

From Search to Dialogue: An Experimental Comparison of User Experience, Satisfaction and Success with ChatGPT and Google

Anna Schätzle
 Technical University of Darmstadt
schaetzle@is.tu-darmstadt.de

Danilo Walenta
 Technical University of Darmstadt
danilo.walenta@tu-darmstadt.de
 Peter Buxmann
 Technical University of Darmstadt
peter.buxmann@tu-darmstadt.de

Maren Mehler
 Technical University of Darmstadt
mehler@is.tu-darmstadt.de

Abstract

Generative Artificial Intelligence (GenAI) is increasingly shaping informal learning in everyday life, with tools like ChatGPT becoming part of many individuals' daily routines. While prior research has focused primarily on student learning, the role of GenAI in adult everyday information-seeking remains underexplored. This study examines how GenAI influences user experience, learning satisfaction, and learning outcomes in informal contexts. In a randomized online experiment (N = 120), participants completed learning tasks using both ChatGPT and Google. Results show that ChatGPT significantly improves user experience and learning satisfaction compared to traditional search engines. Additionally, self-efficacy positively impacts learning satisfaction, and both user experience and satisfaction are associated with fewer unknown responses—leading to higher quiz scores and better learning outcomes. These findings underscore the value of conversational GenAI in enhancing informal learning, highlighting its potential to support more effective, satisfying, and self-directed digital knowledge acquisition.

Keywords: Generative Artificial Intelligence, User Experience, Satisfaction, Comparison, Search Engines.

1. Introduction

People's efforts to stay informed and acquire new knowledge in their daily lives are being transformed by emerging technologies like generative artificial intelligence (GenAI), which unlock interactive and personalized learning experiences (Alavi et al., 2024; Brynjolfsson et al., 2025; Feuerriegel et al., 2024). GenAI, designed to simulate human intelligence, can produce novel content from datasets (Noy & Zhang, 2023), generate information in previously unseen formats, adapt explanations to different knowledge levels, and develop connections across multiple concepts (Alavi et al., 2024; Brynjolfsson et al., 2025).

Informal learning in everyday life—such as understanding unfamiliar topics or solving daily problems—has become increasingly important. It enables individuals to realize their potential, seize new opportunities, and navigate everyday challenges more effectively (Drucker, 1999; Law, 2024). While people often rely on prior knowledge, they frequently need specific, just-in-time information—for instance, to explore unfamiliar subjects, plan trips, or grasp complex ideas more intuitively (Banh & Strobel, 2023; Feuerriegel et al., 2024).

Traditionally, search engines like Google provided quick access to such information and shaped how people searched and learned online (Alavi et al., 2024; Brynjolfsson et al., 2025). GenAI, however, offers a more conversational, contextual, and personalized form of informal learning, enabled through human-like interaction (Brynjolfsson et al., 2025; Feuerriegel et al., 2024). Unlike classic search engines, GenAI responds in context, adapts to prior interactions, and summarizes complex content in accessible language (Alavi et al., 2024). These features make GenAI especially suitable for informal, everyday learning scenarios, where clarity, speed, and ease of understanding are crucial (Alavi et al., 2024; Feuerriegel et al., 2024). Nonetheless, concerns remain—especially regarding misinformation, lack of source transparency, and the difficulty of verifying generated content (Barrett et al., 2023).

Although the potential of GenAI for learning is widely discussed, most research has focused on formal education, while its role in informal, self-directed learning—especially among adults—remains underexplored (Banh & Strobel, 2023; Law, 2024). This gap is relevant as tools like ChatGPT, Claude.ai, or Google Gemini increasingly support everyday information-seeking beyond professional contexts (Banh & Strobel, 2023; Bockshecker et al., 2022; Brynjolfsson et al., 2025; Feuerriegel et al., 2024). Conceptually, we distinguish between search engines and conversational GenAI, using Google and ChatGPT as exemplars for our study.

Against this background, the main objective of this paper is to examine how GenAI affects user experience, learning satisfaction, and learning success in informal digital learning. We frame our claims at the level of these technology families while grounding the empirical comparison in Google (search engine) and ChatGPT (conversational GenAI). Prior research suggests that GenAI tools can enhance user experience through natural language interaction, higher usability, and a sense of human-like engagement (Kim et al., 2025; Seaborn & Fels, 2025). Additionally, hedonic qualities such as enjoyment and intuitive operation are known to increase user engagement and satisfaction in digital learning environments (Eberts, 1994; Stone et al., 2005). To better understand the potential of GenAI for everyday, self-directed learning, it is important to consider psychological and experiential factors such as user experience, learning satisfaction, and self-efficacy (Alter & Oppenheimer, 2009; Bandura, 1977; Deci et al., 1996; Zimmer & Niemimaa, 2019). These constructs shape how users perceive, interact with, and benefit from GenAI-based learning—particularly in informal, digital contexts (Craik, 1983; Dalgıç et al., 2024; Feuerriegel et al., 2024; Zheng & Xiao, 2024; Zhonggen et al., 2019). This led us to our research questions (RQs): *(RQ1) How does the use of GenAI affect user experience and satisfaction?* And *(RQ2) How does GenAI use influence learning success, in terms of the number of (un)known answers?*

We examined these questions in the context of informal everyday learning through an online experiment with 120 participants, who completed tasks using both ChatGPT and Google. Learning success was measured with a factual quiz directly after the tasks. Results show that ChatGPT significantly improved user experience, satisfaction, and learning outcomes. Higher self-efficacy further increased satisfaction, and participants reported a stronger sense of topic familiarity. These findings highlight the potential of conversational GenAI to support informal digital learning.

2. Theoretical background

2.1. GenAI and everyday information gathering

Since the emergence of tools like ChatGPT, GenAI has transformed both work and private life (Alavi et al., 2024; Feuerriegel et al., 2024). We define GenAI as “computation technologies that are capable of generating seemingly new, meaningful content such as text, images, or audio from training data” (Feuerriegel et al., 2024, p.111). Its aim is to produce content matching or surpassing human-created work (Banh &

Strobel, 2023). To this end, models are trained on vast datasets to recognize patterns and correlations; outputs are probabilistic and depend on factors such as prior conversation and context, varying across models and prompts (Banh & Strobel, 2023; Brynjolfsson et al., 2025; Feuerriegel et al., 2024).

GenAI is designed to simplify and accelerate processes (Alavi et al., 2024; Banh & Strobel, 2023). Tools like ChatGPT are used for everyday tasks—from drafting messages to obtaining trip information (Banh & Strobel, 2023; Brynjolfsson et al., 2025; Feuerriegel et al., 2024). Because GenAI accepts natural-language prompts and responds in a human-like manner, the user experience differs from that of traditional search engines (Kim et al., 2025; Kohnke & Moorhouse, 2025; Yang et al., 2025): rather than listing links, it acts as a conversational partner presenting information in an accessible way (Banh & Strobel, 2023; Brynjolfsson et al., 2025; Feuerriegel et al., 2024). Prior work has examined GenAI for learning (e.g., “learning buddies” and tailored materials) and teacher support (worksheets, exercises), enabling customizable learning when used purposefully (Baidoo-Anu & Ansah, 2023; Michel-Villarreal et al., 2023; Mittal et al., 2024; Pesovski et al., 2024; Qadir, 2023).

In this context, it is important to distinguish different forms of learning. In this study, we conceptualize “informal learning” as learning that occurs outside institutional structures, is initiated and largely directed by the learner, and is embedded in everyday situations without formal assessment or certification (Folkestad, 2006; Johnson & Majewska, 2022). It contrasts with formal learning (institutionalized, curriculum-bound, credentialed) and non-formal learning (organized activities outside formal systems, often with explicit goals) (Johnson & Majewska, 2022). Since GenAI can be integrated into everyday contexts and accessed flexibly through natural language, it offers potential to support informal learning processes.

2.2. User experience

User experience (UX) is commonly defined as the experience that arises through interaction with, and is shaped by, technology (Hassenzahl, 2013, p. 1). In this study, UX is assessed following Schrepp et al. (2017) using the short version of the User Experience Questionnaire (UEQ-S). This instrument captures two meta-dimensions: pragmatic quality (e.g., efficiency, perspicuity, dependability) and hedonic quality (e.g., stimulation, originality). UX is widely regarded as a central success factor of technical innovations, as it evaluates how individuals experience and accept a system (Eberts, 1994; Hassenzahl, 2013; Stone et al.,

2005). Accordingly, products must not only be functional but also support ease of learning, efficiency, controllability, and enjoyment (Hinderks et al., 2020; Schrepp et al., 2017). Different dimensions of UX can be considered, such as usage frequency, subjective descriptions of use, or behavioral indicators (Eberts, 1994; Rohrer, 2014). Measurement can be quantitative or qualitative, attitudinal or behavioral (Hinderks et al., 2020; Rohrer, 2014). In this paper, we focus on attitudinal aspects—feelings, impressions, and attitudes—captured through the validated UEQ-S (Schrepp et al., 2017). We chose this focus because alternative approaches, such as behavioral or performance-based measures, are often more complex to capture consistently in experimental online settings.

Existing research highlights that human-like conversation with immediate, contextualized responses, presented directly within a chat interface, differs fundamentally from the experience of entering keywords into a traditional search engine and scanning a list of hyperlinks (Kim et al., 2025). This distinction has a strong impact on how users perceive and engage with the information-retrieval process (Kim et al., 2025; Kohnke & Moorhouse, 2025). Against this backdrop, we hypothesize that the use of GenAI, in comparison to traditional search engines, has a positive effect on user experience during learning tasks on everyday topics. In line with the ongoing shift from conventional information retrieval toward more interactive and conversational systems, GenAI can be considered a transformative technology in digital learning (Kohnke et al., 2023; Kohnke & Moorhouse, 2025). We therefore state a causal, tool-specific hypothesis *H1: Using ChatGPT as a learning tool positively affects user experience.*

2.3. Learning satisfaction

In e-learning, learning satisfaction reflects the extent to which learners feel motivated and fulfilled during and after the experience (Almaiah et al., 2022; Deci et al., 1996; Sun et al., 2008). It is linked to meeting psychological needs such as autonomy and competence (Deci et al., 1996; Zhonggen et al., 2019), as well as motivation, expectations, and confidence in success (Chiu, 2024). Emerging digital technologies can support these factors and complement approaches like the flipped classroom (Akçayır & Akçayır, 2018; Roman et al., 2025).

Prior studies have shown that interaction is one of the strongest drivers of learning satisfaction (Nortvig et al., 2018). In e-learning settings, this interaction can take place both between the learner and the tool, as well as between learners, peers, and instructors (Deci et al., 1996; Nortvig et al., 2018; Zhonggen et al., 2019). For

example, research during the COVID-19 pandemic demonstrated that higher levels of interaction increased both learner success and satisfaction (Pangarso & Setyorini, 2023). Autonomy in learning and the individual's ability to determine learning style and pace are further factors that positively affect satisfaction (Borah et al., 2024; Deci et al., 1996; Zhonggen et al., 2019). Considering these aspects, ChatGPT differs significantly from traditional search engines like Google. By enabling human-like communication and presenting results in a conversational format, ChatGPT provides a more interactive and autonomous learning experience, which is likely to foster higher satisfaction (Brynjolfsson et al., 2025; Feuerriegel et al., 2024). Based on this reasoning, we anticipated a difference between ChatGPT and Google in terms of learner satisfaction and hypothesized a positive relationship between the use of GenAI and the reported satisfaction after completing the learning session. Specifically, we state *H2: Using ChatGPT as a learning tool positively affects learning satisfaction.*

2.4. Self-efficacy

Self-efficacy refers to the initiation and maintenance of goal-directed behavior in the face of challenges (Bandura, 1977; Getenet et al., 2024; McCoy, 2010). It builds on existing skills, motivation, and the belief that a task can be solved (Bandura, 1977; Getenet et al., 2024). Described as “the most central and pervasive mechanism of personal agency” (Bandura, 1977, p. 2), it is a core element of social cognitive theory, where human motivation depends on expected outcomes (Bandura, 1977; Zimmerman, 2000). Unlike personal qualities, self-efficacy emphasizes performance capabilities (Zimmerman, 2000).

Research has shown that self-efficacy in students in online learning models positively influences the students' engagement socially, cognitively, behaviorally, and collaboratively. Additionally, it can increase overall course satisfaction (Getenet et al., 2024; Zheng & Xiao, 2024). Self-efficacy is influenced by digital literacy and attitude (Getenet et al., 2024), and learners' self-regulatory behavior has also been shown to influence self-efficacy in the digital space (Zheng & Xiao, 2024). In the context of GenAI, relevant competencies indirectly enhance critical-thinking awareness by boosting general self-efficacy and learning motivation (Jia & Tu, 2024). In addition, self-efficacy has been found to be a key mechanism influencing cognitive learning processes with GenAI (Jia & Tu, 2024). Self-efficacy also plays an important role in the context of learning, as it influences how confidently individuals approach new content and their belief in their ability to master it (Getenet et al., 2024;

Zimmerman, 2000). Finally, self-efficacy beliefs are significantly positively related to the scores achieved in knowledge tests, such as exams (Jackson, 2002). These findings underline the central role of self-efficacy in the digital learning context, especially in combination with GenAI-supported learning environments. Self-efficacy acts not only as a mediator, but also as a driving force for successful, motivated, and satisfying learning (Bandura, 1977; Zimmerman, 2000). Research has further indicated that self-efficacy is closely linked to learners' perceived competence and autonomy—two psychological needs that are key predictors of learning satisfaction (Chiu, 2024; Deci et al., 1996). A strong belief in one's own capabilities not only fosters motivation but also enhances the feeling of accomplishment and fulfillment during the learning process, which in turn increases satisfaction. Therefore, we anticipated that self-efficacy would positively affect learning satisfaction in our experiment. We hypothesized the following: *H3: Self-efficacy positively affects learning satisfaction.*

2.5. Learning outcome

Learning outcomes refer to the cognitive or affective acquisition of knowledge or skills, usually assessed at the end of the learning process (Caspersen et al., 2017; Dalgıç et al., 2024; Hussey & Smith, 2008). They can be measured in different ways—such as quizzes, oral exams, or practical tests—depending on the content (Wu & Yu, 2024). In this study, which focuses on theoretical knowledge, we used a factual quiz to assess learning success. Outcomes were expressed as the percentage of correct answers.

Prior research highlights the role of engagement and interaction in e-learning (Wagino et al., 2024). With GenAI chatbots, short interactions often improve learning more than longer ones (Wu & Yu, 2024), as their novelty and interactivity can boost individual impact (Boubker, 2024; Wu & Yu, 2024). ChatGPT thus supports interactive learning, mediated by digital literacy (Dalgıç et al., 2024). Moreover, self-determination in learning pace and content fosters outcomes (Deci et al., 1996; Dalgıç et al., 2024). Learning outcomes are also linked to critical thinking, knowledge gains, and feedback (Chiu, 2024; Dalgıç et al., 2024), and can be sustained over time through GenAI support (Edwards-Fapohunda & Adediji, 2024).

Therefore, the next three hypotheses pertain to the effect of learning satisfaction and the user experience on the learning outcome. The existing literature shows that learners' engagement can increase if learning satisfaction can be increased through greater motivation, enjoyment, and clarity of purpose (Chiu, 2024; Nortvig et al., 2018). Moreover, a positive user

experience leads to more interaction with and use of the tool (Eberts, 1994; Hinderks et al., 2020). It has also been shown that user-friendliness, fun, and satisfaction during learning reduce frustration (Hassenzahl, 2013; Kim et al., 2025), which is beneficial to the learning process. We hypothesize that the number of unknown answers decreases when learning satisfaction and user experience increase. Learning success is measured by the frequency with which users did not select the 'I don't know the answer' option in the post-learning quiz. Each question included three possible answers and a fourth option allowing participants to indicate that they did not know the correct answer. The unknown answers are those for which the user actively chooses the multiple-choice option "I do not know the answer" and therefore is aware of their knowledge gaps. This resulted in *H4: User experience has a negative effect on unknown answers in the learning outcome measurement.* And *H5: Learning satisfaction has a negative effect on unknown answers in the learning outcome measurement.*

We assume that the number of "unknown" responses reflects participants' actual knowledge level: the more often "I do not know" is selected, the greater the unfamiliarity with the content. Since participants were instructed not to guess, a higher number of unknowns indicates lower knowledge, which should translate into fewer correct answers and a lower overall score. Therefore, we hypothesized: *H6: The number of unknown answers in the learning outcome measurement has a negative effect on the number of correctly answered quiz questions.*

Our measure of "learning success" (unknown vs. correct answers) captures both knowledge and metacognitive awareness. Fewer "unknowns" may reflect confidence as much as competence. We therefore interpret it as perceived knowledge acquisition and note that future studies should add more robust performance tests. Our hypotheses resulted in the following research model (Figure 1). In this case, GenAI assistance specifically refers to the use of ChatGPT.

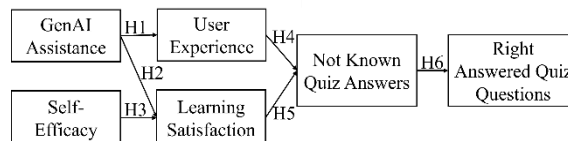


Figure 1. Research model.

3. Research methodology

To answer our research questions, we conducted an online experiment, a common method in information systems research on human–AI collaboration and adult online learning (Fügener et al., 2022; Getenet et al., 2024). We compared whether user experience and learning satisfaction increased when GenAI was used

versus a traditional search engine, and examined effects on unknown and correctly answered quiz questions (learning success). Participants learned about two unfamiliar topics: the Omo¹ River and the Okapi². These topics were chosen because they are relatively unknown, ensuring a genuine learning experience. As both are located in Africa, we excluded participants from that region to minimize prior familiarity.

3.1. Research design and measurement

Our design reflects these features of informal learning by requiring self-directed tasks on unfamiliar topics outside institutional settings. The study employed a 2×2 within-subjects design with two factors: Information artifact (Google vs. ChatGPT) × Learning topic (Omo River vs. Okapi), with counterbalanced presentation order across participants. Each participant had to learn about both topics, randomly assigned to one of the two learning technologies Google or ChatGPT. One group used ChatGPT first, and the other used Google first. To reduce technical challenges as much as possible, we built a website embedded with the tools so that participants could use Google directly within the survey or begin a chat with ChatGPT. They did not have to open a separate browser or the ChatGPT app. The questionnaire was implemented on our website, allowing us to include the use of Google and ChatGPT seamlessly with the provided questions concerning the items of the measured constructs.

The study was conducted as follows: After agreeing to a privacy policy and reading the conditions for storing the given answers and data, the participants began the experiment. First, they answered questions about their ability to apply AI to determine AI literacy as control variable and provide information about their demographics like age or gender. Afterward, they were randomly assigned to the group that would learn with Google first or the group that would learn with ChatGPT first. Whether they should start with the Omo River or Okapi as the first learning topic was also randomly assigned. After being instructed that they should spend at least 5 minutes on the respective topic, which has been found to be the average session length of users browsing different topics (Gadiraju et al., 2018), they were directed to Google or ChatGPT embedded within the experiment on our website. Both learning sessions imposed the same minimum duration of five minutes

(first session: ≥5 min; second session: ≥5 min) to standardize exposure time across conditions.

After the minimum time limit, participants could continue learning about the Omo River or Okapi for as long as they wanted. They were then directed to a quiz in which they were asked to answer 15 questions about the learning topic. This quiz was used to measure learning success. The questions covered different aspects of the Omo River or okapi, for example, the geographical location of the Omo River or the eating habits of the Okapi. All questions had multiple choice items including “I do not know the answer.” They then answered questions addressing the constructs learning satisfaction and user experience. Subsequently, those who initially worked with Google were directed to use ChatGPT, and vice versa. They were then presented with the learning topic (Omo River or okapi) that they had not yet covered. It is important to note that the interaction modes differed systematically: using ChatGPT required participants to formulate prompts and engage in conversational exchanges, whereas using Google involved browsing, evaluating multiple links, and synthesizing content. Thus, our comparison does not assume identical cognitive processes, but rather contrasts two distinct information-access modalities. After the second learning session, questions regarding the newly assigned learning topic were asked to measure learning success. Additionally, the questions related to constructs of learning behavior were repeated to assess differences between the learning methods used. When the participants used Google, we recorded search queries and the links that were opened. When they used ChatGPT, we saved the input prompts as well as the response messages generated by ChatGPT.

We measured the user experience with the items by Schrepp et al. (2017) (eight items on a scale, e.g., “obstructive – supportive”). These items are widely used in the existing literature and we directly adopted them without modification, as they represent a validated and established measure of user experience. Learning satisfaction was measured with the item by Almaiah et al. (2022) (four items on 7-point Likert scale, e.g., “The system contributes effectively to my acquisition of new skills in learning”). This scale has been established in the literature and was directly adopted for our study to ensure reliability and comparability with prior research. The construct of self-efficacy was measured with the items by McCoy (2010) (five items on 7-point Likert scale, e.g., “I can solve most problems if I invest the necessary effort”). To measure the number of unknown questions for each participants, we calculated the

¹ The Omo River is a roughly 760 km long river in Ethiopia that flows into Lake Turkana, serving as a vital water source for indigenous groups. It is famous for its archaeological sites, which have yielded important fossils of early human ancestors.

² The okapi is a rare, forest-dwelling mammal native to the Democratic Republic of Congo, and it is closely related to the giraffe. It has a reddish-brown body with zebra-like stripes on its legs, helping it blend into its dense rainforest habitat.

number questions marked as “I don’t know” from the quizzes as an integer value (Omo and Okapi) that contained 15 questions per quiz. For the score, as a measurement of the learning success, we used the number of correctly answered questions from the quizzes as an integer value. We adapted the items of the constructs slightly to our context. For example, we replaced “technology” with “ChatGPT.” Moreover, we asked the participants about their demographic data, and as a control variable, their ability to apply AI (six items on 7-point Likert scale, e.g., “I can communicate gainfully with AI in everyday life”). During the experiment, we ensured validity by utilizing established constructs and pre-testing the survey. We maintained reliability by employing multi-item constructs whenever feasible. Before the main study, we pretested the experiment with three IS researchers and two student researchers to identify inconsistencies. Based on their feedback, we refined wording and task descriptions.

3.2. Data collection and sample selection

Participants were recruited via Prolific from residents of non-African countries (eligibility restricted to non-African locations to reduce prior topic familiarity). We contacted our participants through Prolific in 2025. Participants were eligible if they were adults, fluent in English, and met Prolific quality criteria; invitations were distributed using Prolific’s random allocation within these eligibility filters. We paid our 122 participants the platform-standard compensation of \$9 per hour. Overall, the experiment was expected to take 25–30 minutes, including a minimum of 5 minutes for using Google and 5 minutes for using ChatGPT (Gadiraju et al., 2018).

Age	Gender		Activity		
18–29	51	Female	45	Agriculture	6
30–49	50	Male	75	Service	35
50–70	17	Other	0	Marketing	5
>70	2			Communications	1
				Human Resources	4
				Finance	5
Management level	Firm size (employees)				
Higher	27	1-9	36	Sales & Distribution	4
Lower	33	9-49	23	Research	5
Non	60	50–249	29	Production	7
		> 249	32	Other	48

Table 1. Demographics of survey participants.

The participants required, on average, 28 minutes. Overall, 122 people participated in the experiment. Two participants failed the attention check with the statement “please select strongly disagree” in the survey, so we excluded their answers from the analyses. We did not detect any straightliners. This resulted in a final sample

size of 120 participants, of whom 60 were assigned to the group first using Google and 60 to the group first using ChatGPT. The participants’ demographic data are shown in Table 1. The participants reported an average prior GenAI experience of 5.80 meaning they have GenAI experience (measured with six items on a 7-point Likert scale where 7 = highest experience level, Cronbach’s alpha = .958, SD = 0.078).

4. Results

In the following, we report the data analysis regarding reliability and validity of our constructs. Internal consistency was assessed using Cronbach’s alpha, composite reliability (CR), and average variance extracted (AVE) (Fornell & Larcker, 1981). As shown in Table 2, all constructs exceeded the recommended thresholds (Cronbach’s alpha > 0.7, CR > 0.7, AVE > 0.5; Hair et al., 2019), indicating satisfactory reliability and convergent validity. To examine discriminant validity, we applied the Fornell–Larcker criterion, which was met since the square root of each construct’s AVE was greater than its correlations with other constructs. We also checked the heterotrait–monotrait (HTMT) ratio of correlations (Henseler et al., 2015). All HTMT values remained below the critical threshold of 0.90, suggesting no issues with discriminant validity. Furthermore, all factor loadings exceeded the 0.7 benchmark, confirming that items strongly represented their intended constructs.

Construct	Factor loadings	Cronbach’s alpha	CR	AVE
User Experience	0.873–0.901	0.863	0.917	0.785
Learning Satisfaction	0.798–0.921	0.907	0.935	0.784
Self-Efficacy	0.760–0.859	0.827	0.882	0.652

Table 2. Assessment of reliability and validity.

Some variables showed particularly strong correlations, such as “User Experience” with “Learning Satisfaction” (0.401), see Table 3. Moreover, the correlation of “Unknown Answers” was negatively correlated with “User Experience,” “Learning Satisfaction,” and “Self-Efficacy” as well as “Correctly Answered Quiz Questions.” All correlations were significant except the correlation between “Self-Efficacy” and “GenAI Assistance” and “Self-Efficacy” and “Correctly Answered Quiz Questions,” as well as “Unknown Answers” and “GenAI Assistance”.

	UE	LS	SE	GenAI	NA	CA
UE	1					
LS	0.401**	1				
SE	0.154*	0.196**	1			

GenAI	0.142*	0.190**	0.050	1		
NA	-0.295**	-0.284**	-0.186**	0.039	1	
CA	0.195**	0.167**	0.048	-0.049	-0.608**	1

Table 3. Correlation matrix.

We conducted paired t-tests using SPSS to compare user experience and learning satisfaction between ChatGPT and Google. Levene’s test confirmed homogeneity of variances ($p > .05$), and normal distribution was assumed based on the sample size ($N = 60$ per group) (Stone, 2010), justifying the use of paired t-tests. Results showed significantly higher user experience ratings for ChatGPT ($M = 5.10, SD = 1.29$) than for Google ($M = 4.47, SD = 1.35$), $t = 3.50, p < .001$. Similarly, learning satisfaction was higher with ChatGPT ($M = 5.88, SD = 0.95$) than with Google ($M = 5.47, SD = 1.13$), $t = 3.43, p < .001$. Both constructs were measured on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), with higher scores indicating more positive experiences. These findings indicate that ChatGPT significantly enhances usability and satisfaction compared to traditional search. As a second step, we tested our hypotheses using PLS-SEM and a bootstrapping procedure with 5,000 subsamples in SmartPLS 4 (Davison & Hinkley, 1997). We evaluated the structural model via R^2 values and the significance of path coefficients (Shmueli & Koppius, 2010). All hypotheses (H1–H6) were supported, as shown in Table 4 and Figure 2.

Hypotheses	β	SD	t	p
H1	0.297	0.133	2.187	0.028
H2	0.399	0.120	3.308	0.001
H3	0.189	0.064	2.711	0.007
H4	-0.217	0.077	2.810	0.005
H5	-0.218	0.076	2.861	0.004
H6	-0.600	0.041	14.594	0.000

Table 4. Assessment of the hypotheses.

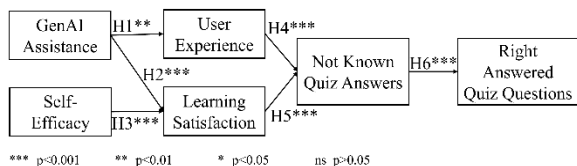


Figure 2. Research model results.

5. Discussion and conclusion

GenAI enhances informal learning by shifting knowledge acquisition from traditional search tools to human-like interaction (Alavi et al., 2024; Feuerriegel et al., 2024). Our study experimentally compares ChatGPT and Google, showing effects on user experience, satisfaction, and learning outcomes, thereby

offering new insights into GenAI’s role in self-directed learning. First, we found that GenAI assistance via ChatGPT significantly improved user experience in our learning experiment. This aligns with prior research showing that enjoyment enhances user experience in GenAI (Kim et al., 2025), potentially leading to better learning outcomes. Human-like interaction is a key driver of user experience and likely explains why ChatGPT use yielded higher scores. While user experience is subjective and challenging to measure (Hinderks et al., 2020; Rohrer, 2014; Schrepp et al., 2017), we addressed this by using validated, widely accepted measurement tools. Our results suggest that GenAI can enhance user experience in adult learning, with ChatGPT’s conversational nature playing a crucial role. Information was perceived as more accessible and engaging, contributing to interdisciplinary model-building that includes ethical and psychological aspects alongside information science.

Second, we showed that both self-efficacy and the use of GenAI have a significant positive influence on learning satisfaction. This finding aligns with prior research demonstrating that motivation and the belief in successfully mastering a learning task are key factors for learning (Baidoo-Anu & Ansah, 2023; Getenet et al., 2024; Zimmerman, 2000). In addition, our findings support the notion that human-like conversation contributes to increased satisfaction (Pangarso & Setyorini, 2023). Our results suggest that self-efficacy, is closely linked to GenAI-supported learning. Since GenAI enables exploratory and dialogical learning while maintaining user autonomy, our findings can be situated within the framework of self-determination theory (Deci et al., 1996; Deci & Ryan, 2012), which emphasizes autonomy and self-directed engagement in learning processes.

Third, we found that both user experience and learning satisfaction significantly impacted the number of answers participants marked as “unknown” in our quizzes. Moreover, the number of “unknown” responses had a significant negative effect on the number of correctly answered quiz questions. In other words, participants who felt more certain tended to perform better overall. This finding is particularly relevant in light of existing research on learner experience, which suggests that positive experiences and higher satisfaction can lead to longer engagement due to increased enjoyment and fulfillment (Hassenzahl, 2013; Kim et al., 2025). This indicates that GenAI use does not only affect knowledge acquisition but also shapes learners’ metacognitive awareness—their recognition of what they do and do not know. While higher certainty was associated with better quiz performance, this does not necessarily imply deeper or more durable learning. Our findings therefore underscore the need to

distinguish between perceived knowledge gains (metacognitive awareness) and actual learning outcomes, especially when evaluating GenAI-supported education. Moreover, in real-world informal learning scenarios, users may also lack the skills or resources to critically verify AI-generated information. In such cases, reliance on GenAI could reinforce perceived certainty without adequate validation. This underscores the continued relevance of traditional search methods, which expose learners to multiple independent sources, and the importance of fostering critical evaluation skills when engaging with GenAI tools.

Regarding the practical contribution of this study, we encourage educational institutions, teachers, and learning platforms offering digital learning opportunities to embed GenAI systematically into learning environments. By integrating GenAI tools such as ChatGPT into digital self-learning processes, user experience and learning satisfaction can be significantly improved. However, we recommend that the introduction of GenAI-supported learning should be accompanied by appropriate pedagogical and technological support to avoid digital divides among learners and to further promote collaboration, communication, and the exchange of knowledge between people. Additionally, learners need guidance in critically assessing AI-generated content to counter risks of misinformation. GenAI integration should not be treated as a one-off technological upgrade, but as a continuous development that aligns with evolving learner needs.

A critical reflection on our work reveals several limitations and directions for future research. Current GenAI tools such as ChatGPT are prone to generating hallucinations, i.e., inaccurate or fabricated information and references (Alavi et al., 2024; Brynjolfsson et al., 2025; Feuerriegel et al., 2024). This poses risks in education, as learners may accept misleading content; future studies should therefore compare different GenAI systems (e.g., Copilot) and emphasize the importance of critical source evaluation skills. Our learning outcome measure captures only immediate effects through post-task quizzes and does not assess long-term retention or critical reflection. Moreover, it may conflate certainty with actual knowledge: fewer “unknown” responses could reflect confidence rather than competence. Although guessing was reduced by offering an “I don’t know” option, future work should add delayed and performance-based measures. In addition, the five-minute minimum learning duration restricts the depth of engagement, meaning our findings reflect initial learning experiences rather than sustained knowledge acquisition. Longer sessions and repeated exposure should be explored in future work. The study also focused solely on adults aged 18 and above with

sufficient digital literacy, limiting generalizability to younger or less digitally skilled populations. Methodologically, the exclusion of African participants relied on a simplification; future studies should use pre-tests of topic familiarity instead. While tasks were standardized in terms of topic, exposure time, and interface, we did not prescribe guiding questions to preserve autonomy, which limits strict comparability. In addition, ChatGPT prompting and Google browsing involve fundamentally different cognitive processes, so our comparison reflects information-access modalities rather than equivalent learning strategies. Although we collected behavioral traces (queries, prompts, links, responses), these were not analyzed here and represent a promising avenue for future research. Finally, user experience remains a subjective construct despite our use of validated scales, and future studies should examine whether effects hold across different content types, including emotional or controversial topics.

To conclude, GenAI offers intuitive and interactive support for informal learning, addressing challenges in accessing comprehensible information on unfamiliar topics. Our findings show that ChatGPT, compared to Google, significantly enhances user experience, learning satisfaction, and learning success. Moreover, self-efficacy further increases satisfaction, while positive experience and satisfaction reduce knowledge gaps, underlining the transformative potential of GenAI in self-directed learning.

6. References

- Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education, 126*, 334–345. <https://doi.org/10.1016/j.compedu.2018.07.021>
- Alavi, M., Leidner, D. E., & Mousavi, R. (2024). Knowledge Management Perspective of Generative Artificial Intelligence. *Journal of the Association for Information Systems, 25*(1), 1–12. <https://doi.org/10.17705/1jais.00859>
- Almaiah, M., Alfaisal, R., Salloum, S., Al-Otaibi, S., Al Sawafi, O., Al-Marouf, R., Lutfi, A., Alrawad, M., Mulhem, A., & Awad, A. (2022). Determinants Influencing the Continuous Intention to Use Digital Technologies in Higher Education. *Electronics, 11*(18), 2827. <https://doi.org/10.3390/electronics11182827>
- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review, 13*(3), 219–235. <https://doi.org/10.1177/1088868309341564>
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52–62.
- Bandura, A. (1977). Social learning theory. *Englewood Cliffs*.

- Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33(1), 63. <https://doi.org/10.1007/s12525-023-00680-1>
- Barrett, C., Boyd, B., Bursztein, E., Carlini, N., Chen, B., Choi, J., Jagielski, M., Mirheidari, S., Roberts, M., Ristenpart, T., Tramer, F., Welinder, P., & Yang, D. (2023). Identifying and mitigating the security risks of generative AI. *Foundations and Trends in Privacy and Security*, 6(1), 1–52. <https://doi.org/10.1561/33000000023>
- Bockshecker, A., Ebner, K., & Smolnik, S. (2022, January). Technology-Enhanced Learning Environments and Adaptive Learning Systems-Development of Functionality Taxonomies. In *HICSS* (pp. 1-10).
- Borah, A. R., Nischith, T. N., & Gupta, S. (2024). Improved learning based on GenAI. 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), 1527–1532.
- Boubker, O. (2024). From chatting to self-educating: Can AI tools boost student learning outcomes? *Expert Systems with Applications*, 238, 121820.
- Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at Work. *The Quarterly Journal of Economics*, qjae044. <https://doi.org/10.1093/qje/qjae044>
- Caspersen, J., Smeby, J.-C., & Olaf Aamodt, P. (2017). Measuring learning outcomes. *European journal of education*, 52(1), 20–30.
- Chiu, T. K. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, 6, 100197.
- Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 302(1110), 341–359.
- Dalgıç, A., Yaşar, E., & Demir, M. (2024). ChatGPT and learning outcomes in tourism education: The role of digital literacy and individualized learning. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 34, 100481. <https://doi.org/10.1016/j.jhlste.2024.100481>
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application* (Nummer 1). Cambridge university press.
- Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. *Handbook of theories of social psychology*, 1(20), 416–436.
- Deci, E. L., Ryan, R. M., & Williams, G. C. (1996). Need satisfaction and the self-regulation of learning. *Learning and Individual Differences*, 8(3), 165–183. [https://doi.org/10.1016/S1041-6080\(96\)90013-8](https://doi.org/10.1016/S1041-6080(96)90013-8)
- Drucker, P. F. (1999). *Managing for the future*. Truman Talley Books/Dutton.
- Eberts, R. E. (1994). *User interface design*. Prentice-Hall, Inc.
- Edwards-Fapohunda, M. O., & Adediji, M. A. (2024). Sustainable development of distance learning in continuing adult education: The impact of artificial intelligence. *IRE Journals*, 8(1), 113–114.
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Folkestad, G. (2006). Formal and informal learning situations or practices vs formal and informal ways of learning. *British journal of music education*, 23(2), 135-145. <https://doi.org/10.1017/S0265051706006887>
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39. <https://doi.org/10.2307/3151312>
- Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, 33(2), 678–696. <https://doi.org/10.1287/isre.2021.1079>
- Gadiraju, U., Yu, R., Dietze, S., & Holtz, P. (2018). Analyzing Knowledge Gain of Users in Informational Search Sessions on the Web. *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval - CHIIR '18*, 2–11. <https://doi.org/10.1145/3176349.3176381>
- Getenet, S., Cante, R., Redmond, P., & Albion, P. (2024). Students' digital technology attitude, literacy and self-efficacy and their effect on online learning engagement. *International Journal of Educational Technology in Higher Education*, 21(1). <https://doi.org/10.1186/s41239-023-00437-y>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hassenzahl, M. (2013). User experience and experience design. *The encyclopedia of human-computer interaction*, 2, 1–14.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hinderks, A., Winter, D., Schrepp, M., & Thomaschewski, J. (2020). Applicability of User Experience and Usability Questionnaires. *JOURNAL OF UNIVERSAL COMPUTER SCIENCE*, 25, 1717–1735.
- Hussey, T., & Smith, P. (2008). Learning outcomes: A conceptual analysis. *Teaching in higher education*, 13(1), 107–115.
- Jackson, J. W. (2002). Enhancing Self-Efficacy and Learning Performance. *The Journal of Experimental Education*, 70(3), 243–254. <https://doi.org/10.1080/00220970209599508>
- Jia, X.-H., & Tu, J.-C. (2024). Towards a New Conceptual Model of AI-Enhanced Learning for College Students: The Roles of Artificial Intelligence Capabilities, General Self-Efficacy, Learning Motivation, and Critical Thinking Awareness. *Systems*, 12(3), 74. <https://doi.org/10.3390/systems12030074>
- Johnson, M., & Majewska, D. (2022). Formal, Non-Formal, and Informal Learning: What Are They, and How Can We Research Them? Research Report. Cambridge University Press & Assessment.

- Kim, J. S., Kim, M., & Baek, T. H. (2025). Enhancing User Experience With a Generative AI Chatbot. *International Journal of Human-Computer Interaction*, 41(1), 651–663. <https://doi.org/10.1080/10447318.2024.2311971>
- Kohnke, L., & Moorhouse, B. L. (2025). Enhancing the Emotional Aspects of Language Education through Generative Artificial Intelligence (GenAI): A Qualitative Investigation. *Computers in Human Behavior*, 108600. <https://doi.org/10.1016/j.chb.2025.108600>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *Relc Journal*, 54(2), 537–550.
- Lancy, D. F., Bock, J. C., & Gaskins, S. (Hrsg.). (2010). *The anthropology of learning in childhood*. AltaMira Press.
- Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 6, 100174. <https://doi.org/10.1016/j.caeo.2024.100174>
- McCoy, C. (2010). Perceived self-efficacy and technology proficiency in undergraduate college students. *Computers & Education*, 55(4), 1614–1617. <https://doi.org/10.1016/j.compedu.2010.07.003>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>
- Mittal, U., Sai, S., Chamola, V., & Sangwan, D. (2024). A Comprehensive Review on Generative AI for Education. *IEEE Access*, 12, 142733–142759. <https://doi.org/10.1109/ACCESS.2024.3468368>
- Nortvig, A.-M., Petersen, A. K., & Balle, S. H. (2018). A literature review of the factors influencing e-learning and blended learning in relation to learning outcome, student satisfaction and engagement. *Electronic Journal of E-learning*, 16(1), pp46-55-pp46-55.
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192. <https://doi.org/10.1126/science.adh2586>
- Pangarso, A., & Setyorini, R. (2023). The drivers of E-learning satisfaction during the early COVID-19 pandemic: Empirical evidence from an Indonesian private university. *Cogent Education*, 10(1), 2149226. <https://doi.org/10.1080/2331186X.2022.2149226>
- Pesovski, I., Santos, R., Henriques, R., & Trajkovic, V. (2024). Generative AI for customizable learning experiences. *Sustainability*, 16(7), 3034.
- Qadir, J. (2023). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. *2023 IEEE Global Engineering Education Conference (EDUCON)*, 1–9. <https://doi.org/10.1109/EDUCON54358.2023.10125121>
- Rohrer, C. (2014). When to use which user-experience research methods. *Nielsen Norman Group*, 12, 21.
- Roman, C., Delgado, M. Á., & García-Morales, M. (2025). Embracing the efficient learning of complex distillation by enhancing flipped classroom with tech-assisted gamification. *Education for Chemical Engineers*, 50, 14–24. <https://doi.org/10.1016/j.ece.2024.11.001>
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103. <https://doi.org/10.9781/ijimai.2017.09.001>
- Seaborn, K., & Fels, D. I. (2025). Keeping up with generative AI: Effects of engagement characteristics on user experience. *Behaviour & Information Technology*, 44(3), 215–230. <https://doi.org/10.1080/0144929X.2025.2483788>
- Shmueli, G., & Koppius, O. (2010). Predictive Analytics in Information Systems Research. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1606674>
- Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. Elsevier.
- Stone, E. R. (2010). T test, paired samples. *Encyclopedia of research design*, 1560–1565.
- Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & education*, 50(4), 1183–1202.
- Wagino, W., Maksum, H., Purwanto, W., Simatupang, W., Lapisa, R., & Indrawan, E. (2024). Enhancing Learning Outcomes and Student Engagement: Integrating E-Learning Innovations into Problem-Based Higher Education. *International Journal of Interactive Mobile Technologies (IJIM)*, 18(10), 106–124. <https://doi.org/10.3991/ijim.v18i10.47649>
- Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, 55(1), 10–33.
- Yang, Y., Tan, Y., Lin, Y. C., King, J. T., Liu, Z., & Lee, Y. C. (2025, April). Understanding How Psychological Distance Influences User Preferences in Conversational versus Web Search. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1-18).
- Zheng, Y., & Xiao, A. (2024). A structural equation model of online learning: Investigating self-efficacy, informal digital learning, self-regulated learning, and course satisfaction. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1276266>
- Zhonggen, Y., Ying, Z., Zhichun, Y., & Wentao, C. (2019). Student satisfaction, learning outcomes, and cognitive loads with a mobile learning platform. *Computer Assisted Language Learning*, 32(4), 323–341. <https://doi.org/10.1080/09588221.2018.1517093>
- Zimmer, M., & Niemimaa, M. (2019, July). *Navigating in the Digital Jungle: Articulating Combinatory Affordances of Digital Infrastructures for Collaboration*.
- Zimmerman, B. J. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, 25(1), 82–91. <https://doi.org/10.1006/ceps.1999.1016>