse ergeben ein sehr zerstreutes Bild der Erinnerungsleistung, wenn man die Intensität der Nutzung hinzufügt.
Nachfolgend illustrieren zwei Strichdiagramme von beiden Messzeitpunkten dieses zerstreute Ergebnis.

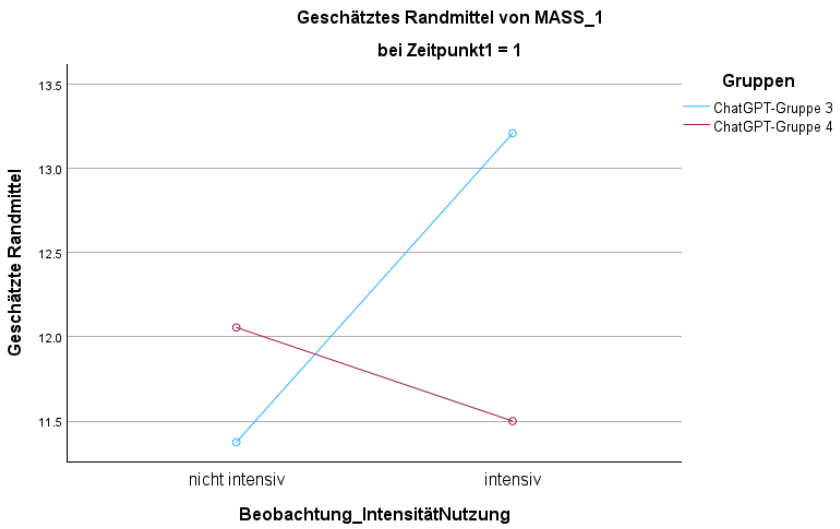


Abbildung 11: Strichdiagramm Erinnerungsleistung mit Beobachtungsvariable Messzeitpunkt 1 (SPSS,2024)

Wie in der Grafik (s. *Abbildung 11*) zu erkennen ist, hat eine intensive Nutzung keine bessere oder schlechtere Erinnerungsleistung zu bedeuten. Das lässt sich durch die weite Schere zwischen den zwei intensiven Gruppen erkennen. Dasselbe Bild zeigt sich auch in der nächsten Grafik.

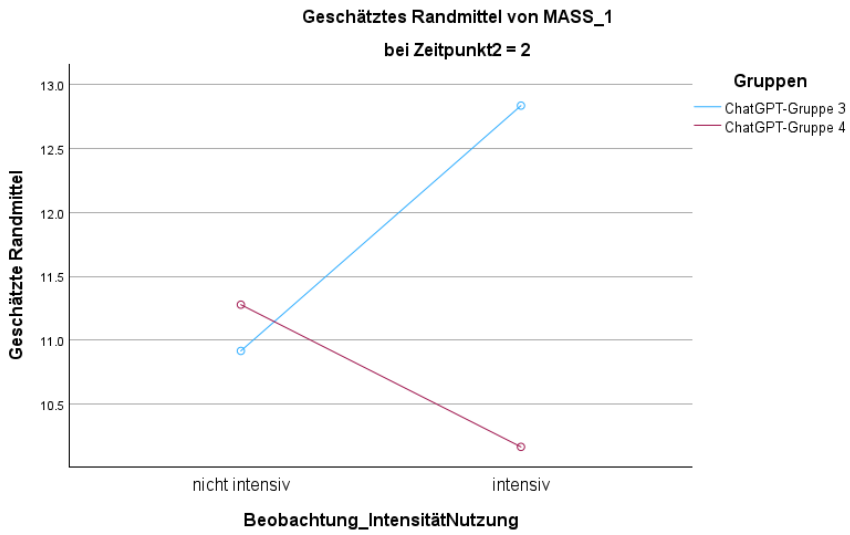


Abbildung 12: Strichdiagramm Erinnerungsleistung mit Beobachtungsvariable Messzeitpunkt 2 (SPSS, 2024)

Auch eine nicht-intensive Nutzung bringt keine bessere oder schlechtere Erinnerungsleistung hervor, wobei die nicht intensiven Gruppen von dem Mittelwert der Erinnerungsleistung her näher beieinander liegen und tendenziell etwas schlechtere Mittelwerte erreichten, wobei diese wie erwähnt in keiner Weise irgendwelche Signifikanzen aufwiesen.

4.3 Deskriptive Statistiken und qualitative Ergebnisse des Fragebogens

Im Fragebogen wurden den Teilnehmenden Aussagen vorgelegt, Multiple-Choice-Fragen sowie offene Fragen rund um ChatGPT gestellt. Die Antworten der Teilnehmenden sollten mögliche Erklärungen und Hintergrundinformationen zu den Ergebnissen aus dem Erinnerungstest liefern. Die Fragen wurden von allen Teilnehmenden beantwortet, einschliesslich der Google-Gruppen. Nachfolgend werden nur die wichtigsten Ergebnisse aus dem Fragebogen vorgestellt. Die einzelnen Items von «ChatGPT und Erinnerung», «ChatGPT-Nutzung» und «Speicherort Aussagen» wurden jeweils als eigene Variable geprüft. Es wurde in SPSS geschaut, ob die erstellten Variablen «Nutzung», «Erinnerung» und «Speicherort» mit der Erinnerungsleistung interagierten. Auch wenn die einzelnen Items der Frage miteinander korrelierten, hatten diese als Variable die Erinnerungsleistungen der Gruppen negativ beeinflusst, sodass sogar die Signifikanzen der einzelnen Messzeitpunkte verschwanden.

4.3.1 ChatGPT und Erinnerung

Den Teilnehmenden wurden Aussagen vorgelegt und sie durften sich auf einer Likert-Skala von 1-6 einordnen (1 = stimme überhaupt nicht zu und 6 = stimme voll und ganz zu). Da alle Antworten der Teilnehmenden einen Mittelwert von mind. 4 und aufwärts hatten, lässt sich

ableiten, dass ein Grossteil der Probanden den Statements *eher zustimmte* (theoretische Mitte der Likert-Skala 3.5).

1. ChatGPT beeinflusst das Wissen von Menschen. **(*M* = 4.40; *SD* = 0,963)**
2. ChatGPT beeinflusst die Erinnerungsleistung / Gedächtnis der Menschen. **(*M* = 4.00; *SD* = 1,348)**
3. ChatGPT beeinflusst die Art, wie Informationen erinnert werden. **(*M* = 4.09; *SD* = 1,203)**

4.3.2 Erinnerungsmethodik

Die Probanden wurden ebenfalls gefragt, welche Strategie sie verwendeten, um sich die Antworten zu merken (Plechatá et al., 2019). Hierfür wurden ihnen die folgenden Mehrfachantworten vorgelegt. Die genauen Nennungen der Merkmethoden befinden sich im Anhang K. Die meistverwendete Merkmethode der Teilnehmenden stellt das «Auswendiglernen» dar. An zweiter Stelle wurde die Visualisierung der Gegenstände im Kopf genannt und am wenigsten verwendet wurde die Methode des Geschichtenerzählens.

4.3.3 Einschätzung der Wiedererkennung der Antworten bei Aufgabe 2

Die Probanden wurden ebenfalls gefragt, ob sie die korrekten Antworten aus der Aufgabe 2 (unterstützte Reproduktion) wiedererkennen konnten. Die Probanden konnten sich auf einer Skala von 1-5 eintragen (1 = definitiv nein, 5 = definitiv ja). Die theoretische Mitte dieser 5er-Likert-Skala entspricht somit dem Wert 2.5. Der Mittelwert der Antworten ist mit dem Mittelwert 2.80 leicht über dieser theoretischen Mitte (*SD* = 1,290). Somit haben die Probanden die Antworten eher wiedererkannt («eher ja»).

4.3.4 ChatGPT Nutzung

Den Probanden wurden Aussagen zur Nutzung von ChatGPT vorgelegt, bei denen sie sich ebenfalls auf einer 6er-Likert-Skala eintragen mussten (1 = trifft überhaupt nicht zu, 6 = trifft voll und ganz zu). Die theoretische Mitte war hiermit 3.5. Nachfolgend sind die Aussagen aus dem Fragebogen mit den Mittelwerten aufgeführt:

1. Ich nutze ChatGPT oft auch für kleine Aufgaben. Es nimmt mir Arbeit ab. **(*M* = 3.11; *SD* = 1,434)**
2. Ich nutze ChatGPT absichtlich nicht oft, um mich nicht zu stark daran zu gewöhnen und darauf angewiesen zu sein. **(*M* = 2.13; *SD* = 1,079)**
3. Ich nutze ChatGPT für Ideengenerierung. **(*M* = 4.56; *SD* = 1,307)**
4. Ich nutze ChatGPT, um meine bereits entwickelten Konzepte weiterzuentwickeln. **(*M* = 3.56; *SD* = 1,575)**
5. Ich nutze ChatGPT öfter als noch vor 6 Monaten. **(*M* = 4.02; *SD* = 1,574)**
6. Ich finde ChatGPTs Antworten vertrauenswürdig. **(*M* = 3.18; *SD* = 0,936)**
7. Ich hinterfrage ChatGPTs Antworten mit meinem eigenen Fachwissen. **(*M* = 4.89; *SD* = 1,191)**
8. Ich prüfe ChatGPTs Antworten mit anderen Quellen. **(*M* = 4.49; *SD* = 1,359)**

Alle Aussagen bis auf die Aussagen 1, 2 und 6 sind im Mittelwert über der theoretischen Mitte von 3.5. Die befragten Personen geben an, ChatGPT eher *nicht* für kleinere Aufgaben einzusetzen (*M* = 3.11), ebenfalls scheinen die Probanden in die Antworten von ChatGPT bislang eher wenig Vertrauen zu haben (*M* = 3.18). ChatGPT absichtlich nicht oft zu nutzen,

um in keine Abhängigkeit zu geraten, scheint für die befragten Personen eher weniger der Fall zu sein (*M* = 2.13). Die befragten Personen benutzen ChatGPT eher für die Ideengenerierung und Konzeptentwicklung. Ebenso wird ChatGPT von den meisten Teilnehmenden nunmehr öfter genutzt als in den letzten 6 Monaten. Des Weiteren hinterfragen die Probanden ChatGPTs Antworten unter Rückgriff auf eigenes Fachwissen und prüfen eher noch unter Einbeziehung anderer Quellen nach.

4.3.5 Speicherort Aussagen

Etwas abstrakter waren die Aussagen zum «Speicherort». Mit diesen Aussagen sollte die Komponente «Zugriff / kein Zugriff» auf ChatGPT oder Google wie in den Studien von (Nagam, 2023; Sparrow et al., 2011; Storm et al., 2017) konkretisiert werden. Die Probanden konnten sich auf einer 6er-Likert-Skala einordnen (theoretische Mitte 3.5). Alle Mittelwerte und Standardabweichungen der Aussagen zum Speicherort befinden sich im Anhang L. Die befragten Personen sind der Meinung, dass sie Informationen, die sie interessant finden oder mit denen sie sich intensiv beschäftigen, besser erinnern als Informationen, die unverständlich sind. Wo die Informationen gespeichert sind und ob diese später wieder abrufbar sind, scheint gemäss den Teilnehmenden keinen Einfluss auf die Erinnerungsleistung zu haben.

4.3.6 Einstellungsveränderung

Die erste qualitative Frage im Fragebogen sollte die Einstellungsveränderung zu ChatGPT im letzten Jahr ermitteln. Die Einstellungsentwicklung der Probanden zeigt ein positives Bild. Negative Aussagen gab es von keinem Probanden. Die meisten äusserten sich eher positiv gegenüber ChatGPT. Negativere Äusserungen gehen eher in die Richtung einer Skepsis gegenüber der Vertrauenswürdigkeit der Antworten und der Weiterentwicklung von ChatGPT. Die Antworten der Probanden wurden in drei Kategorien unterteilt: skeptisch, neutral und positiv. Das Kategoriensystem befindet sich im Anhang M und die Ankeraussagen sind optisch hervorgehoben.

4.3.7 Veränderung der Erinnerung seit ChatGPT

Die zweite qualitative Frage sollte ermitteln, wie die Probanden ihre Erinnerungsleistung bzw. eine mögliche Veränderung dieser seit ChatGPT einschätzen.
Der Grossteil der Probanden geht davon aus, dass die Erinnerungsleistung seit Einführung von ChatGPT unverändert geblieben ist. Bezüglich Verschlechterung gibt es mehr Nennungen, wobei eher von einer kleinen Veränderung ausgegangen wird. Ebenso spannend waren Aussagen dahingehend, dass die «Nicht-Erarbeitung» von Informationen dazu führt, dass diese nicht erinnert werden. Diese Aussage passt zu den quantitativen Ergebnissen aus den Erinnerungstests der ChatGPT-Gruppen. In der Diskussion wird

dieses Ergebnis genauer erläutert. Die Antworten wurden in drei Kategorien aufgeteilt: «Unverändert / Neutral», «Verbessert / Positiv» und «Verschlechtert / Negativ». Das Kategoriensystem mit den Probandenaussagen und Ankerbeispielen befindet sich im Anhang N.

Wie bereits zu Beginn des Kapitels «Deskriptive Statistiken des Fragebogens» ausgeführt, wurden die unterschiedlichen Bewertungen aus dem Fragebogen zur ChatGPT-Nutzung, zu Erinnerung und ChatGPT sowie zum Speicherort nach der Variablenerstellung anschliessend mit einer ANOVA geprüft. Es wurde geschaut, ob die Variablen eine Interaktion anzeigen und die mögliche Kovariate zu einem Unterschied in der Erinnerungsleistung führt.
Bei beiden Variablen war das nicht der Fall und es konnte kein Einfluss auf die Erinnerungsleistung nachgewiesen werden. Im Gegenteil: Die Variablen haben die Signifikanzen bei der Erinnerungsleistung zu Zeitpunkt 1 und 2 vermindert und entfernt. Deshalb sind die Berechnungen der ANOVA mit Kovarianz nicht nochmals im Detail aufgeführt.

5 Diskussion

Die Ergebnisse zeigen somit auf, dass die Nutzung von ChatGPT gegenüber der Nutzung von Google keine Auswirkungen auf die Erinnerungsleistung hat. Hierfür gibt es eine Vielzahl an möglichen Begründungen, die nachfolgend erläutert werden. Zunächst werden allerdings die Hauptergebnisse noch einmal zusammengefasst und früheren Forschungsergebnissen gegenübergestellt.

Die Studie brachte keine signifikanten Unterschiede bei der Erinnerungsleistung hervor. Wie erwähnt, ist ein leichter Trend erkennbar, dass die ChatGPT-Gruppen schlechtere Leistungen aufweisen als die beiden Google-Gruppen. Dieser ist statistisch jedoch nicht signifikant.
Ebenfalls ist die Komponente des Zugriffs (die Überzeugung der Probanden, dass die Informationen gespeichert werden und somit weiterhin für sie zugänglich bleiben) gemäss den Ergebnissen nicht signifikant. Es wurde kein Unterschied in den Gruppen nachgewiesen, alle Gruppen verhielten sich ähnlich. Zu wissen, dass im zweiten Teil des Experiments kein Zugriff mehr auf ChatGPT oder Google möglich sein würde, hätte gemäss der Literaturrecherche einen signifikanten Unterschied herbeiführen sollen. Dies war nicht der Fall. Unter Umständen liefern die Antworten aus den Fragebogenitems zum Speicherort einen Hinweis darauf. Denn gemäss den Probanden hat der Ablageort und die Abrufmöglichkeit von Informationen für sie keinen Einfluss auf ihr Erinnerungsvermögen. Unabhängig davon könnte ein weiterer Grund darin liegen, dass die Menschen heutzutage daran gewöhnt sind, Informationen in Systemen abzulegen und sich eher zu merken, wo diese Informationen gespeichert sind, statt die Informationen selbst zu behalten (Sparrow et al., 2011). Warum sich die Probanden trotz Aufforderung an gewisse Informationen nicht erinnern konnten, kann daran liegen, dass solche Situationen in der heutigen Zeit kaum mehr vorkommen und die Probanden einfach davon ausgingen, weiterhin Zugriff auf Recherche-Möglichkeiten zu erhalten – ein bisschen nach dem Motto: «Das Internet vergisst nie. Also weshalb sollte ich mir diese Informationen merken».

Signifikant nachgewiesen werden konnte der Erinnerungszerfall über die Zeit hinweg. Nachvollziehbarerweise konnten sich sämtliche Probanden nach 2-3 Wochen signifikant schlechter erinnern als noch bei dem ersten Erinnerungstest. Im Einklang mit der Theorie zum Gedächtnis ist ein Erinnerungszerfall ohne Üben (Rehearsal) nach gewisser Zeit erkennbar (Gruber, 2018).

Es konnte nachgewiesen werden, dass die unterstützte Reproduktion (mithilfe der Trivia-Fragen) jeweils bessere Werte erzielte als die freie Reproduktion. Diese Werte waren sogar signifikant.

Die angenommenen Hypothesen (H1.1, H1.2, H2.1 und 2.2) konnten nicht bestätigt werden. Wie bereits erwähnt, sind leichte Trends spürbar, denn die Google-Gruppen wiesen eher eine bessere Erinnerungsleistung auf. Es ist jedoch möglich, dass Google und ChatGPT bei der Nutzung geringere Unterschiede aufweisen als die jeweiligen Vergleichsobjekte aus der Literatur. Dort wurde stets die Internetnutzung mit herkömmlichen Mitteln der Informationsbeschaffung, wie z. B. dem Rückgriff auf vorheriges Wissen und Bücherrecherchen, verglichen. Dieser Vergleich wäre heutzutage jedoch nicht mehr praxisnah.

Ein weiterer Grund für die nicht signifikanten Werte könnte die unterschiedliche Nutzung von Google in den Google-Gruppen sein. Im Experiment konnte die Autorin beobachten, wie unterschiedlich die Probanden der ChatGPT-Gruppen die Aufgabe lösten. Während der Experimente war es auffallend, dass nur wenige Probanden der ChatGPT-Gruppen die Antworten von ChatGPT tatsächlich hinterfragten. Ein Grossteil der ChatGPT-Probanden übernahm die Antworten von ChatGPT direkt. Hingegen überprüften jene Probanden, die im Screener angegeben hatten, bereits erfahrene ChatGPT-Nutzende zu sein, ihre Antworten. Sie stellten entweder Rückfragen wie «Bist du sicher?» oder sie stellten andere Fragen, um die Perspektive von ChatGPT zu erweitern und zu schauen, ob dieser immer noch die korrekte Antwort geben würde.
Beruhend auf diesen Beobachtungen wurden Post-hoc-Tests berechnet, die jedoch keinen nachweislichen oder signifikanten Einfluss auf die Erinnerungsleistung aufzeigen.
Ebenfalls konnte während der Experimente beobachtet werden, dass die Google-Gruppen sich bei der Recherche viel intensiver mit den Antworten auseinandersetzten. Die Probanden recherchierten sicherheitshalber mehrmals, selbst wenn die Antwort bei Google direkt als erster Eintrag vorgeschlagen wurde. Nach diesen Beobachtungen ist es nicht erstaunlich, dass sich die Google-Gruppen etwas besser als die ChatGPT-Gruppen erinnern konnten. Diese leichten Unterschiede sind, wie eingangs erwähnt, aber nicht signifikant.

Alle Experimente der drei Hauptstudien wurden in Amerika durchgeführt. Es ist also möglich, dass kulturelle Faktoren bei der Durchführung des Experiments eine Rolle spielen, so zum Beispiel hinsichtlich der Komponente des Zugriffs. Mit Blick auf die Literatur konnte man davon ausgehen, dass die Antworten trotz expliziter Aufforderung, sie sich zu merken, nicht

erinnert werden, wenn die betroffene Person davon ausgeht, dass die Antworten irgendwo gespeichert werden. Vielleicht sind die Schweizer etwas genauer, weshalb sie der expliziten Aufforderung eher Folge leisten als Probanden in Amerika.

Ein weiter, nicht unwichtiger Faktor stellt die Relevanz der Themen dar. Es ist gut möglich, dass die Trivia-Fragen den Probanden nicht relevant genug erschienen waren, um sich während der Recherche verstärkter damit auseinanderzusetzen und sich die Antworten anschliessend zu merken.

In zwei der drei Hauptstudien, die zur Entwicklung des Experimentdesigns herangezogen wurden, wurden weniger Probanden getestet als in der vorliegenden Studie und trotzdem konnten dort signifikante Ergebnisse berechnet werden (Sparrow et al., 2011; Storm et al., 2017). Dieser Umstand lässt die Vermutung zu, dass die Ergebnisse der vorliegenden Studie nicht wegen der Anzahl der Probanden nicht signifikant sind.

Beim Lernen der Trivia-Antworten könnte auch der Interferenzeffekt (Gruber, 2018) eine Rolle gespielt haben. Da die recherchierten Antworten, sei es mit Google oder ChatGPT, von den echten Lösungen abweichen konnten, waren wenige Probanden gezwungen, vorher bereits «gelernte» Inhalte im Kontakt mit neuen zu verlernen. Es gab sogar Fälle, in denen Probanden zwei bis drei Wochen nach dem Experiment noch die recherchierte Antwort statt der korrekten Antwort aus den Lösungen angaben. Dies könnte eine mögliche Interferenz der Erinnerung andeuten.

Gemäss dem Kapitel Ergebnisse kann die Forschungsfrage «Wie wirkt sich die ChatGPT-Nutzung auf die Erinnerungsleistung von Swisscom-Mitarbeitenden aus?» wie folgt beantwortet werden: Die fehlende Signifikanz lässt stark annehmen, dass die Nutzung von ChatGPT keinen Einfluss auf die Erinnerung hat. Wie bereits oben erwähnt, sind Trends in den Daten spürbar. Die Probanden, die im Experiment mit ChatGPT gearbeitet haben, scheinen sich leicht weniger gut erinnern zu können. Die Gruppen sind im Vergleich zu den Google-Gruppen jedoch nicht so viel schlechter, dass es signifikant wäre. Somit sind diese Ergebnisse gerade auch für die Praxispartnerin nicht alarmierend, sodass in Zukunft zum Wohl der Mitarbeitenden vermehrt auf die Nutzung von ChatGPT verzichtet oder diese stark angepasst werden sollte.

5.1 Implikationen der Ergebnisse

Die Ergebnisse der Studien sind für die Vorhaben der Praxispartnerin Swisscom strategisch relevant. Die Studie weist nach, dass die ChatGPT-Nutzung keinen Einfluss auf die Erinnerungsleistung der Swisscom-Mitarbeitenden hat. Dies sollte zukünftige AI-Projekte, bei denen ChatGPT eingesetzt werden soll, simpler gestalten, da eine mögliche Auswirkung nicht stark berücksichtigt werden muss.
Da es jedoch noch generell wenig Forschung zu ChatGPT-Einflüssen auf arbeitspsychologisch relevante Konstrukte hat, sollte der Einsatz von ChatGPT immer auf die Umstände der Mitarbeitenden Bedacht nehmen. Bislang kann ChatGPT Menschen bei der Arbeit unterstützen, aber sie noch nicht in allen Tätigkeitsbereichen ersetzen. Es sollte auch primär nicht darum gehen zu ersetzen, sondern gemeinsam zusammenzuarbeiten. Das bedingt jedoch, dass die Mitarbeitenden sich mit ChatGPT auseinandersetzen und dessen Schwächen und Stärken kennen.

Die Ergebnisse zur unterschiedlichen Erinnerungsleistung bei freier und unterstützter Reproduktion zeigen, dass Menschen immer stärker Informationen an Systeme abgeben und diese durch die Unterstützung der Systeme auch schneller wieder aus dem Gedächtnis abrufen können. Dies lässt eine Weiterentwicklung des transaktionalen Gedächtnisses vermuten. Wir speisen dem System Informationen ein und müssen uns diese folglich nicht mehr merken, da das System diese verwaltet. Durch diese Speicherung von Informationen in anderen Systemen kann jedoch auch eine mögliche Abhängigkeit entstehen.
Diese Ergebnisse befürworten aktuelle Produktentwicklungen wie zum Beispiel Copilot Notebook (Ortiz, 2024). Mit Copilot Notebook wird den Nutzenden die Möglichkeit gegeben, alles, was jemals auf dem Gerät angeschaut wurde, zu einem späteren Zeitpunkt wieder abzurufen. Hierfür muss der Nutzende Copilot Notebook nur fragen und dieser sucht in der Datenbank die gespeicherten Screenshots heraus (Ortiz, 2024). Solche Systeme unterstützen also unser transaktionales Gedächtnis stark.
In diesem Zusammenhang wird es auch für den Nutzenden dieser Systeme immer wichtiger, gute Fragen oder Prompts stellen zu können, um wieder an die Informationen zu gelangen. Wie sich auch schon in den Beobachtungen gezeigt hat, begaben sich die intensiven ChatGPT-User eher in ein Gespräch mit dem Chatbot und prüften dessen Antworten mit Gegenfragen. Diese Dialogs- oder auch Prompting-Kompetenzen werden in Zukunft mehr an Relevanz gewinnen. Da dieses Verhalten jedoch nur eine kleine Anzahl der Probanden zeigte, ist es fraglich, in welche Richtung sich diese Nutzung weiterentwickeln wird. Auch die Kompetenz des «Critical Thinking» (des kritischen Hinterfragens/Denkens) ist in diesem Kontext relevant und in Unternehmen sowie in der Bildung zu fördern.

5.2 Reflektion und Limitationen

Die vorliegende Studie weist Stärken und Schwächen auf, die nun diskutiert werden. Klare Stärken dieser Studie waren das bis ins kleinste Detail geplante Experiment mit intensiver Literaturrecherche, um eine gewisse Reliabilität sicherzustellen. Bei der Durchführung des Experiments sowie bei der Planung wurde bereits darauf geachtet, jegliche Störungen auf ein Minimum zu reduzieren. Es war ein Erfolg, denn in allen 48 Experimenten gab es nur drei technische Störungen, die erfolgreich innert fünf bis zehn Minuten behoben werden konnten.
Die Aufgabenstellung wurde vorab intensiv diskutiert und die Autorin entschied sich schliesslich für kulturell passende Trivia-Fragen, da die Aufgabenstellung kein Vorwissen benötigen und nicht Personen bevorteilen sollte, die sich in einem spezifischen Arbeitskontext sehr gut auskennen. Alle Probanden sollten die gleichen Chancen erhalten. Natürlich lässt sich dabei nicht ausschliessen, dass sich gewisse Personen lieber mit Allgemeinwissensfragen oder lustigen Fakten auseinandersetzen (wie zum Beispiel im Rahmen von Quizspielen). Das hätte ein mögliches Ausschlusskriterium für das Experiment sein können, was die Autorin nun zu den negativen Aspekten der Studie bringt.
Das Bewertungsraster wurde nach den ersten Experimenten entwickelt und hätte vorab bereits definiert werden können, um unter Umständen eine noch objektivere Bewertung zu gewährleisten.
Den Probanden wurde in der Einladung erklärt, man würde kognitive Leistungen testen. Jedoch wurde einigen Probanden nach dem Lesen der Instruktionen schnell klar, dass es um das Gedächtnis ging. Daraufhin gab es einige Fragen, die umgangen werden mussten. Des Weiteren wurde in den Instruktionen nicht genau definiert, welche Aufgaben nach der Pause folgen würden. Dies wurde in keinem Pretest erkannt und somit zog es sich durch alle Experimente, dass die Teilnehmenden Fragen zur Aufgabe nach der Pause stellten. Dieser Interpretationsspielraum könnte einen Einfluss darauf gehabt haben, wie gut die Teilnehmenden sich an die Antworten erinnern wollten, und liefert möglicherweise einen Grund dafür, weshalb die Komponente «Zugriff» keinen Einfluss zu haben schien.

Es ist schwer, eine Antwort auf die Frage zu finden, wie dieses Studiendesign verbessert werden könnte. Denn das Experiment wurde gemäss der wenigen vorhandenen Literatur so ähnlich designt, wie es in Anbetracht der neuen Technologie ChatGPT möglich war. Die Anzahl der Probanden wäre gemäss Berechnung ausreichend gewesen, um statistische Relevanz zu gewährleisten, wenn die Unterschiede tatsächlich signifikant gewesen wären. Wie bereits oben erwähnt, hatten nicht alle der anderen Studien mehr Probanden rekrutiert. Die aktuellste Studie von (Nagam, 2023) hatte sogar nur zwanzig Probanden getestet.

Ob noch schwierigere oder relevantere Fragen hätten verwendet werden können, ist nicht schlüssig zu beantworten. Was die Probanden als einfache oder schwierige Fragen definieren, hängt stark von deren Vorwissen ab und dieses war in der Stichprobe sehr divers.
Im Screener hätte das Erinnerungsvermögen etwas detaillierter erfragt werden können. Dies hätte unter Umständen zu anderen Verteilungen der Probanden in den Gruppen geführt.
Wie oben beschrieben wäre es nicht praxisnah gewesen, dieses Experiment wie in der Literatur mit analogen Medien (Büchern) als Vergleichsobjekt durchzuführen, da heute kaum Mitarbeitende ohne Internet arbeiten.

5.3 Handlungsempfehlungen zur Nutzung von ChatGPT

Es gibt unterschiedliche Prompt-Techniken, die je nach Zielsetzung angewendet werden können.
Eine der bekanntesten bei Swisscom ist die folgende Prompting-Technik: «Du bist ein Forscher, du forschst zum Thema Erinnerungsleistung im Rahmen einer Masterarbeit; suche mir relevante Literatur von den Jahren XY im Zusammenhang mit XY.»
Prompting-Techniken werden zurzeit stark erforscht und es gibt sehr viele unterschiedliche Strategien, in Bezug auf die es sich lohnt, sich vertieft damit auseinanderzusetzen. Da dies den Rahmen dieser Arbeit jedoch sprengen würde, werden diese Techniken hier nicht weiter beschrieben.

Die Nutzenden aus den Experimenten, die sich mit den generierten Inhalten stärker auseinandersetzten, verwendeten alle gewisse Strategien, auf die nachfolgend genauer eingegangen wird.

Dialogführung
Einige der Probanden führten mit dem Chatbot eine Unterhaltung, bei der weitere, für die Frage relevante Gesichtspunkte aufkamen. Ausgangslage, Kontext und Auftrag waren hier die wichtigsten Elemente. Der Chatbot wurde schrittweise an das Thema herangeführt, wobei dieser Prozess im Experiment aufgrund der Trivia-Fragen eher kurz ausgefallen ist.
Es ist jedoch ein empfehlenswerter Prozess, da man ChatGPT bitten kann, alle vorherigen Fragen und Antworten in der finalen Antwort zu berücksichtigen, um somit ein ausführliches und genaues Ergebnis zu erhalten.

Rückfragen

Probanden, die Rückfragen stellten, konnten in der Regel schneller eruieren, ob die von ChatGPT gelieferte Antwort stimmen konnte.
Es wurden Fragen gestellt wie: «Ich denke, dass das nicht stimmt.», «Ist das die einzige Option?» oder «Bist du sicher?». Aufgrund der möglichen Halluzinationen, die ChatGPT teilweise aufweist, sind solche Rückfragen in Fällen von Unsicherheit nützlich und hilfreich.

Perspektivenwechsel

Was ebenfalls eine interessante und erfolgreiche Strategie darstellte, war der Perspektivenwechsel. Probanden versuchten, Fragen zu stellen, die dazu führten, dass ChatGPT dazu gezwungen wurde, ein grösseres Bild der Ausgangslage zu erhalten. Somit wurden ähnliche Fragen gestellt, jedoch aus anderen Bereichen, oder es wurde etwas «herausgezoomt». Diese Strategie wurde verwendet, um zu prüfen, ob die Antworten von ChatGPT korrekt bleiben, wenn der Kontext erweitert wird oder etwas anders ist. Diese Strategie bedingt jedoch Experten- oder Fachwissen, um die Antworten von ChatGPT zu prüfen.

Die Handlungsempfehlungen für das Prompten sind nicht abschliessend, sie bieten aber drei wichtige Strategien: Dialogführung, Rückfragen und Perspektivwechsel, um die Antworten von ChatGPT zu verbessern und herauszufordern.

5.4 Ausblick auf zukünftige Forschung

Die vorliegenden Ergebnisse zeigen, dass die Nutzung von ChatGPT keinen signifikanten Einfluss auf die Erinnerungsleistung der Mitarbeitenden hat. Dennoch wurden nicht signifikante Trends beobachtet, die darauf hindeuten, dass Probanden, die mit ChatGPT gearbeitet haben, eine leicht geringere Erinnerungsleistung aufweisen, als es bei Google der Fall ist. Diese Studie gibt diesbezüglich Einblick darin, dass sich diese Unterschiede bei der Internetnutzung (Google) und ChatGPT-Nutzung sehr ähnlich verhalten.
Aufgrund der Tatsache, dass es wenige Studien zu diesem Thema gibt, ist es also empfehlenswert, eine Studie in ähnlichem und verbessertem Setting nochmals durchzuführen.
Für zukünftige Forschung bieten sich mehrere interessante Ansätze an, um diese Ergebnisse weiter zu untersuchen und zu vertiefen:
Da die vorliegende Studie möglicherweise kurzfristige Effekte untersucht hat, könnte es aufschlussreich sein, die Langzeiteffekte der ChatGPT-Nutzung auf die Erinnerungsleistung mit einer Längsschnittstudie zu analysieren. Hierbei könnte beispielsweise untersucht

werden, ob sich die Erinnerungsleistung über mehrere Monate hinweg verändert und ob sich mit der Zeit signifikante Unterschiede zwischen den Nutzenden von ChatGPT und anderen Informationsquellen wie Google ergeben.

Es wäre interessant zu untersuchen, wie sich die Nutzung von ChatGPT in unterschiedlichen Arbeitskontexten auf die Erinnerungsleistung auswirkt. Beispielsweise könnten verschiedene Arten von Aufgaben (z. B. kreative vs. analytische Aufgaben) oder unterschiedliche Komplexitätsgrade der zu erinnernden Informationen untersucht werden.

Die Untersuchung individueller Unterschiede, wie z. B. in Bezug auf kognitive Fähigkeiten, Vorerfahrungen mit AI-Technologien oder technologische Affinität, könnte aufzeigen, ob bestimmte Personengruppen stärker oder weniger stark von den Auswirkungen der ChatGPT-Nutzung betroffen sind. Darüber hinaus könnten Moderatoren wie die Art und Weise der Nutzung (intensiv vs. gelegentlich) oder die spezifischen Einsatzszenarien (z. B. Recherche vs. Entscheidungsunterstützung) analysiert werden.

Der Vergleich mit anderen AI-Tools (wie dem GAI Copilot oder Jasper für Texterstellung) ermöglicht einen weiteren Forschungsansatz. Die Studie könnte dahingehend erweitert werden, dass die Auswirkungen der Nutzung von ChatGPT mit anderen AI-basierten Tools verglichen werden. Dies könnte helfen zu verstehen, ob die beobachteten Trends spezifisch für ChatGPT sind oder ob sie auch bei der Nutzung anderer AI-Technologien auftreten.

Ergänzend zu quantitativen Analysen könnten qualitative Studien, wie Interviews oder Fokusgruppen, durchgeführt werden, um ein tieferes Verständnis dafür zu gewinnen, wie Mitarbeitende ChatGPT nutzen, welche Strategien sie dabei anwenden und wie sie die Nutzung wahrnehmen. Solche Studien könnten Aufschluss darüber geben, warum und wie die Nutzung von ChatGPT die Erinnerungsleistung beeinflusst. Auch Mixed-Methods-Ansätze mit weiteren Experimenten oder quantitativen Erhebungen unter Einbezug von qualitativen Methoden bieten eine weitere Option, um einerseits tiefgründige, gehaltvolle Aussagen und andererseits repräsentative Daten zu erhalten.

Künftige Studien könnten experimentelle Manipulationen einführen, um spezifische Mechanismen zu identifizieren, die der Beziehung zwischen ChatGPT-Nutzung und Erinnerungsleistung zugrunde liegen. Beispielsweise könnte untersucht werden, wie sich unterschiedliche Arten der Interaktion mit ChatGPT (z. B. dialogorientiert vs.

informationsabruforientiert) auf die Gedächtnisleistung auswirken – analog zu den Feststellungen dieser Interaktionsarten aus den Beobachtungen.

Die vorliegende Studie liefert wichtige erste Erkenntnisse für weitere Forschung. Die aufgezeigten Forschungspotenziale bieten vielfältige Möglichkeiten, um das Verständnis der Auswirkungen von AI-Nutzung auf die kognitiven Prozesse und das psychologische Wohl der Mitarbeitenden weiter zu vertiefen. Dies könnte nicht nur theoretische Implikationen für die kognitive Psychologie (Lernen und Gedächtnisforschung) und die Mensch-Technik-Interaktion (unter Einbezug der psychologischen Sicherheit) haben, sondern auch praktische Implikationen für den Einsatz von AI-Technologien am Arbeitsplatz.

6 Literaturverzeichnis

Arntz, M., Gregory, T., & Zierahn, U. (2016). *The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis.*

Baddeley, A. D., Eysenck, M. W., Anderson, M., & Anderson, M. C. (2010). *Memory* (Reprinted). Psychology Press.

Bai, L., Liu, X., & Su, J. (2023). ChatGPT: The cognitive effects on learning and memory. *Brain‐X,* 1(3), e30. https://doi.org/10.1002/brx2.30

Bernhardt, T., Maurer, K., & Frölich, L. (2002). Der Einfluss eines alltagsbezogenen kognitiven Trainings auf die Aufmerksamkeits- und Gedächtnisleistung von Personen mit Demenz. *Zeitschrift für Gerontologie und Geriatrie*, 35(1), 32–38. https://doi.org/10.1007/s003910200004

Bhandari, S. (2022, Februar 10). Künstliche Intelligenz vs. menschliche Intelligenz: Unterschied und Vergleich. Verfügbar unter https://askanydifference.com/de/difference-between-artificial-intelligence-and-human-intelligence/

Bundesrat. (2017). Auswirkungen der Digitalisierung auf Beschäftigung und Arbeitsbedingungen – Chancen und Risiken. Verfügbar unter https://www.newsd.admin.ch/newsd/message/attachments/50248.pdf

ChatGPT. (2023, Juni 11). Unterschiede zwischen GPT-3.5 und GPT-4. Verfügbar unter https://chatgpt.ch/unterschiede-zwischen-gpt-3-5-und-gpt-4/

CHIP. (2023, Juli 18). ChatGPT Grundlagen: Die Basics für KI-Anfänger erklärt. Verfügbar unter https://praxistipps.chip.de/chatgpt-grundlagen-die-basics-fuer-ki-anfaenger-erklaert_156259

Delcker, J. (2023, März 31). Wie verändert künstliche Intelligenz die Gesellschaft? Verfügbar unter https://www.dw.com/de/k%C3%BCnstliche-intelligenz-wie-revolutioniert-ki-die-gesellschaft/a-65192292

Ebbinghaus, H. (1885). *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie.* Leipzig, Duncker & Humblot.

Edwards, B. (2024, März 20). GPT-5 might arrive this summer as a “materially better” update to ChatGPT. Verfügbar unter https://arstechnica.com/information-technology/2024/03/openais-gpt-5-may-launch-this-summer-upgrading-chatgpt-along-the-way/

Fregin, M.-C., & Stops, M. (2023). Künstliche Intelligenz: Wer mit KI-Technologien arbeiten erfolgreich arbeiten will, sollte die Wirkungen valide abschätzen können. *IFO Schnelldienst, 76*(8/2023), 12–15.

Gruber, T. (2018). *Gedächtnis* (2. Auflage). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-56362-5

Haase, J., & Hanel, P. H. P. (2023). Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity. *Journal of Creativity, 33*(3), 100066. https://doi.org/10.1016/j.yjoc.2023.100066

Hobmair, H. (2013). *Psychologie* (5. Auflage). Bildungsverlag EINS.

Hundertmark, S. (2023, Juli 21). ChatGPT für Zusammenfassungen—Ein Leitfaden und Best Practices. Verfügbar unter https://www.sophiehundertmark.com/chatgpt-fuer-zusammenfassung/

Hünebeck, D. (2024). Keynote Speaker & Referent für Künstliche Intelligenz & ChatGPT. Verfügbar unter https://www.daniel-huenebeck.ch/chatgpt-schulung-kuenstliche-intelligenz-ki-fuer-unternehmen/

IBM. (2023, Juli 18). Was ist ein Chatbot? Verfügbar unter https://www.ibm.com/de-de/topics/chatbots

KI-Cafe. (2023, März 3). Was sind KI-Prompts und wie schreibst du sie? (Z.B. für ChatGPT, Dall-E etc.). Verfügbar unter https://ki-cafe.de/prompt-engineering/was-sind-prompts/

Krauss, P. (Hrsg.). (2023). Was ist Künstliche Intelligenz*? Künstliche Intelligenz und Hirnforschung: Neuronale Netze, Deep Learning und die Zukunft der Kognition.* Springer. https://doi.org/10.1007/978-3-662-67179-5_11

Marr, B. (2023, März 22). ChatGPT: What Are Hallucinations And Why Are They A Problem For AI Systems. Verfügbar unter https://bernardmarr.com/chatgpt-what-are-hallucinations-and-why-are-they-a-problem-for-ai-systems/

McKinsey. (2023). Studie: Generative KI kann zum Produktivitätsbooster werden. Verfügbar unter https://www.mckinsey.de/news/presse/genai-ist-ein-hilfsmittel-um-die-produktivitaet-zu-steigern-und-das-globale-wirtschaftswachstum-anzukurbeln

Müsseler, J., & Rieger, M. (Hrsg.). (2017). *Kontextinterferenz (contextual interference)* (3. Auflage). Springer. https://doi.org/10.1007/978-3-642-53898-8

Myers, D. G., Hoppe-Graff, S., & Myers, D. G. (2014). *Psychologie* (3., vollst. überarb. und erw. Aufl). Springer. https://doi.org/10.1007/978-3-642-40782-6

Nagam, V. M. (2023). Internet use, users, and cognition: On the cognitive relationships between Internet-based technology and Internet users. *BMC Psychology*, *11*(1), 82. https://doi.org/10.1186/s40359-023-01041-5

Ortiz, S. (2024). I tried Copilot Notebook: Microsoft's new AI tool offers two handy prompt features. Verfügbar unter https://www.zdnet.com/article/i-tried-copilot-notebook-microsofts-new-ai-tool-offers-two-handy-prompt-features/

Plechatá, A., Sahula, V., Fayette, D., & Fajnerová, I. (2019). Age-Related Differences With Immersive and Non-immersive Virtual Reality in Memory Assessment. *Frontiers in Psychology*, *10*, 1330. https://doi.org/10.3389/fpsyg.2019.01330

Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, *20*(1), 98–108. https://doi.org/10.1038/mp.2014.105

Raab, G., Unger, A., & Unger, F. (2022). Theorien zur Informationsverarbeitung. In G. Raab, A. Unger, & F. Unger (Hrsg.), *Marktpsychologie: Grundlagen und Anwendung* (S. 119–132). Springer Fachmedien. https://doi.org/10.1007/978-3-658-37688-8_8

Schaller, D., Wohlrabe, K., Wolf, A., Demary, V., Mertens, A., Fregin, M.-C., Stops, M., Gillhuber, A., Johannes Heckmann, D. W., & Grunwald, A. (2023). Künstliche Intelligenz: Chance oder Gefahr? Wie verändert der Einsatz von KI unsere Gesellschaft? *ifo Schnelldienst, 76*(08), 03–28.

Simplilearn.com. (2024). Types of Artificial Intelligence That You Should Know in 2024. Verfügbar unter https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/types-of-artificial-intelligence

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. *Science, 333*(6043), 776–778. https://doi.org/10.1126/science.1207745

Stangl, W. (2024). Einspeichermodell. Verfügbar unter. https://lexikon.stangl.eu/5116/einspeichermodell

Statista. (2024). Einfluss auf Arbeitsleistung. Verfügbar unter https://de.statista.com/statistik/daten/studie/1401314/umfrage/chatgpt-einfluss-auf-arbeitsleistung/

Sternberg, R. J. (1997). The concept of intelligence and its role in lifelong learning and success. *American Psychologist, 52*(10), 1030–1037. https://doi.org/10.1037/0003-066X.52.10.1030

Storm, B. C., Stone, S. M., & Benjamin, A. S. (2017). Using the Internet to access information inflates future use of the Internet to access other information. *Memory, 25*(6), 717–723. https://doi.org/10.1080/09658211.2016.1210171

Stuck, N. (2020, November 6). Erinnerungsvermögen: Wie das Gehirn Informationen speichert. Verfügbar unter https://cloud-minded.de/magazin/erinnerungsvermoegen/

Swisscom. (2024). Vision, Purpose, Werte, Strategie und Ziele. Verfügbar unter https://www.swisscom.ch/de/about/unternehmen/purpose-vision-values-and-goals.html

Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology, 61*(6), 923–929. https://doi.org/10.1037/0022-3514.61.6.923

7 Abbildungsverzeichnis

Abbildung 1: Mehr-Speicher-Modell von Atkinson und Shiffrin (Zoelch et al., 2019) 13
Abbildung 2: eigene Darstellung (2024).....14
Abbildung 3: Miro-Board-Experimentaufbau zur Illustration des Experiments (eigene Darstellung, 2024) 28
Abbildung 4: Schritt-für-Schritt-Versuchsaufbau (eigene Darstellung, 2024).....29
Abbildung 5: Altersverteilung der Stichprobe (Microsoft-Forms-Darstellung, 2024) 39
Abbildung 6: Strichdiagramme zur Variable Zugriff Messzeitpunkt 1 (links) Messzeitpunkt 2 (rechts) (SPSS, 2024) 40
Abbildung 7: Strichdiagramm beide Erinnerungstest (SPSS, 2024).....42
Abbildung 8: Strichdiagramm alle Gruppen Erinnerungsleistung über zwei Zeitpunkte (SPSS, 2024) 43
Abbildung 9: Detaillierte Strichdiagramme zu den einzelnen Erinnerungstests; Links: Erinnerungstests freie Reproduktion; Rechts: Erinnerungstest unterstützte Reproduktion (SPSS, 2024) 43
Abbildung 10: Strichdiagramm zu beiden Erinnerungsaufgaben (freie und unterstützte Reproduktion) Ergebnisse aller Gruppen (SPSS, 2024) 44
Abbildung 11: Strichdiagramm Erinnerungsleistung mit Beobachtungsvariable Messzeitpunkt 1 (SPSS,2024) 46
Abbildung 12: Strichdiagramm Erinnerungsleistung mit Beobachtungsvariable Messzeitpunkt 2 (SPSS, 2024) 47

8 Anhang

A: Instruktionen für die Probanden

Liebe/r Versuchsteilnehmende/r
Vielen Dank für dein Interesse sowie Unterstützung an meiner wissenschaftlichen Studie. Bitte lies dir die nachfolgenden Informationen sorgfältig durch.

Ziel des wissenschaftlichen Vorhabens

Die Studie hat zum Ziel den Einfluss von ChatGPT auf die Leistung des Menschen bei der Arbeit zu erforschen. Da grosse Sprachmodelle wie ChatGPT unsere Arbeitsprozesse und Aufgaben in Zukunft stark beeinflussen wird, möchte ich dieses Thema aus psychologischer Sicht genauer erforschen. Mit deiner Teilnahme hilfst du mir dabei, neue Erkenntnisse zu sammeln. Es erfolgt eine Aufnahme des Versuchs via MS Teams. Diese dient dazu, nach dem Versuch eine objektive Auswertung zu ermöglichen.

Aufgabe

Du wirst später im Miroboard Fragen sehen, die nacheinander erscheinen werden.
Deine Aufgabe besteht darin mit Google / ChatGPT die Antwort auf die Frage vor Ablauf des Timers zu finden.
Sobald der Timer der Frage abläuft, erscheint die nächste Frage und der nächste Timer wird aktiviert.
Für schwierigere Fragen erhältst du mehr Zeit.
Insgesamt wirst du 14 Fragen beantworten. Nach Beantwortung der 14 Fragen hast du nochmals 3 Min Zeit die Lösungen durchzulesen.

Wichtig: Google/ ChatGPT wird später nicht mehr zugänglich sein, deshalb wäre es gut, wenn du dir die korrekten Antworten merkst.
ODER
Wichtig: Du wirst später weiterhin Zugriff auf Google/ ChatGPT haben dennoch bitte ich dich dir die korrekten Antworten zu merken.

Nach dem Versuch

Nach dem Versuch erhältst du nach ca. 2-3 Wochen einen Fragebogen zugestellt, den du ohne Hilfsmittel beantworten sollst.

Einverständnis

Deine Antworten und Ergebnisse dieser Studie werden anonymisiert und vertraulich behandelt.
Wenn du mit dem Vorgehen wie oben beschrieben einverstanden bist, dann bitte ich dich die Einverständniserklärung zu unterschreiben

Probandennummer: Kondition (T1, T2, K1 oder K2) + Zahl des Geburtsmonats + 1.Buchstabe vom Vornamen der Mutter (Bsp. T104C)

Deine Probandennummer ist: _____

Besten Dank.

B: Einverständniserklärung für die Probanden

Fachhochschule Nordwestschweiz
Hochschule für Angewandte Psychologie

Einverständniserklärung
zur Teilnahme an der wissenschaftlichen Studie „Veränderung der Arbeitswelt durch ChatGPT psychologische Auswirkungen“

_____
Vor- und Nachname in Druckschrift

_____

Ich habe die Informationen zur Studie „Veränderung der Arbeitswelt durch ChatGPT kognitive Auswirkungen“ gelesen und verstanden. Eine Kopie dieser Informationen kann dir auf Wunsch per E-Mail zugesandt werden. Ausserdem hatte ich ausreichend Gelegenheit Fragen (z.B. zu Inhalt, Ziel, Verlauf und Risiken) zu stellen.

Ich erkläre mich damit einverstanden, an der Studie teilzunehmen. Meine Teilnahme erfolgt freiwillig.

Sämtliche von mir erhobenen Daten werden vollkommen anonym gespeichert. Ich habe Gelegenheit, meine Daten jederzeit zurückzuziehen und löschen zu lassen.

Ich weiss, dass ich die Möglichkeit habe, meine Teilnahme an dieser Studie jederzeit und ohne Angabe von Gründen abzubrechen, ohne dass mir daraus Nachteile entstehen.

Die Audio- und Videoaufzeichnung wird nur zur Datenauswertung der gegenwärtigen Studie verwendet und da- nach gelöscht.

Meine E-Mail-Adresse wird zur einmaligen Rückmeldung von Ergebnissen zur gegenwärtigen Studie verwendet.

Datenschutzklausel:
Ich erkläre, dass ich mit der im Rahmen der Studie erfolgenden Aufzeichnung von Studiendaten und ihrer Verwendung in anonymisierter Form einverstanden bin.

Hiermit bestätige ich, dass ich diese Erklärungen gelesen habe und damit einverstanden bin.

Ort, Datum und Unterschrift

C: Checkliste zur Überprüfung der ethischen Unbedenklichkeit von Untersuchungen in Projekten in der Aus- und Weiterbildung an der Hochschule für Angewandte Psychologie FHNW

1. Angaben zur Untersuchung

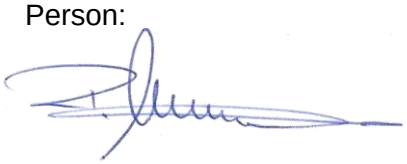
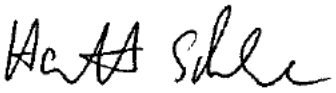
Name der verantwortlichen Person**:** Rosina Genovese
Name der betreuenden Person Dr. Hartmut Schulze
Name des Moduls bzw. des Weiterbildungsprogramms: Masterarbeit
Thema/Titel des Vorhabens**:** Veränderung der Arbeitswelt durch ChatGPT – Einfluss von ChatGPT auf die Erinnerungsleistung

2. Checkliste (siehe Erläuterungen ab Seite 3)

		Ja	Nein
1.	Besteht die Gefahr, dass den Teilnehmenden durch ihr Verhalten in der Studie oder durch ihre Nichtteilnahme an der Studie Nachteile entstehen?	☐	☒
2.	Ist die Teilnahme von minderjährigen Personen (d.h. Kinder und Jugendliche) und/oder Personen, die in ihrer Urteilsfähigkeit eingeschränkt sind, vorgesehen oder zumindest möglich?	☐	☐
3.	Wird es notwendig sein, dass Personen an der Studie teilnehmen, ohne dies vorab zu wissen und ohne ein informiertes Einverständnis gegeben zu haben (z.B. verdeckte Beobachtung von Personen)?	☐	☒
4.	Werden die Teilnehmenden mit Absicht unvollständig oder falsch über die Ziele und das Verfahren der Studie informiert (z.B. durch manipulierte Rückmeldungen über ihre Leistungen)?	☐	√
5.	Werden die Teilnehmenden gebeten, persönliche Erfahrungen (z.B. belastende Erlebnisse), sensitive Informationen (z.B. Drogenkonsum) oder sensitive oder belastende Einstellungen (z.B. politische Präferenzen) preiszugeben?	☐	√
6.	Falls während der Untersuchung die psychische Integrität der Teilnehmenden beeinträchtigt wird (z.B. reduzierte	☐	√

	Konzentrationsfähigkeit, Induktion von negativen Emotionen): Können nach der Studienteilnahme negative psychische Folgen eintreten?		
7.	Falls während der Studie die soziale Integrität der Teilnehmenden tangiert wird (z.B. in einer Gruppenstudie): Kann die Teilnahme auf der sozialen Ebene negative Folgen haben (z.B. bei den anderen Teilnehmenden erworbener «Ruf»)?	☐	√
8.	Wird den Teilnehmenden ein finanzieller Anreiz angeboten, der über die übliche Vergütung der Teilnahme an der Studie hinausgeht?	☐	√
9.	Verlangt der Auftragsgeber/die Auftragsgeberin eine Begutachtung der Studie durch eine Ethikkommission?	☐	√
10.	Verlangt der Auftragsgeber/die Auftragsgeberin oder die Gesetzgebung eine Registrierung der Studie?	☐	√

3. Unterzeichnung

Ort und Datum:	Name der verantwortlichen Person:	Unterschrift der verantwortlichen Person:
Olten, 14.12.23	**Rosina Genovese**	
Ort und Datum:	Name der betreuenden Person:	Unterschrift der betreuenden Person:
Olten, 14.12.23	**Dr. Hartmut Schulze**	

Falls mindestens eine der Fragen 1-10 der Checkliste mit JA beantwortet wurde, ist ein Ethikantrag bei der Studiengang- bzw. Programmleitung einzureichen.

Anhang: Erläuterungen zu den Fragen

Frage 1: Falls Studienteilnehmende in einem Abhängigkeitsverhältnis mit dem/der StudienleiterIn oder AuftraggeberIn der Studie stehen (z.B. weil sie in einem Anstellungsverhältnis stehen), muss Sorge getragen werden, dass den Teilnehmenden keine nachteiligen Konsequenzen aus der Teilnahme an der Studie entstehen können – z. B. eine schlechtere Bewertung ihrer Leistung, wenn sie an der Studie nicht in zufriedenstellender Weise teilgenommen haben. Das kann dadurch erreicht werden, dass die Anonymität der Teilnehmenden gegenüber der Person, zu der das Abhängigkeitsverhältnis besteht, gewahrt wird. Wenn die Anonymität sichergestellt ist, kann Frage 1 mit NEIN beantwortet werden.

Frage 2: Beispiele für nicht voll urteilsfähige Personen sind Kinder, Jugendliche, Personen mit geistigen Behinderungen, Personen mit Demenz, Personen mit psychischen Krankheiten. Für zusätzliche Anforderungen bei solchen besonders verletzbaren Personen siehe HFG, Kapitel 3.

Frage 3: Diese Frage bezieht sich auf Untersuchungen, bei denen das Verhalten von Personen ohne deren Wissen beobachtet oder experimentell beeinflusst wird. In allen anderen Fällen ist vorab eine Einverständniserklärung einzuholen, in der die potentiellen Teilnehmenden über folgende Punkte aufgeklärt werden: 1) den Zweck der Forschung; 2) die erwartete Dauer der Studie und das Vorgehen; 3) ihr Recht darauf, die Teilnahme abzulehnen oder sie auch während der Studie zu beenden; 4) absehbare Konsequenzen der Nichtteilnahme oder der vorzeitigen Beendigung der Teilnahme bzw. explizite Darlegung, dass keine Konsequenzen entstehen; 5) absehbare Faktoren, die möglicherweise die Teilnahmebereitschaft beeinflussen, wie z.B. potenzielle Risiken, Unbehagen oder mögliche anderweitige negative Auswirkungen, die über alltägliche Befindlichkeitsschwankungen hinausgehen; 6) den erwarteten Erkenntnisgewinn durch die Forschungsarbeit; 7) die Gewährleistung von Vertraulichkeit und Anonymität sowie ggf. deren Grenzen (d.h. wie werden die Daten anonymisiert [wenn möglich], wie lange werden sie aufbewahrt und wer hat Zugang zu den Daten); 8) evtl. Teilnahmeentschädigung und 9) an wen sie sich mit Fragen zum Forschungsvorhaben und zu ihren Rechten als Teilnehmende wenden können. Zudem werden potenziellen Teilnehmenden die Gelegenheit gegeben, Fragen zum Forschungsvorhaben zu stellen.

Frage 4: Diese Frage bezieht sich auf Untersuchungen, bei denen die Teilnehmenden gezielt getäuscht werden («Deception»). Das heisst, dass die Teilnehmenden über für sie wesentliche Aspekte der Untersuchung gezielt im Unklaren gelassen oder in die Irre geführt werden, so dass sie sich, wenn sie davon erfahren, getäuscht fühlen müssten. Dazu gehört beispielsweise falsche Leistungsrückmeldung und falsche Informationen über die Ziele der Untersuchung. Nicht gemeint ist mit dieser Frage der Umstand, dass Teilnehmende in aller Regel nicht vollständig über die wissenschaftlichen Hintergründe einer Untersuchung und über die Hypothesen aufgeklärt werden. Bei einer gezielten Täuschung muss ein Ethikgesuch eingereicht werden, in dem klar beschrieben wird, wie und mit welcher Begründung die Teilnehmenden getäuscht werden und wie ihnen das nachträglich so bald als möglich erklärt wird (d.h. sie erhalten ein Debriefing, inkl. der Möglichkeit, ihre Teilnahme und ihre Daten zurückzuziehen).

Frage 5: Die Frage bezieht sich auf die Erhebung von Informationen, die aus einem von zwei Gründen sensitiv sind. Zum einen geht es um Informationen, die in besonderem Masse vertraulich behandelt werden müssen, weil ihre Weitergabe für die Person zu Nachteilen führen könnte (z.B. politische Einstellungen oder Drogenkonsum). Zum anderen geht es um Informationen, deren Preisgabe für die Person mit starken Emotionen verbunden sein kann (z.B. traumatische Erfahrungen), so dass die Erhebungssituation eine unzumutbare emotionale Belastung mit sich bringen kann. Besonders heikel

sind z.B. Fragen über Gewalt in der Familie, sexuelles Verhalten, Suchtverhalten, Krankheiten, Freundschaften und Feindschaften, Zufriedenheit mit sich selbst und dem Umfeld, mögliche oder tatsächliche strafrechtlich relevante Handlungen, und sensitive, potentiell belastende - etwa beschämende - Einstellungen (z.B. politische Präferenzen).

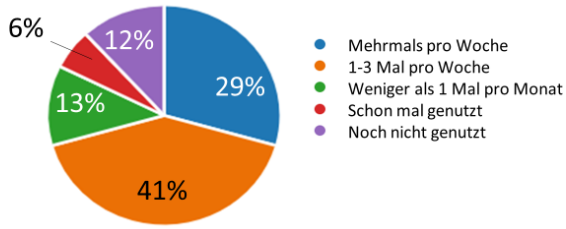
Frage 6: Hier ist zu unterscheiden, ob die Folgen der Intervention bedenklich sind oder nicht. Fragen oder Manipulationen, die sich in ihrer Intensität und Qualität nur geringfügig und nur vorübergehend auf die Teilnehmenden auswirken, sind unbedenklich. Hingegen sind Fragen oder Manipulationen, die eine längere und intensive Auswirkung auf die Teilnehmenden haben können (z.B. das Zeigen von Kriegs- und Verstümmelungsbildern) bedenklich, weil sie mit grossen Risiken und Belastungen verbunden sind (z.B. heftige emotionale Reaktionen). Teilnehmenden dürfen keine grösseren Risiken als solche, denen sie im Alltag ausgesetzt sind, zugemutet werden.

Frage 7: Nicht jede Gruppenstudie ist ethisch bedenklich, aber sie bringen unter Umständen die Gefahr mit sich, dass Personen in für sie unangenehme soziale Situationen gebracht werden, etwa wenn in einer Gruppendiskussion eine Konkurrenzsituation entsteht, bei der einzelne Teilnehmende offensichtlich unterliegen, oder wenn Personen die Situation als peinlich empfinden. Auch hier muss, wie bei den vorangegangenen beiden Fragen, eine Grenze gezogen werden zwischen kleineren Unannehmlichkeiten, die alltäglich und zumutbar sind (wie z.B. die Nervosität, die manche Menschen empfinden, wenn sie in Gruppen sprechen müssen) und solchen, die die Zumutbarkeitsgrenze überschreiten (wie z.B. persönliche Informationen mit Arbeitskolleginnen und -kollegen zu teilen).

Frage 8: Hier wird unterschieden zwischen der in einem Forschungsgebiet üblichen Aufwandsentschädigung für die Teilnahme an einer Untersuchung und einem finanziellen Anreiz, der spezifisch in der geplanten Studie eingesetzt wird, um ein Ziel der Studie (z.B. besonders hohe Leistungsmotivation) zu erreichen. Die DGPs gibt vor: «Psychologinnen und Psychologen vermeiden übertriebene oder unverhältnismäßige finanzielle oder anderweitige Anreize bei der Anwerbung von an der Forschung teilnehmenden Personen, wenn anzunehmen ist, dass solche Anreize zu einer Teilnahme nötigen würden" (siehe DGPs, 2016, Kapitel. 3.7, Artikel 7a).

Fragen 9 und 10: Wenn der Forschungsträger eine Begutachtung durch eine Ethikkommission verlangt, ist die Ethikkommission der APS die erste Anlaufstelle, sofern nicht ausdrücklich eine andere Ethikkommission verlangt oder aufgrund der Art der Untersuchung zuständig ist. Die Registrierungspflicht ist davon unabhängig.

D: ChatGPT-Nutzung im Alltag



ChatGPT Vorkenntnisse/ Erfahrung Verteilung aus MicrosoftForms Rekrutierungsfragebogen und -Screener

E: Verteilung Geschlecht

Gruppen * Geschlecht Kreuztabelle

Anzahl

		Geschlecht		
		männlich	weiblich	Gesamt
Gruppen	Googlegruppe 1	4	8	12
	Googlegruppe 2	5	7	12
	ChatGPT-Gruppe 3	7	5	12
	ChatGPT-Gruppe 4	5	7	12
Gesamt		21	27	48

Kreuztabelle Geschlechtsverteilung SPSS

F: Gruppenverteilung

Zwischensubjektfaktoren

		Wertbeschriftung	N
Gruppen	1.00	Googlegruppe 1	11
	2.00	Googlegruppe 2	10
	3.00	ChatGPT Gruppe 3	12
	4.00	ChatGPT Gruppe 4	12

Anzahl Probanden pro Gruppe

G: Zugriffskomponente

Tests der Innersubjekteffekte

Maß: MASS_1

Quelle		Typ III Quadratsumme	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Zeitpunkt1	Sphärizität angenommen	793.592	1	793.592	103.026	<.001	.715
	Greenhouse-Geisser	793.592	1.000	793.592	103.026	<.001	.715
	Huynh-Feldt (HF)	793.592	1.000	793.592	103.026	<.001	.715
	Untergrenze	793.592	1.000	793.592	103.026	<.001	.715
Zeitpunkt1 * Zugriff_Technologie	Sphärizität angenommen	6.686	1	6.686	.868	.357	.021
	Greenhouse-Geisser	6.686	1.000	6.686	.868	.357	.021
	Huynh-Feldt (HF)	6.686	1.000	6.686	.868	.357	.021
	Untergrenze	6.686	1.000	6.686	.868	.357	.021
Zeitpunkt1 * GooglevsChatGPT	Sphärizität angenommen	.334	1	.334	.043	.836	.001
	Greenhouse-Geisser	.334	1.000	.334	.043	.836	.001
	Huynh-Feldt (HF)	.334	1.000	.334	.043	.836	.001
	Untergrenze	.334	1.000	.334	.043	.836	.001
Zeitpunkt1 * Zugriff_Technologie * GooglevsChatGPT	Sphärizität angenommen	.599	1	.599	.078	.782	.002
	Greenhouse-Geisser	.599	1.000	.599	.078	.782	.002
	Huynh-Feldt (HF)	.599	1.000	.599	.078	.782	.002
	Untergrenze	.599	1.000	.599	.078	.782	.002
Fehler(Zeitpunkt1)	Sphärizität angenommen	315.817	41	7.703			
	Greenhouse-Geisser	315.817	41.000	7.703			
	Huynh-Feldt (HF)	315.817	41.000	7.703			
	Untergrenze	315.817	41.000	7.703			
Zeitpunkt2	Sphärizität angenommen	436.123	1	436.123	186.455	<.001	.820
	Greenhouse-Geisser	436.123	1.000	436.123	186.455	<.001	.820
	Huynh-Feldt (HF)	436.123	1.000	436.123	186.455	<.001	.820
	Untergrenze	436.123	1.000	436.123	186.455	<.001	.820
Zeitpunkt2 * Zugriff_Technologie	Sphärizität angenommen	.004	1	.004	.002	.968	.000
	Greenhouse-Geisser	.004	1.000	.004	.002	.968	.000
	Huynh-Feldt (HF)	.004	1.000	.004	.002	.968	.000

	Untergrenze	.004	1.000	.004	.002	.968	.000
Zeitpunkt2 * GooglevsChatGPT	Sphärizität angenommen	4.882	1	4.882	2.087	.156	.048
	Greenhouse-Geisser	4.882	1.000	4.882	2.087	.156	.048
	Huynh-Feldt (HF)	4.882	1.000	4.882	2.087	.156	.048
	Untergrenze	4.882	1.000	4.882	2.087	.156	.048
Zeitpunkt2 * Zugriff_Technologie * GooglevsChatGPT	Sphärizität angenommen	2.544	1	2.544	1.088	.303	.026
	Greenhouse-Geisser	2.544	1.000	2.544	1.088	.303	.026
	Huynh-Feldt (HF)	2.544	1.000	2.544	1.088	.303	.026
	Untergrenze	2.544	1.000	2.544	1.088	.303	.026
Fehler(Zeitpunkt2)	Sphärizität angenommen	95.900	41	2.339			
	Greenhouse-Geisser	95.900	41.000	2.339			
	Huynh-Feldt (HF)	95.900	41.000	2.339			
	Untergrenze	95.900	41.000	2.339			
Zeitpunkt1 * Zeitpunkt2	Sphärizität angenommen	57.167	1	57.167	23.275	<.001	.362
	Greenhouse-Geisser	57.167	1.000	57.167	23.275	<.001	.362
	Huynh-Feldt (HF)	57.167	1.000	57.167	23.275	<.001	.362
	Untergrenze	57.167	1.000	57.167	23.275	<.001	.362
Zeitpunkt1 * Zeitpunkt2 * Zugriff_Technologie	Sphärizität angenommen	2.284	1	2.284	.930	.340	.022
	Greenhouse-Geisser	2.284	1.000	2.284	.930	.340	.022
	Huynh-Feldt (HF)	2.284	1.000	2.284	.930	.340	.022
	Untergrenze	2.284	1.000	2.284	.930	.340	.022
Zeitpunkt1 * Zeitpunkt2 * GooglevsChatGPT	Sphärizität angenommen	1.323	1	1.323	.539	.467	.013
	Greenhouse-Geisser	1.323	1.000	1.323	.539	.467	.013
	Huynh-Feldt (HF)	1.323	1.000	1.323	.539	.467	.013
	Untergrenze	1.323	1.000	1.323	.539	.467	.013
Zeitpunkt1 * Zeitpunkt2 * Zugriff_Technologie * GooglevsChatGPT	Sphärizität angenommen	6.185	1	6.185	2.518	.120	.058
	Greenhouse-Geisser	6.185	1.000	6.185	2.518	.120	.058
	Huynh-Feldt (HF)	6.185	1.000	6.185	2.518	.120	.058
	Untergrenze	6.185	1.000	6.185	2.518	.120	.058
Fehler(Zeitpunkt1*Zeitpunkt2)	Sphärizität angenommen	100.700	41	2.456			
	Greenhouse-Geisser	100.700	41.000	2.456			
	Huynh-Feldt (HF)	100.700	41.000	2.456			

Untergrenze	100.700	41.000	2.456			

a. Unter Verwendung von Alpha = .05 berechnet

H: ANOVA Messwiederholung alle Daten

Tests der Innersubjekteffekte

Quelle		Typ III Quadratsumme	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Zeitpunkt	Sphärizität angenommen	793.592	1	793.592	103.026	<.001	.715
	Greenhouse-Geisser	793.592	1.000	793.592	103.026	<.001	.715
	Huynh-Feldt (HF)	793.592	1.000	793.592	103.026	<.001	.715
	Untergrenze	793.592	1.000	793.592	103.026	<.001	.715
Zeitpunkt * Gruppen	Sphärizität angenommen	7.294	3	2.431	.316	.814	.023
	Greenhouse-Geisser	7.294	3.000	2.431	.316	.814	.023
	Huynh-Feldt (HF)	7.294	3.000	2.431	.316	.814	.023
	Untergrenze	7.294	3.000	2.431	.316	.814	.023
Fehler(Zeitpunkt)	Sphärizität angenommen	315.817	41	7.703			
	Greenhouse-Geisser	315.817	41.000	7.703			
	Huynh-Feldt (HF)	315.817	41.000	7.703			
	Untergrenze	315.817	41.000	7.703			
Erinnerungstest	Sphärizität angenommen	436.123	1	436.123	186.455	<.001	.820
	Greenhouse-Geisser	436.123	1.000	436.123	186.455	<.001	.820
	Huynh-Feldt (HF)	436.123	1.000	436.123	186.455	<.001	.820
	Untergrenze	436.123	1.000	436.123	186.455	<.001	.820
Erinnerungstest * Gruppen	Sphärizität angenommen	7.261	3	2.420	1.035	.387	.070
	Greenhouse-Geisser	7.261	3.000	2.420	1.035	.387	.070
	Huynh-Feldt (HF)	7.261	3.000	2.420	1.035	.387	.070
	Untergrenze	7.261	3.000	2.420	1.035	.387	.070
Fehler(Erinnerungstest)	Sphärizität angenommen	95.900	41	2.339			
	Greenhouse-Geisser	95.900	41.000	2.339			
	Huynh-Feldt (HF)	95.900	41.000	2.339			
	Untergrenze	95.900	41.000	2.339			
Zeitpunkt * Erinnerungstest	Sphärizität angenommen	57.167	1	57.167	23.275	<.001	.362
	Greenhouse-Geisser	57.167	1.000	57.167	23.275	<.001	.362

	Huynh-Feldt (HF)	57.167	1.000	57.167	23.275	<.001	.362
	Untergrenze	57.167	1.000	57.167	23.275	<.001	.362
Zeitpunkt * Erinnerungstest * Gruppen	Sphärizität angenommen	9.111	3	3.037	1.237	.309	.083
	Greenhouse-Geisser	9.111	3.000	3.037	1.237	.309	.083
	Huynh-Feldt (HF)	9.111	3.000	3.037	1.237	.309	.083
	Untergrenze	9.111	3.000	3.037	1.237	.309	.083
Fehler(Zeitpunkt*Erinnerungstest)	Sphärizität angenommen	100.700	41	2.456			
	Greenhouse-Geisser	100.700	41.000	2.456			
	Huynh-Feldt (HF)	100.700	41.000	2.456			
	Untergrenze	100.700	41.000	2.456			

I: ANOVA mit Messwiederholung mit Vergleich beider Zeitpunkte

Tests der Innersubjekteffekte

Maß: MASS_1

Quelle		Typ III Quadratsumme	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Zeitpunkt1	Sphärizität angenommen	793.592	1	793.592	103.026	<.001	.715
	Greenhouse-Geisser	793.592	1.000	793.592	103.026	<.001	.715
	Huynh-Feldt (HF)	793.592	1.000	793.592	103.026	<.001	.715
	Untergrenze	793.592	1.000	793.592	103.026	<.001	.715
Zeitpunkt1 * Gruppen	Sphärizität angenommen	7.294	3	2.431	.316	.814	.023
	Greenhouse-Geisser	7.294	3.000	2.431	.316	.814	.023
	Huynh-Feldt (HF)	7.294	3.000	2.431	.316	.814	.023
	Untergrenze	7.294	3.000	2.431	.316	.814	.023
Fehler(Zeitpunkt1)	Sphärizität angenommen	315.817	41	7.703			
	Greenhouse-Geisser	315.817	41.000	7.703			
	Huynh-Feldt (HF)	315.817	41.000	7.703			
	Untergrenze	315.817	41.000	7.703			
Zeitpunkt2	Sphärizität angenommen	436.123	1	436.123	186.455	<.001	.820

	Greenhouse-Geisser	436.123	1.000	436.123	186.455	<.001	.820
	Huynh-Feldt (HF)	436.123	1.000	436.123	186.455	<.001	.820
	Untergrenze	436.123	1.000	436.123	186.455	<.001	.820
Zeitpunkt2 * Gruppen	Sphärizität angenommen	7.261	3	2.420	1.035	.387	.070
	Greenhouse-Geisser	7.261	3.000	2.420	1.035	.387	.070
	Huynh-Feldt (HF)	7.261	3.000	2.420	1.035	.387	.070
	Untergrenze	7.261	3.000	2.420	1.035	.387	.070
Fehler(Zeitpunkt2)	Sphärizität angenommen	95.900	41	2.339			
	Greenhouse-Geisser	95.900	41.000	2.339			
	Huynh-Feldt (HF)	95.900	41.000	2.339			
	Untergrenze	95.900	41.000	2.339			
Zeitpunkt1 * Zeitpunkt2	Sphärizität angenommen	57.167	1	57.167	23.275	<.001	.362
	Greenhouse-Geisser	57.167	1.000	57.167	23.275	<.001	.362
	Huynh-Feldt (HF)	57.167	1.000	57.167	23.275	<.001	.362
	Untergrenze	57.167	1.000	57.167	23.275	<.001	.362
Zeitpunkt1 * Zeitpunkt2 * Gruppen	Sphärizität angenommen	9.111	3	3.037	1.237	.309	.083
	Greenhouse-Geisser	9.111	3.000	3.037	1.237	.309	.083
	Huynh-Feldt (HF)	9.111	3.000	3.037	1.237	.309	.083
	Untergrenze	9.111	3.000	3.037	1.237	.309	.083
Fehler(Zeitpunkt1*Zeitpunkt2)	Sphärizität angenommen	100.700	41	2.456			
	Greenhouse-Geisser	100.700	41.000	2.456			
	Huynh-Feldt (HF)	100.700	41.000	2.456			
	Untergrenze	100.700	41.000	2.456			

J: Resultate ANOVA mit Messwiederholung schwierige Fragen

	Gruppen	Mittelwert	Standardabweichung
Messzeitpunkt1_SchwierigeFragen_freie_Reproduktion	Googlegruppe 1	5.5909	3.08073
	ChatGPT-Gruppe 3	5.3750	1.72053
	Googlegruppe 2	6.7500	3.38502
	ChatGPT-Gruppe 4	5.3750	2.21693
Messszeitpunkt2_SchwierigeFragen_freie_Reproduktion	Googlegruppe 1	3.8636	3.14715
	ChatGPT-Gruppe 3	3.2083	2.22034
	Googlegruppe 2	3.0500	2.65047
	ChatGPT-Gruppe 4	2.6250	2.93199
Messzeitpunkt1_SchwierigeFragen_unterstützte_Reproduktion	Googlegruppe 1	6.5455	3.04512
	ChatGPT-Gruppe 3	5.7917	2.08303
	Googlegruppe 2	6.6500	2.69825
	ChatGPT-Gruppe 4	5.9167	2.09798
Messzeitpunkt2_SchwierigeFragen_unterstützte_Reproduktion	Googlegruppe 1	5.0000	2.82843
	ChatGPT-Gruppe 3	4.5417	2.09391
	Googlegruppe 2	4.8500	2.98189
	ChatGPT-Gruppe 4	4.2917	3.31976

Zeitpunkt:

$F(1, 41) = 33{,}750$; $p < 0{,}001$

Erinnerungstest:

$F(1, 41) = 55{,}089$; $p < 0{,}001$

K: Merkmethode Häufigkeiten

- Auswendiglernen (Wiederholung im Kopf) - **40**
- Sortieren der Gegenstände in Kategorien (Obst, Gemüse, Kosmetika ...) - **2**
- Visualisierung der Gegenstände im Kopf - **16**
- Assoziationen zwischen Gegenständen (z. B. Löffel gehört auf den Teller) - **8**
- Visualisierung des Weges oder der Verteilung der Artikel im Supermarkt - **3**
- Erfinden einer Geschichte aus den Gegenständen auf der Liste - **2**
- andere Antwort

Andere Antworten:

Persönliche Assoziation zum Gegenstand, Auswendiglernen unbekannter Antworten, Angeknüpft an bestehendes Wissen

L: Ergebnisse Aussagen zum Speicherort

1. Wenn ich Informationen jederzeit online abrufen kann, fällt es mir schwer mich an diese zu erinnern. **(M= 2.93; SD: 1,116)**
2. Wenn ich mich intensiv mit einem Thema beschäftige, dann kann ich mich gut daran erinnern, obgleich Informationen abrufbar sind. **(M= 4.40; SD: 0,809)**
3. Wenn mich ein Thema weniger interessiert, dann habe ich Mühe mich daran zu erinnern. **(M= 3.93; SD:1,156)**
4. Indem ich mir gewisse Informationen nicht merke, schaffe ich mir mehr kognitive Ressourcen für andere, relevantere Informationen. **(M= 2.87; SD:1,057)**
5. Wenn ich weiss, wo ich Informationen finde, erinnere ich mich weniger gut an die Information selbst. **(M= 3.04; SD:1,167)**
6. Ich merke mir Informationen eher, wenn ich das Gefühl habe, diese nicht wiederfinden zu können. **(M= 2.96; SD:1,147)**
7. Ich erinnere mich eher an Informationen, die ich interessant finde. **(M= 4.24; SD:1,090)**
8. Wenn ein Thema unverständlich ist, dann habe ich Mühe mich daran zu erinnern. **(M= 3.78; SD:1,064)**
9. Ich merke mir Informationen, obgleich ich sie online abrufen kann. **(M= 3.53; SD:1,217)**

M: Kategoriensystem zur Einstellung

Skeptisch	Neutral	Positiv	Entwicklung
• Anfang 2023 war ich eher skeptisch und ehrlich gesagt auch ein bisschen genervt, dass ChatGPT gefühlt das einzige Thema war, das Leute interessiert hat. Heute benutzte ich ChatGPT selbst ab und zu, wenn auch noch nicht in einem grossen Ausmass. (K105B) • Ich war etwas skeptisch, ChatGPT in meinem beruflichen Umfeld zu verwenden, aber als ChatGPT schnell ausgereift war und ich viele mit ChatGPT implementierte Anwendungsfälle erlebte, war ich immer mehr davon überzeugt, ChatGPT zu verwenden und mein Team zu ermutigen, es auch in seinem persönlichen Umfeld zu verwenden als professionelle Arbeitsumgebungen. (K107S) • Grundsätzlich war und bleibe ich etwas skeptisch. Aktuell bin ich nur ein gelegentlicher User, bin jedoch über den jeweiligen Output positiv überrascht. Ich gehe davon aus, zukünftig ChatGPT als unterstützendes Instrument vermehrt einzusetzen. (K108B) • **Kaum, ich vertraue ChatGPT noch immer bedingt und recherchiere die Aussagen und Behauptungen von ChatGPT selbst nochmals nach, bevor ich sie für 'bare Münze' nehme.(K101C)**	• Es wurde weniger kritisch, aber mehr hinzu als Nutzung eines Tool und weniger als KI zur Benutzung (K105D) • Zu Beginn habe ich es öfters genutzt, um mich zu inspirieren lassen (Ideen, Texte, Wissen). Mittlerweile benutze ich es selten, v.a. weil es oft nicht funktioniert (entweder aus Security Gründen oder Überlastung). (K108P) • **Keine wahrnehmbare Veränderung in meiner Einstellung.** (K207W) • keine Veränderung (T102P) • Es ist ein mächtiges Tool, das Arbeit abnehmen kann. Ich selbst bin noch nicht so aufgesprungen. So Fragen beantworten, wie es im Experiment war, mache ich über Google. Dort habe ich die Möglichkeit verschiedene Quellen zu prüfen und zu sehen, wo die Info steht. (T112P) • Ich finde die ganze Thematik sehr interessant, wende es aber noch zuwenig an. Auch denke ich, dass man sehr vorsichtig sein muss	• Positiv in dem Sinn, dass der Nutzen für meine Arbeit ständig steigt und auch die Nutzungsdauer. Zudem sind meine Erfahrungen gewachsen. (T109R) • ich bin neugieriger geworden. Hätte ich Zeit, würde ich gerne mehr mit ChatGPT arbeiten und bin überzeugt, dass wir in Zukunft damit noch viel mehr Berührungspunkte haben werden. (T106C) • Ich sehe grossen Mehrwert beim Einsatz von ChatGPT. Jedoch ist es zwingend notwendig, dass man sich aktiv mit der Funktionsweise und den Rahmenbedingungen befasst, um die gewünschten Resultate unter Einhaltung von Regulatorien im geschäftlichen Einsatz zu erhalten. (T207U) • **Ich finde ChatGPT ein sehr hilfreiches Tool um innerhalb küzester Zeit Kommunikationsaufgaben zu erledigen oder um mich oberflächlich in ein Thema einlesen zu können. Grundsätzlich ist meine Einstellung sehr positiv, aber ich spüre auch eine gewisse Ineffizienz / Abhängigkeit wenn es mir für Kommunikationsaufgaben bspw. nicht zur Verfügung steht.** (T201R)	• Zu Beginn war ich etwas zurückhaltend und skeptisch. Durch die vermehrte Anwendung und die gewonnene Erkenntnis sehe ich ihn als nützliche Unterstützung in diversen Bereichen und wende ihn aktiver an. (T101C) • **Zuerst war ich sehr gehyped und habe alles mögliche damit ausprobiert. Vor allem im Studium nutze ich ChatGPT täglich. Danach trat eine Art Gewöhnungseffekt ein. Dies führte dazu, dass ich das Interesse zunehmend verloren habe. Ich nutze das NLP-Tool heute noch gerne für schnelles googeln, jedoch stehe ich dem Wahrheitsgehalt gewisser Aussagen/Antworten von ChatGPT inzwischen kritischer als am Anfang gegenüber - es scheint**

• Zuerst war ich sehr interessiert an der neuen Technologie und wollte es unbedingt ausprobieren. Nach meine ersten Versuchen habe ich meine ersten Schlüsse gezogen und fand es ein cooles Tool für bestimmte Analysen und Verständnisfragen - es hat mir Beispielsweise immer wieder geholfen in Statistikfragen (ich musste meine Bachelorarbeit mit R-Statistics analysieren). **Aber die Skepsis würde auch immer grösser, weil ChatGPT nur so gut ist, wie die Daten, die er gefüttert bekommt.** (K103G) • Ich bin kritischer geworden, was der Zuverlässigkeit der Antworten betrifft, nutze es aber mehr, um erste Entwürfe von Texten zu erstellen. (K207M) • Nicht gross - betrachte es immer noch kritisch & das Experiment hat wieder bestätigt, dass man mit den Antworten vorsichtig sein muss. (T209M1) • Mein Vertrauen hat stark abgenommen, da ich bemerkte, dass es auf Fragen, zu denen ich selbst gut Bescheid weiss, ganz klare Falschinformationen oder bei kleinen Nachfragen total andere Ergebnisse geliefert hat. ChatGPT ist ein bisschen wie Wikipedia: super, zum einen Ideenanstoss zu erhalten, aber ich verwende es nicht, wenn ich ein vertrauenswürdiges Ergebnis brauche. (T212A)	betreffend Fake News usw. (T207R) • Ich sehe die Vorteile eher als zuvor. Aber die Risiken (fehlende Spitzenqualität, kritisches Denken leidet, teilweise fehlende Effizienz bei Spezialthemen, fehlende Regelungen im Umgang mit Daten, etc.) die damit verbunden sind, sehe ich gleich wie im letzten Jahr. (T208S) • Ich habe Chat GPT bisher nicht wirklich gebraucht. Ich denke dies bietet eine gute Möglichkeit, Wissen zu sammeln, man muss aber die Antworten nicht einfach blind annehmen sondern auch selber nachdenken / interpretieren. (T209M)	• Es hat immer mehr an Bedeutung gewonnen. Viele Firmen und Privatpersonen nutzen dies im Alltag. (K203V) • Mittlerweile nutze ich Chat GPT regelmässig um Inspiration zu Themen zu erhalten. (K201B) • Anfang 2023 war ich eher skeptisch und ehrlich gesagt auch ein bisschen genervt, dass ChatGPT gefühlt das einzige Thema war, das Leute interessiert hat. Heute benutzte ich ChatGPT selbst ab und zu, wenn auch noch nicht in einem grossen Ausmass. (K105B) • Bin immer noch sehr begeistert. (K203S) • Ich habe viel über ChatGPT gelernt. Ich habe andere Anwendungsbeispiele gesehen, als nur write me a text about... (K108C) • Sehr positiv. Ich brauche ChatGPT täglich, jedoch eher für private Zwecke (Ausbildung etc.) (K108V) • Ich sehe für meine Arbeit aber auch im privaten unheimliches Potenzial (K101I) • Grundsätzlich war und bleibe ich etwas skeptisch. Aktuell bin ich nur ein gelegentlicher User, bin jedoch über den jeweiligen Output positiv überrascht. Ich gehe davon aus, zukünftig ChatGPT als unterstützendes Instrument vermehrt einzusetzen. (K108B) • Positiv in dem Sinn, dass der Nutzen für meine Arbeit ständig steigt und auch die Nutzungsdauer. Zudem sind	**als wäre die Maschine vorsichtiger/restriktiver mit seinen Aussagen geworden.** (T101F)

• Zu Beginn war ich etwas zurückhaltend und skeptisch. Durch die vermehrte Anwendung und die gewonnene Erkenntnis sehe ich ihn als nützliche Unterstützung in diversen Bereichen und wende ihn aktiver an. (T101C) • Ich war eigentlich überrascht, dass ChatGPT auch bei vermeintlich einfachen Fragen, mir eine falsche Antwort lieferte. Ich würde nun meine Fragen für ChatGPT unterschiedlich formulieren und die Antworten vergleichen. (T101M) • Ich hinterfrage viel mehr die Richtigkeit der Antworten von ChatGPT. (T101E) • Nicht gross, seit dem Versuch bin ich jetzt aber skeptischer, was die Richtigkeit der Antworten von ChatGPT angeht. (T209M1) • Ich sehe immer mehr Anwendungsmöglichkeiten. Die Antworten von ChatGPT versuche ich noch intensiver weiter zu verifizieren. (T205C)		meine Erfahrungen gewachsen. (T109R) • ich bin neugieriger geworden. Hätte ich Zeit, würde ich gerne mehr mit ChatGPT arbeiten und bin überzeugt, dass wir in Zukunft damit noch viel mehr Berührungspunkte haben werden. (T106C) • Anfänglich war ich begeistert, Neugierig. Heute bin ich das immer noch und nutze es. Vom Spielen, Testen ins Nutzen. (T105Y) • Ich habe Chat GPT erst im letzten Jahr kennengelernt. D.h. ich nutze es jetzt und finde es gut, weil es einem helfen kann. (T110E) • Hat sich in eine positive Richtung entwickelt. Es gibt viele spannende Anwendungsfelder aus meiner Sicht. (T112S) • Ich habe die Potenziale von ChatGPT etwas mehr kennengelernt, was mich grundsätzlich positiver einstellt als zuvor. (T209P) • Ich kenne ChatGPT erst seit ein paar Monaten und habe es bisher nur selten gebraucht. Ich empfinde es als sehr hilfreich zur Erstellung von längeren Texten wie Mails, SoMe-Posts, Artikel, etc. Meine Einstellung war grundsätzlich immer positiv. (T204J) • From good to great! I am a total fan! (T204B)	

N: Kategoriensystem zur Veränderung

Unverändert/ Neutral	Verbessert/ Positiv	Verschlechtert/ Negativ
1. Ich muss mir sowieso alles aufschreiben, also ist sie nicht schlechter geworden. (K10B) 2. Nicht viel, **(ich nutze es nur als Sprachverbesserungstool)** (K105D) **3. kaum, weil ich ChatGPT nur sehr vereinzelt und selten nutze** (K103G) 4. Ich hätte gesagt unverändert (K108P) 5. Unverändert, Google und Wikipedia hatten einen stärkeren Einfluss. (K207M) 6. nutze es zu wenig als dass es Einfluss haben könnte (K207K) 7. noch immer gleich schlecht… (T106C) 8. sehe keine Veränderung, nutze chatgpt nicht so oft (T102P) 9. Nicht negativ ggü. früher. (T110E) 10. keine große Veränderung (T101E) 11. sehe keine Veränderung (T207R) 12. Wurde nicht beeinflusst (T209M) 13. keine Änderung, es war immer schlecht (T204B) 14. Ich habe das Gefühl gar nicht. (T207U) 15. Unverändert (T205C) 16. Keine Veränderung (K209M) 17. Gleich (K207W) 18. nicht viel (T101F) 19. mässig (K101I) 20. Gar nicht. (K207M1) 21. Gar nicht (K201B) 22. gar nicht (T109R) 23. Gar nicht T201R)	24. Mein Gedächtnis profitierte von der Erinnerung an das Langzeitgedächtnis (K107S) 25. Verständnis von Dingen wurde besser, genaue/spezifische Antworten gleichbleibend mit Tendenz zu schlechter (K101C) 26. Die Infos sind leichter zugänglich. Muss mir nicht mehr so viel merken (K203V) 27. Ich habe nicht das Gefühl. Dass die Sachen, die ich in chat gpt gesucht habe, ich mich erinnern mag, weil ich sie in Chat GPT gesucht habe. Ich lerne dadurch. (T105Y) 28. nicht mehr wissen, sondern auswendig lernen (T101M) **29. Gar nicht, ich lerne visuell durch meine eingegebenen Antworten** (T112P) 30. jene Sachen, welche ich mit ChatGPT gesucht habe, konnte ich mir relativ gut merken, jedoch noch kein Bestandteil meines Alltages (T112S) 31. Schwierig zu sagen, vor allem bei Differenzen zu meinen eigenen Antworten konnte ich mir die Resultate merken (T209M1) **32.** Kommt auf das Thema drauf an, bestimmte Themen sind wichtiger, die kann ich mir gut merken, andere weniger (T209P)	33. Marginal (K108B) 34. Langfristiger (K108V) 35. Ich nutze es nicht so oft aber ich merke, dass ich mich an den Inhalt nicht langfristig erinnern kann. (K112P) 36. ich bin überrascht, dass ich jetzt noch soviel weiß. Kurzfristig konnte ich mich nicht mehr daran erinnern. -> vergesslicher geworden seit ChatGPT (T109A) 37. Dieser Test zeigt, dass sie sich verschlechtert hat. Hätte ich nicht gedacht. (T101C) 38. nicht mehr wissen, sondern auswendig lernen (T101M) 39. **keine Selbsterarbeitung mehr? => Informationen bleiben weniger lange im Gedächtnis (ich merke mir Dinge vor allem, durch den Weg)** (T212A) **40. Ich arbeite zwar schneller, versuche mir aber weniger die Informationen zu merken da sie immer innerhalb kürzester Zeit zur Verfügung stehen.** (T201R) 41. Wohl eher verschlechtert (T204J) 42. Verschlechtert (T211X)

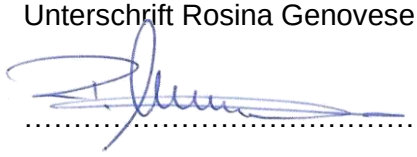
Anderes:

43. Kann ich nicht beurteilen (K205K)
44. ich habe ChatGPT selber kaum gebraucht. Aus diesem Grund kann ich diese Frage nicht wirklich beantworten (K108C)
45. Nicht verfolgt (K205B)

O: Eidesstaatliche Erklärung

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig, ohne Mithilfe Dritter und nur unter Benutzung der angegebenen Quellen verfasst zu haben. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit war in gleicher oder ähnlicher Form noch nicht Bestandteil einer anderen Studien- oder Prüfungsleistung.

Ort, Datum	Unterschrift Rosina Genovese
Olten, 20.06.2024	
……	……....