Understanding the Effect of Risk Perception on the Acceptance and Use of Large Language Models Among University Students

MICHAEL T. RÜCKER, Friedrich-Schiller-Universität Jena, Germany
CAROLIN BÜCHTING, HU Berlin, Germany
THOMAS KOSCH, HU Berlin, Germany

The rise of Large Language Models (LLMs) in education introduced powerful tools such as ChatGPT, which students increasingly use for academic purposes. However, these technologies present significant challenges for higher education, such as the risk of undermining academic integrity through AI-assisted performance and the uncertainty around proper use. This paper seeks to understand the benefits and risks university students perceive regarding LLM usage and how those influence their acceptance and use of related services in higher education. To this end, we employed a mixed-method approach. Using the UTAUT2 model extended by a *Risk Expectancy* construct, we conducted an online survey and follow-up interviews with university students. The results indicated that while students perceive considerable risks related to LLMs, those do not impact their usage and behavioral intention. We discuss this phenomenon based on the qualitative interview analysis and suggest research directions for future work.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**.

Additional Key Words and Phrases: Large Language Models, Higher Education, Students, Technology Acceptance

ACM Reference Format:
Michael T. Rücker, Carolin Büchting, and Thomas Kosch. 2025. Understanding the Effect of Risk Perception on the Acceptance and Use of Large Language Models Among University Students. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW512 (November 2025), 21 pages. https://doi.org/10.1145/3757693

1 Introduction

Large Language Models (LLMs) significantly impacted the educational landscape. According to a survey conducted by von Garrel et al. [49], by mid-2023, over 63% of students in Western Europe had used an AI tool in their studies, with nearly half of them relying on ChatGPT[1] specifically. LLMs, trained on vast datasets, offer powerful functionalities with applications across various academic fields. However, while universities are aware of students' increasing usage of these tools, there is little clarity on how and why they are used, leaving institutions uncertain about how to adapt to this new paradigm. At the same time, not engaging with such technologies could leave students unprepared for future professional environments that require the use or demand a critical reflection of LLMs.

[1]https://openai.com/chatgpt/overview

Authors' Contact Information: Michael T. Rücker, michael.ruecker@uni-jena.de, Friedrich-Schiller-Universität Jena, Jena, Germany; Carolin Büchting, carolin.buechting@web.de, HU Berlin, Berlin, Germany; Thomas Kosch, thomas.kosch@hu-berlin.de, HU Berlin, Berlin, Germany.

ACM 2573-0142/2025/11-ARTCSCW512
https://doi.org/10.1145/3757693

The main issue centers around the ambiguity of LLM usage in higher education [1, 19]. On one hand, LLMs are viewed as a productivity tool in academic research [18] or as a potential asset to support adaptive teaching and learning practices [36]. On the other hand, various authors also caution against the unethical use of LLM-based services and potentially detrimental effects on education, citing plagiarism, misinformation, privacy, hallucination, or copyright concerns [23, 24, 28, 29, 35]. Given this situation, very little is known about how students perceive various benefits and risks related to these technologies and how they affect their usage decisions. Consequently, the study of how users interact and adopt LLM-generated content gained interest in the computer-supported cooperative work and social computing community for investigating trust, safety, and responsibility in AI [27].

The growing discourse around both the potential benefits and risks of LLMs highlights the need for theoretical models that not only account for their perceived usefulness but also systematically incorporate users' risk perceptions. Models such as the Technology Acceptance Model (TAM) [10, 32] and the Unified Theory of Acceptance and Use of Technology (UTAUT2) [46] are established methods to predict technology adoption, but primarily measure intended system qualities, such as performance or usability, while negative side effects are not covered. Work to include risk perception into these models exists mainly in the context of e-commerce and has produced inconclusive results [17]. Given the above-mentioned discourse on the potentially harmful effects of LLMs in particular, it is reasonable to believe that perceived risk may play a more prominent role in usage decisions regarding these systems. Although work by von Garrel et al. [49] provides initial insights into the prevalence of related tool usage, a comprehensive account of students' risk perceptions and their effect on usage behaviors regarding LLM-based services for their studies is missing.

We conducted a mixed-methods [9] investigation to understand how students use LLM-based services, what factors drive their adoption, and how they perceive their use in academic settings. Using the UTAUT2 extended by a *Risk Expectancy* construct, we examined the factors driving students' acceptance and use of LLM-based services in their studies. In our work, we understand "risk" as students' perceived potential for negative consequences when using LLM-based services, including privacy concerns, academic integrity issues, and uncertainty about the reliability and objectivity of generated content. Subsequently, we conducted several in-depth interviews to gain more insights into students' perceptions and lived experiences regarding LLMs. This dual-method design enables us to capture a wide range of perspectives and contribute a new understanding of the academic use of LLMs. Our findings suggest that although students recognize significant risks associated with LLMs, these concerns do not appear to influence their usage or behavioral intentions. Further interviews with participants affirm these sentiments. Our research shows a disconnect between students' awareness of LLM-related risks and their actual usage behavior, suggesting that traditional acceptance models may underestimate the influence of motivational and contextual factors. This raises a broader question explored in this paper: Why do students continue to use LLM-based services despite perceiving them as risky? We examine this tension by combining survey data with qualitative insights, positioning our findings within broader debates around the behavioral paradoxes of technology use, such as the privacy paradox [16]. Understanding this gap is important for designing more effective educational policies and interventions that promote responsible and informed use of LLM-based services in academic settings. To this end, we discuss explanations for this observation drawn from the qualitative interview analysis and recommend an alternative theoretical framework for future research.

Contribution Statement

In this paper, we make three key contributions to understanding the acceptance of LLM-based services in higher education. First, we extend the UTAUT2 model with a novel *Risk Expectancy* construct to examine how perceived risks influence students' behavioral intentions. Second, a mixed-methods study combining survey data and in-depth interviews reveals that while students report high awareness of various risks, these do not significantly diminish their use of LLMs. Third, our qualitative analysis provides insight into why this gap exists, revealing how students rationalize risk and how motivational orientations shape their usage decisions. We frame our findings within known behavioral patterns such as the privacy paradox and established motivational theory, particularly goal orientations. In doing so, we expand the conceptual toolkit for understanding AI adoption in everyday academic life.

2 Theory and Related Work

Based on a survey of over 6000 university students in June 2023, von Garrel et al. [49] found that almost two-thirds were using AI-based tools for their studies, with a lower adoption rate among persons who self-identify as women. The literature suggests a diverse set of potentially valuable applications of LLM-based services, particularly ChatGPT, to study-related tasks. These include fostering central key qualifications such as language proficiency [15] and various research-related tasks [4, 19, 31]. Huge potential benefits have been proposed in generating individualized learning material and feedback [13, 19]. LLMs may also increase technology accessibility, for example, by generating image descriptions or providing other language-based interfaces [19].

However, various authors also caution against the unethical use of LLM-based services, again primarily ChatGPT, and potentially detrimental effects on education [28, 29, 35]. Apart from general concerns about the automation of human elements in education [8], arguably one of the most widely voiced concerns is the use of LLM-based services to cheat in examinations or to plagiarise academic work [20]. Issues related to the academic credit system notwithstanding, this may undercut the development of corresponding cognitive abilities and hinder learning [4, 13]. Moreover, model hallucinations and biases, giving rise to misinformation, can be detrimental to learning as well [6, 38]. Copyright and privacy concerns are also being discussed regarding training data and generation prompts [19, 31]. It is notable, however, that many of these concerns primarily exhibit an administrator or educator perspective. Whether and to what extent similar risks and benefits are being perceived and weighed by students remains unclear.

Factors influencing and predicting the acceptance and use of LLM-based services among students is an active and growing field of research. According to the technology acceptance model (TAM), the intention to adopt a particular technology can be predicted by two key variables: perceived usefulness and perceived ease of use [11, 32]. Later work has, respectively, recast these constructs as performance expectancy and effort expectancy, and extended them by various other factors, including hedonic motivation, age, gender, or habit, to formulate a Unified Theory of the Acceptance and Use of Technology (UTAUT and later UTAUT2) [44, 46]. While many potential benefits of using LLMs outlined above can be readily construed in terms of performance or effort expectancy, associated risks or other detrimental effects of usage are seemingly more challenging to integrate with these models.

Based on 373 self-reports collected from 373 post-graduate students, Tian et al. [43] investigated their intention to use AI chatbots in their studies. The authors found that performance expectancy was positively correlated with behavioral intention, while effort expectancy demonstrated no statistically significant correlation. Strzelecki [39] surveyed over 500 students regarding their intention to use ChatGPT and found performance expectancy, habit, and hedonic motivation to have the highest predictive power. This indicates that some students may use LLM-based services, particularly chatbots, simply for fun rather than to achieve a particular goal. Using survey data from 400 Spanish university students Romero-Rodríguez et al. [37] report similar results. However, contrary to the above-cited study by von Garrel et al. [49], they found no gender differences.

Based on their survey of over 600 US undergraduates, Albayati [5] found perceived usefulness and ease of use statistically significant predictors of usage intention regarding ChatGPT. Moreover, the authors found privacy concerns and social desirability to influence students' perceptions, indicating awareness of related potential risks. Using UTAUT as a framework, Menon and Shilpa [33] interviewed 32 students from India on their usage of ChatGPT and identified privacy concerns as an additional factor, whose statistical significance remains unclear.

Summary and Research Question

Overall, related work generally corroborates UTAUT's central variables, particularly performance expectancy, as valid predictors for students' acceptance and use of LLMs. However, given the broad versatility of LLM-based applications, particularly chatbots, perceived performance expectancy may vary greatly depending on the task. In that regard, existing research is much less informative. Moreover, insights into other forms of LLM-based services, as well as the role of perceived risks, remain much more tentative as well. On the one hand, students' lack of risk perception would readily explain the absence of related factors in predictive models. On the other hand, given the various concerns educators and policymakers raise, such a lack of awareness would hardly be desirable. In other words, an empirical account, however robust, of behavioral intention may still fall short of normative educational expectations. In our work, we define "risk" broadly to encompass multiple perceived threats related to LLM use in academic settings, including privacy breaches, academic misconduct (e.g., plagiarism or dependence), and concerns over the reliability or bias of generated content. These categories of risk, though often conflated in public discourse, may have distinct implications for student behavior. Therefore, by investigating the following research question, our study provides a more differentiated and student-centered perspective on the factors influencing the acceptance and use of LLM-based services for study-related purposes.

RQ: What benefits and risks do students perceive concerning using LLM-based services for their studies, and how do these perceptions influence their corresponding usage behavior?

Based on previous research and the general discourse on LLMs, perceived LLM risk denotes the perceived potential for unwanted detrimental side effects, which may fall into a variety of categories: privacy risks, referring to concerns over data security and information exposure when using LLM tools; academic integrity risks, related to issues such as plagiarism, reliance on AI-generated content, and potential violations of institutional policies; and trust-related risks, concerning students' confidence in the accuracy, fairness, and potential biases of LLM-generated responses. Three general scenarios appear plausible. Students may only perceive negligible risk, which then, naturally, does not have a notable effect on their usage behavior. However, given the prevalence of risk in the public discourse on AI in general and LLMs in particular, we find that scenario unlikely. Instead, we expect that students do perceive notable risks, which then either do or do not have a significant effect on usage. In either case, a more in-depth understanding of the mechanisms underlying or preventing such effects in the context of LLMs is desirable.

3 Methodology

This study employs a sequential explanatory mixed-methods design [9], which aligns with our dual goals of identifying generalizable usage patterns of LLM-based services and unpacking the reasoning behind those patterns. We first conducted a quantitative online survey to examine students' acceptance of LLMs using the UTAUT2 framework, extended by a novel *Risk Expectancy* construct. This step enabled us to identify key predictors of behavioral intention through statistical analysis. Subsequently, we conducted in-depth interviews with a purposefully sampled subset of survey participants to contextualize and explore students' lived experiences, motivations, and perceptions in greater depth. This design strengthens the empirical robustness of our findings but also improves their interpretability by connecting statistical analysis with individual sensemaking.

The remainder of the paper is structured as follows. In Section 4 and Section 5, we present the procedures and results of the survey and interview studies, respectively. Section 6 then jointly discusses the implications and limitations. Section 7 concludes the paper with a summary and outlook.

4 Online Survey

We conducted an online survey to understand which factors influence the use of LLMs in higher education. The online survey aims to understand the demographics of the students who use LLMs, how frequently they use them, and how they adopt and use LLMs for their studies.

4.1 Participants

We recruited 156 participants via mailing lists, student networks, and social media platforms (e.g., LinkedIn, WhatsApp, and Signal groups). Overall, the study involved students currently enrolled in universities. We used Google Forms to collect the answers. Participation was voluntary, and participants did not receive compensation.

In total, 156 participants (80 self-identified as female, 71 self-identified as male, one self-identified as diverse, and four participants preferred not to describe their gender) met the enrolment criteria. The mean age was $\bar{x} = 22.81$ ($s = 1.96$). The students studied STEM subjects, social sciences, economics, jurisprudence, and medicine.

4.2 Survey Structure

We adapted the constructs from the UTAUT2 to include inquiries about LLM adoption, perception, and usage [45, 47]. As a consequence, we adapted the items from the constructs Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Conditions (FC), Hedonic Motivation (HM), Price Value (PV), Habit (HT) and Behavioral Intention (BI) to focus on the use of LLMs of students. We also designed items on an additional construct, which we termed *Risk Expectancy* (RE), whether students also perceive potential negative consequences of using LLMs in their study and how this might affect their behavioral intention. Table 1 shows the complete list of the adapted items. All items were rated on a seven-point Likert scale, where 1 indicated total disagreement ("Strongly Disagree") and 7 total agreement ("Strongly Agree"). Additionally, we asked the participants for their age, self-identified gender, familiarity with LLMs, and how frequently they use LLMs for their studies. Lastly, participants were asked whether they agreed to be contacted for a follow-up interview (see Section 5).

Table 1. Survey items on LLM-based services for studying. Cronbach's α revealed a consistency for all constructs except the construct Facilitation Conditions (FC). The statements were translated directly from their original language into English. The items were rated on a seven-point Likert item.

Construct	Item	Statement
Performance Expectancy (PE)	PE1	I find LLM-based services useful for my studies.
Cronbach's α = 0.92	PE2	LLM-based services increase my chances of achieving study-relevant goals that are important to me.

Continued on next page

Construct	Item	Statement
	PE3	I can complete study-related tasks more quickly with LLM-based services.
	PE4	I increase my productivity in studies with LLM-based services.
Effort Expectancy (EE)	EE1	It is easy for me to learn to use LLM-based services for my studies.
Cronbach's α = 0.84	EE2	My interaction with LLM-based services in my studies is clear and understandable.
	EE3	I find LLM-based services easy to use in my studies.
	EE4	It is easy for me to become proficient in using LLM-based services for my studies.
Social Influence (SI)	SI1	People who are important to me think that I should use LLM-based services for my studies.
Cronbach's α = 0.88	SI2	People who influence my behavior think that I should use LLM-based services for my studies.
	SI3	People whose opinions I value prefer that I use LLM-based services for my studies.
Facilitating Conditions (FC)	FC1	I have the necessary resources to use LLM-based services for my studies.
Cronbach's α = 0.60	FC2	I have the necessary knowledge to use LLM-based services for my studies.
	FC3	LLM-based services are compatible with other technologies that I use for my studies.
	FC4	I can get help when I have difficulties using LLM-based services for my studies.
Hedonic Motivation (HM)	HM1	Using LLM-based services for my studies is fun.
Cronbach's α = 0.75	HM2	Using LLM-based services in my studies is enjoyable.
	HM3	Using LLM-based services in my studies is very entertaining.
Price Value (PV)	PV1	LLM-based services for studying are inexpensive.
Cronbach's α = 0.89	PV2	LLM-based services provide good value in the study context.
	PV3	At the current price, LLM-based services offer good value for my studies.
Habit (HT)	HT1	Using LLM-based services for my studies has become a habit for me.
Cronbach's α = 0.70	HT2	I am addicted to using LLM-based services in my studies.
	HT3	I must use LLM-based services for my studies.
	HT4	The use of LLM-based services for my studies feels natural to me.

Continued on next page

Construct	Item	Statement
Risk Expectancy (RE)	RE1	I believe that using LLM-based services in my studies involves uncertainties that should be considered.
Cronbach's α = 0.80	RE2	I see risks associated with using LLM-based services in my studies.
	RE3	I believe that using LLM-based services in my studies is associated with certain risks.
	RE4	I find the use of LLM-based services in my studies risky.
Behavioral Intention (BI)	BI1	I intend to use LLM-based services for my studies in the future.
Cronbach's α = 0.93	BI2	I will always try to use LLM-based services in my daily studies.
	BI3	I plan to use LLM-based services regularly for my studies in the future.

4.3 Results

The following section presents the results of our survey. We begin by assessing the internal consistency within the constructs. Then, we present the answers regarding familiarity, frequency of LLM use, UTAUT2, and, as suggested by previous work [43], correlations between behavioral intentions and frequency of use.

4.3.1 Assessing the Internal Consistency Within Constructs. We begin by investigating the internal consistency of the items by calculating Cronbach's α. Cronbach's α measures the consistency of the items within a construct [42]. We discarded constructs with a Cronbach's α of less than 0.7 and kept constructs with an α value equal or higher to 0.7. Previous work suggested that values above 0.7 are acceptable and provide sufficient internal consistency for measuring the same construct [41]. Our results showed that all constructs except FC showed sufficient reliability (see Table 1). Consequently, we removed the construct *Facilitating Conditions* from our analysis while keeping all other constructs for further investigation.

4.3.2 Familiarity and Frequency of Using LLMs for Education. Figure 1 and Figure 2 show the distribution of the familiarity with LLMs and the use frequency. Asking the question "*How familiar are you with LLM-based services?*" showed that twelve participants indicated no familiarity, 39 participants had low familiarity, 55 participants had some familiarity, 41 had a high familiarity, and finally, nine participants had a very high familiarity (see Figure 1). Asking the question "*How often do you use LLM-based services?*" showed that 20 participants never used LLMs, 49 used them a few times, 30 used them one to three times per month, 40 used them three times per week, and 17 used them almost every day (see Figure 2).

4.3.3 UTAUT2. We used the UTAUT2 questionnaire to investigate the user acceptance and usage of LLMs of our participants (see Table 1). We calculated the average student responses and reported them along with the standard deviation. Figure 3 visualizes the mean and distribution of the answers for each construct. Table 2 shows the mean values for each build and their respective Spearman rank correlations with *Behavioral Intention* and Use Frequency. *Risk Expectancy* exhibited the highest mean score among all constructs, indicating widespread concern among students, yet it showed no significant correlation with either Behavioral Intention or actual Use Frequency. This suggests a potential disconnect between students' perception of risk and their actual usage behavior, contrary

ID	Statement
Q1	I have no familiarity with LLMs.
Q2	I have a low familiarity with LLMs.
Q3	I have some familiarity with LLMs.
Q4	I have a high familiarity with LLMs.
Q5	I have a very high familiarity with LLMs.

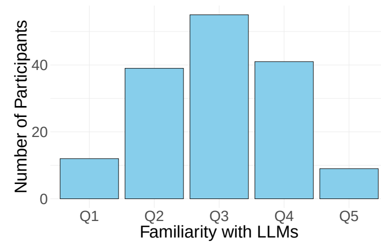

Fig. 1. Distribution of LLM familiarity among the participants. Our results show that most students are familiar with using LLMs.

ID	Statement
Q6	Never.
Q7	A few times.
Q8	One to three times per month.
Q9	One to three times per week.
Q10	Almost every day.

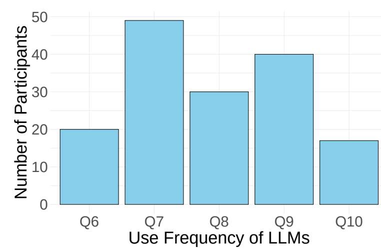

Fig. 2. Use frequency of LLMs. Most participants have used LLMs a few times before.

to what might be assumed based on the prominence of risks in academic and public discourse around LLMs. Using follow-up interviews we further investigated the underlying motivations and reasoning behind those patterns.

Table 2. Mean and standard deviation of each UTAUT2 construct with a Cronbach's $\alpha \geq 0.7$. We calculated Spearman's rank correlation coefficient for the UTAUT2 constructs with the construct *Behavioral Intention* and the *Use Frequency*. All correlations with *Behavioral Intention* and the other constructs were significant ($p < .001$) except for *Risk Expectancy* ($p = .067$). Similar to *Behavioral Intention*, all correlations with *Use Frequency* and the other constructs were significant ($p < .001$) except for *Risk Expectancy* ($p = .197$). Bold numbers indicate the highest and lowest numbers.

Constructs	$\bar{x}$	s	Spearman's Correlation Behavioral Intention	Spearman's Correlation Use Frequency
Performance Expectancy	4.18	1.82	**0.87**	0.71
Effort Expectancy	4.33	1.36	0.64	0.58
Social Influence	2.92	1.56	0.61	0.41
Hedonic Motivation	4.37	1.41	0.70	0.55
Price Value	4.50	1.66	0.50	0.39
Habit	**2.27**	1.18	0.85	**0.77**
Risk Expectancy	**5.32**	1.10	**-0.14**	**-0.10**
Behavioral Intention	3.72	1.99	-	**0.77**

5 Follow-Up Interviews

While the survey results reported above, particularly the low predictive value of *Risk Expectancy*, are in line with previous research on UTAUT2, they are nevertheless surprising and certainly

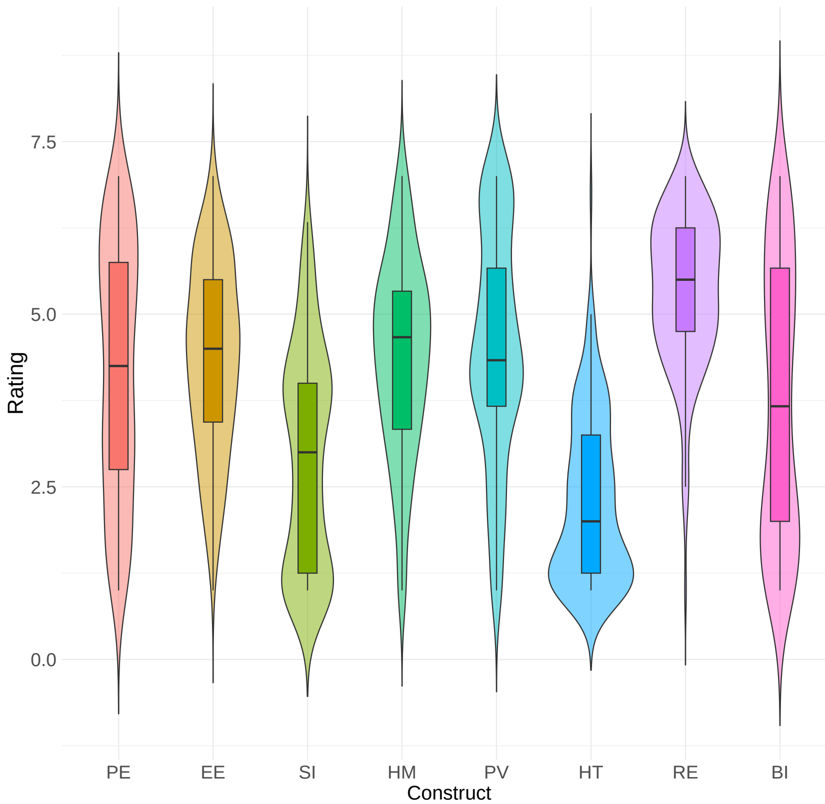

Fig. 3. Mean UTAUT2 values for each construct. *Risk Expectancy* achieved the highest mean score, while Habit received the lowest mean score.

undesirable. From an educational perspective, students *should* factor perceived risk into their usage decisions. To further explore this dynamic and to provide more contextual interpretation for the quantitative findings, we conducted follow-up interviews to investigate how students make sense of perceived risks, what factors drive or override those concerns, and how different usage goals and educational motivations may influence their decisions.

5.1 Sampling and Procedure

49 survey participants agreed to be contacted for a follow-up interview. The interviews aimed to map out various viewpoints on LLM-based services and provide a more differentiated account of perceived benefits and risks. Therefore, we implemented purposive maximum variation sampling [7, p. 114f] with respect to LLM acceptance and use. To this end, we conducted a k-means clustering on the complete survey data, generating four clusters. Also considering age, gender, usage experience, and study discipline, we contacted eight participants for a follow-up interview, two from each cluster. Figure 4 visualizes the four clusters in 2D based on principal component analysis and indicates the selected participants. Table 4 provides additional information on the factors above for each participant. All eight were successfully contacted and interviewed.

The interviews followed a semi-structured guideline divided into three sections, leaving room for spontaneous inquiries. Each section ended with an open question, allowing participants to formulate

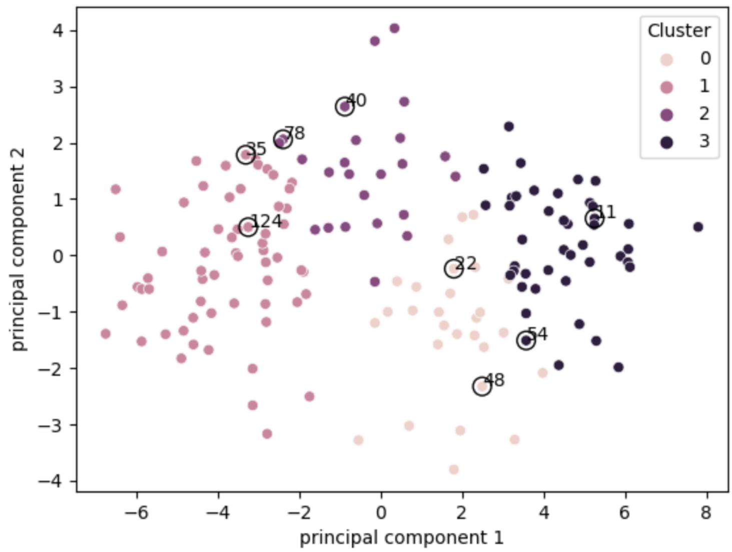

Fig. 4. Visualisation of the four k-means clusters based on a principal component analysis. Circled dots indicate the selected interview participants.

individual responses. We aimed to obtain a more holistic view of different usage intentions and influences. Hence, we prioritized gathering rich and authentic responses rather than striving for comparability between the interviews.

5.1.1 Usage Context: Participants were asked what they understood by LLM-based services, what related services they knew about, and which they were currently using for their studies, and how. Participants not using any LLM-based services were asked what potential applications they knew about and what kept them from using them. This section aimed to gain an overview of the participants' current use of LLM-based services as a basis for the remainder of the interview.

5.1.2 Usage Factors: Based on this, participants were asked more about the factors facilitating or preventing their use of LLM-based services in their studies. If applicable, they were asked when they first started using related services and what prompted them. Subsequently, the interview inquired about the most influential factors identified during the survey study, including performance expectancy, hedonic motivation, habit, and effort expectancy. Participants were presented with their corresponding survey responses and asked to explain why they responded the way they did. We aimed to gain a more differentiated understanding of how participants conceived of these constructs and their influence on their usage behavior.

5.1.3 General Assessment: In the final section, participants were asked for a more general evaluation of the benefits and risks of using LLM-based services in their studies. Analogous to the previous section, they were asked to comment on their survey responses related to *Risk Expectancy*. Moreover, they were asked about their wishes for the future and the necessity of legal regulations around LLM-based services.

Table 3. Main categories and exemplary selected sub-categories of the final coding system. Every line of each category represents a sub-category.

Performance Expectancy	**Effort Expectancy**	**Social Influence**	**Facilitating Conditions**	**Hedonic Motivation**
generate ideas	techn. availability	friends & family	no clear rules	entertainment
improve text	service selection	fellow students	prohibition	work, not fun
summarize inf.	training	(social) media		pleasure
mathematics	user experience	non-issue		frustration
literature citations	prompting			
…	…			
Price Value	**Habit**	**Behavioral Intention**	**Risk Expectancy**	
free version	not established	pos. stmt. of intent	dangerous output	
premium version	established	neg. stmt. of intent	reduced learning	
		general affinity	alleg. of plagiarism	
		general reluctance	privacy	
			…	

5.2 Data Analysis

The interviews were conducted via Zoom and lasted between 29 and 54 minutes. They were audio-recorded and transcribed verbatim. The data was then subjected to a structuring qualitative content analysis [26], using a combination of deductive and inductive coding. The overall aim of the analysis was to gain a better understanding of the UTAUT2 constructs plus *Risk Expectancy* and their possible interactions in the specific context of LLM-based services.

The eight central UTAUT2 factors and *Risk Expectancy* were thus defined as the main categories of analysis and first applied deductively to the interview transcripts. Already during this phase, several passages were identified that were difficult to clearly assign to only one category. We, therefore, allowed multiple categories to be assigned to the same section. For each main category, sub-categories were then developed inductively in order to capture specific expressions for each factor. That is, new sub-codes were defined, for instance, for concrete application contexts (PE), barriers (EE), or concerns (RE) mentioned by the interviewees. In that phase, we also specifically looked for plausible interactions between the main categories, particularly in those sections assigned to multiple categories. The sub-codes were refined during several coding iterations until further revisions were judged negligible. While the general coding focus thus shifted from deductive to inductive over the course of the analysis, these two phases were not entirely separate but partially occurred in parallel, as some rather obvious sub-codes were introduced early, and later revisions could still result in a change of main category.

The analysis was mainly conducted by the second author, who consulted with the other two authors for peer feedback on coding strategies and initial insights. After the analysis was completed, the second author presented their findings to the others to discuss the results and interpretations. Subsequently, the first author also conducted a confirmability audit of the coded data [3] based on the complete MAXQDA project file containing all transcripts, codes, and code assignments. This resulted in minor revisions of the code structure but did not affect the overall results of the analysis. Table 3 shows an overview of the final code system, including all main categories and selected sub-categories.

Table 4. Selected interview participants.

cluster	index	discipline	age	gender	experience
0	48	nutritional science	23	f	very high
	22	natural science	23	m	high
1	54	social science	23	f	high
	11	economics	21	m	medium
2	40	math/computer science	21	f	medium
	78	liberal arts	25	m	low
3	35	education	25	f	medium
	124	political science	20	m	none

5.3 Results

Unsurprisingly, the most commonly mentioned service was ChatGPT, and four of the eight participants did not know about or use any other service. However, other services like QuillBot, DeepL, or Grammarly were mentioned as well, and one participant explicitly stated that they never used just one service by itself, but always a "combination," which together formed "a kind of workflow."[2]

In terms of performance expectancy, one of the main motivations was to save time on specific tasks and get things done faster.

> What motivated me, I think, was the speed at which you can get things done.

Participants mentioned a range of different tasks for which they described LLM-based services as sufficiently performant. These included, above all, the production and improvement of different kinds of text, ranging from written correspondence to term papers to program code. However, the particular kind of usage could differ significantly, from generating first ideas and outlines, improving style, and generating complete solutions.

> Just say, okay, I type this into it, I say here, write this letter (...) using this data.

> You can sometimes input a written text and get a better wording.

> And I once got an idea for an implementation of a simple program from ChatGPT.

> I often use ChatGPT for creative brainstorming.

Simple knowledge questions to "summarize or explain certain terminology" were also mentioned, and one interviewee related an instance where some of their fellow students used ChatGPT to focus their reading assignment:

> What sections can I build on and where should I pay particular attention when I read it myself, to get some food for thought.

Another used ChatGPT to generate feedback to a solution draft:

> Sometimes I present my solution and use ChatGPT to ask if it makes sense. (...) Have I forgotten anything? Are there contradictions?

In that sense, LLM-based services were also perceived as a possible learning support, which, according to one student, may be particularly useful in cases where human tutors are not available. However, participants also reported several instances where the generated results did not meet their expectations. This pertained, for instance, to faulty results, misinformation, or lack of emotion. In

[2]All presented data quotes are translated from the native language verbatim.

particular, several participants reported that ChatGPT seemingly struggled with higher mathematics, and some mentioned the issue of "made-up research studies" and an overall inability to generate proper citations. This point also pertains to *Risk Expectancy* and we will return to this interrelation below.

Concerning effort expectancy, several interviewees described the actual use of LLM-based services as rather accessible, referring, for instance, to natural-language interfaces, for which "you don't have to learn any programming code." However, some also described notable efforts to formulate and re-formulate appropriate prompts in order to arrive at the desired results. Depending on the context, these efforts could outweigh the expected gains such that by the time you get it right "I might as well have done it myself." Some students also described entry thresholds, particularly concerning first-time setups:

> What version do I use? Do I want to spend money? Do I have to install something? Do I need an account? How does it work?

Some also referred to their "affinity to technology" to explain why they didn't find certain services intuitive. Notably, the students who reported on such entry thresholds were also the ones who, according to the prior survey, had less experience using LLM-based services. This corroborates the previously reported quantitative finding that effort expectancy notably influences usage behavior.

The interviews also mentioned the effects of *habit* on usage behavior. While some reported that using LLM-based services had already become routine, the more frequent response was to the contrary. Several interviewees reported that it often did not even occur to them to use LLMs for a particular task:

> I often don't use it even though I could, not because I explicitly decide against it, but just because I don't think of it.

> I completely forget about it. It's just not at the back of my mind.

Habit also appears to play a role in the notable prevalence of ChatGPT in students' responses, even though very similar services have since become available:

> I already know about it and use it and that's why I don't use the other ones. It's not a conscious decision (...) just habit.

For *social influence*, some students mentioned social media, particularly videos on YouTube or Instagram:

> The biggest influence, I think was this somewhat popular Youtuber and I knew from previous experience that his tips for working more efficiently are pretty good. So that made me trust in the technology and I gave it a try.

Personal conversations or peer experience reports, particularly by friends, family and fellow students, were also mentioned as having influenced usage decisions. One student even described a rather competitive fear of being left behind

> I mean if everyone uses it but you don't, then it's a little like WhatsApp, or the smartphone, the computer. At some point you are just cut off.

However, negative influences were described as well, both on an institutional and personal level

> The people I talk about and also the fact that my French uni isn't very positively disposed toward it.

As already mentioned, several students reported that they found using LLM-based services rather effortless and convenient, for example, due to saving time, which several described as a pleasant experience. In that sense, *hedonic motivation* depends on good performance. However, in another sense, it also seemed to exist in a certain tension, if not outright conflict, with performance expectancy:

> Yeah, I don't know. It's amusing to let ChatGPT formulate some dumb hook up lines, but I wouldn't say that's very useful.

> Sometimes you get nonsense answers and that's always rather funny because it sounds so sophisticated and formal, but is completely wrong. That's always funny.

"Dumb" or "nonsense" answers may be a source of entertainment or fun but are also detrimental to utility. The interplay of these effects is likely highly situational and task-dependent. If entertainment is the goal, hilarious output is a source of fun. If the goal is to get actual work done, however, it is just as easily a source of frustration. Compared to other technical systems, generating entertaining content may be a genuinely new quality of LLMs, suggesting a more complex role of hedonic motivation in their use.

Concerning *Risk Expectancy*, students reported on a broad range of perceived risks of using LLM-based services, both for society in general and the specific context of their studies. The former included, for instance, the displacement of jobs, particularly journalistic professions, or the intentional dissemination of political propaganda and fake news, which could be generated en masse using LLM-based services. Note that such *intentional* misinformation differs from the generation of *unwanted* misinformation, which we will discuss below. Data privacy and copyright issues were mentioned as well:

> I have also already put in text containing personal information about me and I don't know if they were stored somehow.

> Because it is always unclear where the information comes from that ChatGPT is referencing.

Apart from the processes and information used to generate certain output, students also expressed uncertainty about the legal status of that output and who owns any rights to it.

The most commonly mentioned risk was that the information generated by LLMs may be unreliable, faulty, biased or discriminating, particularly with respect to racism or sexism.

> I think I see the risk of misinformation and that people might think it's correct because it comes from some online source.

> And I think that ChatGPT and all these services mirror all our human weaknesses and problems, especially with respect to sexism or racism. And you need to know that these services are not perfect, just because they come from a computer.

Indeed, several students described the necessity of a critical attitude toward LLM-generated responses, and practices of double-checking:

> And then I read it again to see, OK, does this make sense?

> I see that people use such systems and don't have a clue about the kind of botched stuff that may come out.

All these issues about misinformation, errors, or biases could also be framed in terms of performance expectancy because the output quality does not meet the user's requirements. Moreover, double-checking the system's output requires further effort and its importance likely increases in the context of high-stakes scenarios such as term papers or theses:

> I am currently in the middle of my Master's thesis, which has to adhere to scientific standards, and because ChatGPT's sources are always very fuzzy, I stay away from it for the time being.

Hence, there likely exists a three-fold interaction between performance expectancy, effort expectancy and *Risk Expectancy*. This provides a plausible partial explanation for why *Risk Expectancy* was not observed to have a notable influence on usage in the quantitative survey, as significant portions of it may instead have been conflated with performance and effort expectancy. However, the interviews also provided additional insights into the construct of *Risk Expectancy*:

> Well the question of using or not using, I'd say, essentially comes down to two things. Whether it is functional and does what I want, but also whether for me personally it makes sense that I don't do it myself.

> Well, one disadvantage can certainly be that you tend to overuse it and maybe reduce your own learning output.

> You make it easy for yourself and then you don't learn it anymore.

These participants exhibit an awareness that delegating certain efforts to an LLM may rob them of valuable learning opportunities, regardless of the system's expected quality of performance. In fact, increased performance expectancy may result in a higher temptation to avoid an otherwise challenging task, further *increasing* the risk of reduced learning. Compare this to the following statement about plagiarism:

> I think the fear to get caught. Because many use these services to just let them write a term paper or something with made-up sources.

Several other students mentioned the risk of being accused of plagiarism, which in an educational context is essentially cheating, an attempt to avoid a cognitive challenge and feign a level of mastery not attained. In contrast to the risk of reduced learning, the risk of being caught cheating may indeed *decrease* with more sophisticated LLMs as their output becomes harder and harder to distinguish from genuine human performance. We believe this is a very important point to make that informs different motivations for using LLMs in a learning context, which we will discuss further in the next section.

Concerning *facilitating conditions* and given the above concerns, particularly regarding data privacy, copyright, and plagiarism, it is unsurprising that several students expressed a favorable opinion regarding clear rules and regulations of LLM use by their universities. While an outright ban was endorsed by no one, several called for boundaries and enforceable consequences of misuse:

> In my opinion there should be regulation. Students should be allowed to use these services. Period. (...) But if I, say, during an exam I take out my phone to ask ChatGPT, then that is obviously cheating.

> I am a little afraid that they just make up some rule, but then you don't know how that is actually enforced. And my uni does this frequently.

In summary, the qualitative findings generally corroborate the quantitative survey results. They also provide insight into the specific meaning of these factors and plausible interrelations between them in the specific context of LLM use in education, particularly between performance expectancy, effort expectancy, hedonic motivation, and *Risk Expectancy*.

6 Discussion

To answer our research question, we investigated the benefits and risks students perceive concerning using LLM-based services for their studies and how these influence their usage behavior. The results of the quantitative survey (cf. Section 4) are consistent with related prior work, which has found the UTAUT2 constructs performance expectancy, effort expectancy, habit, and hedonic motivation to be valid predictors for behavioral intention in that context [5, 37, 40].

Additionally, given the prevalence of the potentially harmful effects of LLMs in the public and academic discourse, we suspected that perceived risks may also influence students' usage decisions. Therefore, we investigated students' *Risk Expectancy* and found that, while many indeed perceive notable usage risks, this does not seem to affect behavioral intention. On the one hand, the finding that *Risk Expectancy* does not contribute to predicting behavioral intention is consistent with UTAUT2 in its present form, further corroborating the theory in this new context. On the other hand, it raises questions about the specific mechanisms that prevent these perceptions from being acted upon. Given the potential adverse effects of LLM usage in educational contexts [29, 35, 38], ignoring perceived risk is hardly desirable behavior.

The analysis of the interview data (see Section 5) shows that participants indeed perceived a large variety of potential benefits and risks of using LLM-based services in their studies, mainly mirroring those put forward in the literature. However, several of the more general risks mentioned by interviewees, such as hallucinated or biased results, are also readily interpretable as reduced performance expectancy. This may have conflated these two constructs in the quantitative survey and thus contributed to explaining the negligible effect of risk perception on usage behavior. Another type of risk mentioned in the interviews and prior work [5, 33] pertains to data protection and privacy issues. Here, the low effect on behavioral intention may be explained by the same kind of "privacy paradox" observed in several other contexts as well [16, 22]. In that sense, LLMs mirror behavioral trends observed in other technology domains, including smartphone overuse, password reuse, and reduced social media privacy concerns. This suggests that perceived benefits, such as efficiency and convenience, overrule potential downsides in their decision-making. Future research should explore targeted interventions, such as educational awareness programs or user-centered design improvements, to assess whether increasing risk awareness alters long-term usage patterns.

However, students also mentioned risks more directly related to their study and learning outcomes, particularly the risk of getting caught when handing in an LLM-generated solution to a task or the concern of impoverishing one's learning process. Those can be framed in terms of motivational goal orientation [12, 14]. Cheating is indicative of performance goals (i.e., seeking good/avoiding bad grades) or even mastery avoidance goals (i.e., dodging a challenge), whereas trying to enhance learning or at least being concerned about not diminishing it is indicative of a mastery approach goal (pursuing learning/seeking challenges). Consequently, what counts as a risk (e.g., getting caught, failing an exam, learning less) vis-à-vis successful performance (e.g., undetected cheating, improved learning) likely depends on individual motivations and attitudes towards learning.

While goal orientation has a long-standing history in research on learner motivation, we are unaware of any research utilizing motivational theories of learning to investigate the use of LLM-based services in education. Nevertheless, we would hypothesize that a student's goal orientations concerning their studies are predictive of the kind of LLM usage they would endorse and practice in that context and the associated benefits and risks. In all likelihood, the performance quality of LLMs will keep improving. Hallucinated, outdated, or inappropriate results keep decreasing with every new model iteration [2, 34], providing ever-easier opportunities to cheat without getting caught. Yet learning happens through cognitive engagement and is diminished if said engagement is offloaded or bypassed. Learning cannot be delegated.

Regarding regulation, this makes it rather difficult to draw a general and clear line between using LLMs in a way that undercuts learning by diminishing relevant cognitive engagement and using them in a way that fosters learning through scaffolding or by focusing the process. How much effort has to be put into a solution attempt until it becomes a legitimate input for LLM-generated feedback? Does generating an abstract solution strategy constitute legitimate scaffolding or an undesirable reduction in task difficulty? Such questions critically depend on the particular task at hand and its particular educational objective. This certainly challenges educators and policymakers who aim to establish consistent rules regarding using LLM-based services in education. Nevertheless, our findings suggest that many students favor precise regulation only to have a reliable guideline about what kind of behavior they might get in trouble for.

Consequently, our findings contribute to ongoing discussions within the computer-supported cooperative work community around the socio-technical dynamics of emerging technologies. In particular, the disconnect between high perceived risk and continued use of LLMs echoes prior work on the privacy paradox in social media and collaborative platforms [16, 22], where user actions diverge from expressed concerns. While students recognized various risks associated with LLMs—including issues of academic integrity, misinformation, and privacy, these concerns do not appear to significantly deter their actual usage. This behavioral paradox mirrors patterns observed in other domains of technology adoption, such as privacy and ethical concerns. Similar tensions have been documented by Wang et al. [50], who found that AI prototypers benefit from in-situ AI tools to better anticipate harms during early design stages, yet such tools are rarely used without deliberate intervention due to limited awareness or motivational barriers. Moreover, Long et al. [30] showed that users of generative AI workflows tend to appropriate and customize tools over time, increasing their perceived utility despite initial concerns or novelty-related skepticism or the novelty effect of AI tools [21, 25, 48]. These insights support our finding that students, after an initial familiarization phase, rationalize or integrate LLM use into their study routines even when risk perceptions remain high. However, our study suggests that this gap is not merely behavioral but tied to deeper motivational tensions: students weigh risks against competing goals such as efficiency, performance, and time pressure. This perspective invites computer-supported cooperative work researchers to move beyond rationalist models of acceptance and engage with motivational theories, such as goal orientation and learned helplessness, as lenses to explain this behavioral dissonance. By surfacing these frictions, our work extends risk-related discourse in computer-supported cooperative work and contributes towards an understanding of how LLMs are being normalized in everyday academic practice.

6.1 Limitations and Future Work

Although our study employs a mixed-methods approach to consider quantitative and qualitative perspectives, we acknowledge several limitations. Due to how the survey participants were recruited, our sample cannot be assumed to be representative of the population of university students recruited through our network. It exhibits certain biases, likely introduced via the distribution networks available to the researchers and evidenced, for instance, by the disproportionately high number of computer science and maths students among the survey respondents. One criticism of the qualitative interviews might be the low number of participants, which comprised eight persons. To counteract this limitation, we aimed to sample a broad range of participants, thereby also deliberately trying to mitigate the quantitative bias mentioned above. While a larger sample size is always preferable, the combination of quantitative and qualitative sampling and analysis employed in the present study also mitigates some of the limitations of either approach. To mitigate the issue of limited representativeness in future work, we will expand our sampling to include a more diverse and stratified participant base across disciplines, institutions, and educational systems. Additionally, incorporating cross-cultural data could further clarify how risk perception interacts with students' individual backgrounds and cultural norms.

Furthermore, the way "risk" was framed in our study may have influenced participants' responses by placing more emphasis on concerns than advantages. While the *Risk Expectancy* construct captured generalized concern, interviews revealed that students interpret risk in diverse and situational ways. This suggests the need for more differentiated risk constructs in future studies, distinguishing between systemic, behavioral, and pedagogical risks to better understand how each uniquely affects user behavior. Future research should consider a framing to explore how both perceived risks and benefits shape behavior. Moreover, as LLM tools become more embedded in educational settings, students' perceptions and usage patterns are likely to change. A longitudinal study that tracks changes in attitudes, mental models, and behavioral responses over time through repeated surveys and interviews would provide deeper insights into these dynamics. Such a study could also incorporate risk-awareness interventions, including AI literacy training, transparency-enhancing design features, or peer-led discussions, to assess their effectiveness in fostering more reflective and responsible use of LLM-based services.

Lastly, our findings strongly suggest that students' perceptions of associated benefits and risks depend on their individual goal orientations. Whether students pursue mastery or performance goals, that is, whether they are genuinely concerned with their own learning process or merely want to get past the next exam or term paper, likely has a significant effect on when and how they employ LLMs in their studies. Therefore, future research should investigate students' acceptance and use of LLM-based services while also controlling for goal orientation.

7 Conclusion and Outlook

Understanding and predicting students' adoption of LLM-based services in higher education is an important and very timely field of research. To that end, prior work has successfully applied established theories of the acceptance and use of technology (e.g., TAM or UTAUT2) and shown that performance and effort expectancy, among others, remain vital predictors. Our survey results are consistent with those findings and corroborate that potential risks, although frequently discussed by educators and policymakers and readily perceived by students, seem to have negligible effects on usage intention. Yet predictive models can only ever capture the status quo of people's behaviors, whereas educational endeavors are chiefly concerned with behavioral change. This is particularly apparent in the context of risk perception, which arguably *should* affect students' decisions to use LLM-based services or not. Here, our qualitative findings strongly suggest the need for a more differentiated view. Some of the risks students perceive, for example, hallucinated or biased output, may be conflated with lowered performance expectancy. At the same time, privacy concerns may be subject to the "privacy paradox" observed in many other contexts as well. When it comes to actual student learning, however, we propose that some of the most prominently discussed risks of LLM use in education, such as cheating, plagiarism, and impoverished learning outcomes, are better understood not through the lens of technology acceptance and use but through motivational goal orientation. Viewed from this perspective, students' motivation to truly master a subject or to merely demonstrate or, if necessary, fake performance can be expected to directly impact their attitudes and behavioral intentions towards LLM-based services for their studies. Further research is needed to corroborate this. Yet, it may contribute to a better understanding of how and why students use LLMs and ultimately inform effective institutional policy.

Acknowledgments

This work is supported by the German Research Foundation (DFG), CRC 1404: "FONDA: Foundations of Workflows for Large-Scale Scientific Data Analysis" (Project-ID 414984028).

References

[1] Ibrahim Adeshola and Adeola Praise Adepoju. 2023. The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments* 0, 0 (2023), 1–14. doi:10.1080/10494820.2023.2253858

[2] Garima Agrawal, Tharindu Kumarage, Zeyad Alghamdi, and Huan Liu. 2024. Can Knowledge Graphs Reduce Hallucinations in LLMs? : A Survey. arXiv:2311.07914 [cs.CL] https://arxiv.org/abs/2311.07914

[3] Sanne Akkerman, Wilfried Admiraal, Mieke Brekelmans, and Heinze Oost. 2008. Auditing Quality of Research in Social Sciences. *Quality & Quantity* 42, 2 (April 2008), 257–274. doi:10.1007/s11135-006-9044-4
[4] Mohammad Awad AlAfnan, Samira Dishari, Marina Jovic, and Koba Lomidze. 2023. ChatGPT as an Educational Tool: Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses. *Journal of Artificial Intelligence and Technology* 3, 2 (March 2023), 60–68. doi:10.37965/jait.2023.0184
[5] Hayder Albayati. 2024. Investigating undergraduate students' perceptions and awareness of using ChatGPT as a regular assistance tool: A user acceptance perspective study. *Computers and Education: Artificial Intelligence* 6 (June 2024), 100203. doi:10.1016/j.caeai.2024.100203
[6] David Baidoo-Anu and Leticia Owusu Ansah. 2023. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. doi:10.2139/ssrn.4337484
[7] Louis Cohen, Lawrence Manion, and Keith Morrison. 2007. *Research Methods in Education.* Routledge, London and New York.
[8] Joseph Crawford, Kelly-Ann Allen, Bianca Pani, and Michael Cowling. 2024. When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention. *Studies in Higher Education* 49, 5 (May 2024), 883–897. doi:10.1080/03075079.2024.2326956 Publisher: SRHE Website _eprint: https://doi.org/10.1080/03075079.2024.2326956.
[9] John W Creswell and Vicki L Plano Clark. 2017. *Designing and conducting mixed methods research.* Sage publications.
[10] Fred D. Davis. 1985. *A Technology Acceptance Model for Empirically Testing New End-User Information Systems : Theory and Results.* Thesis. Massachusetts Institute of Technology.
[11] Fred D Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13, 3 (Sept. 1989), 319. doi:10.2307/249008
[12] Carol S. Dweck and Ellen L. Leggett. 1988. A social-cognitive approach to motivation and personality. *Psychological Review* 95, 2 (1988), 256–273. doi:10.1037/0033-295X.95.2.256 Place: US Publisher: American Psychological Association.
[13] Yogesh K. Dwivedi, Nir Kshetri, Laurie Hughes, Emma Louise Slade, Anand Jeyaraj, Arpan Kumar Kar, Abdullah M. Baabdullah, Alex Koohang, Vishnupriya Raghavan, Manju Ahuja, Hanaa Albanna, Mousa Ahmad Albashrawi, Adil S. Al-Busaidi, Janarthanan Balakrishnan, Yves Barlette, Sriparna Basu, Indranil Bose, Laurence Brooks, Dimitrios Buhalis, Lemuria Carter, Soumyadeb Chowdhury, Tom Crick, Scott W. Cunningham, Gareth H. Davies, Robert M. Davison, Rahul Dé, Denis Dennehy, Yanqing Duan, Rameshwar Dubey, Rohita Dwivedi, John S. Edwards, Carlos Flavián, Robin Gauld, Varun Grover, Mei-Chih Hu, Marijn Janssen, Paul Jones, Iris Junglas, Sangeeta Khorana, Sascha Kraus, Kai R. Larsen, Paul Latreille, Sven Laumer, F. Tegwen Malik, Abbas Mardani, Marcello Mariani, Sunil Mithas, Emmanuel Mogaji, Jeretta Horn Nord, Siobhan O'Connor, Fevzi Okumus, Margherita Pagani, Neeraj Pandey, Savvas Papagiannidis, Ilias O. Pappas, Nishith Pathak, Jan Pries-Heje, Ramakrishnan Raman, Nripendra P. Rana, Sven-Volker Rehm, Samuel Ribeiro-Navarrete, Alexander Richter, Frantz Rowe, Suprateek Sarker, Bernd Carsten Stahl, Manoj Kumar Tiwari, Wil van der Aalst, Viswanath Venkatesh, Giampaolo Viglia, Michael Wade, Paul Walton, Jochen Wirtz, and Ryan Wright. 2023. Opinion Paper: "So What If ChatGPT Wrote It?" Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy. *International Journal of Information Management* 71 (Aug. 2023), 102642. doi:10.1016/j.ijinfomgt.2023.102642
[14] Andrew J Elliot. 2005. A conceptual history of the achievement goal construct. In *Handbook of competence and motivation.* The Guilford Press.
[15] Fauzi Fauzi, Laros Tuhuteru, Ferdinandus Sampe, Abu Muna Almaududi Ausat, and Heliza Rahmania Hatta. 2023. Analysing the Role of ChatGPT in Improving Student Productivity in Higher Education. *Journal on Education* 5, 4 (April 2023), 14886–14891. doi:10.31004/joe.v5i4.2563
[16] Nina Gerber, Paul Gerber, and Melanie Volkamer. 2018. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security* 77 (Aug. 2018), 226–261. doi:10.1016/j.cose.2018.04.002
[17] John Ingham, Jean Cadieux, and Abdelouahab Mekki Berrada. 2015. e-Shopping acceptance: A qualitative and meta-analytic review. *Information & Management* 52, 1 (Jan. 2015), 44–60. doi:10.1016/j.im.2014.10.002
[18] Shivani Kapania, Ruiyi Wang, Toby Jia-Jun Li, Tianshi Li, and Hong Shen. 2024. "I'm categorizing LLM as a productivity tool": Examining ethics of LLM use in HCI research practices. arXiv:2403.19876 [cs.HC] https://arxiv.org/abs/2403.19876
[19] Enkelejda Kasneci, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, Stepha Krusche, Gitta Kutyniok, Tilman Michaeli, Claudia Nerdel, Jürgen Pfeffer, Oleksandra Poquet, Michael Sailer, Albrecht Schmidt, Tina Seidel, Matthias Stadler, Jochen Weller, Jochen Kuhn, and Gjergji Kasneci. 2023. ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences* 103 (April 2023), 102274. doi:10.1016/j.lindif.2023.102274
[20] Mohammad Khalil and Erkan Er. 2023. Will ChatGPT Get You Caught? Rethinking of Plagiarism Detection. In *Learning and Collaboration Technologies*, Panayiotis Zaphiris and Andri Ioannou (Eds.). Springer Nature Switzerland, Cham, 475–487. doi:10.1007/978-3-031-34411-4_32

[21] Agnes M. Kloft, Robin Welsch, Thomas Kosch, and Steeven Villa. 2024. "AI enhances our performance, I have no doubt this one will do the same": The Placebo Effect Is Robust to Negative Descriptions of AI. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. ACM, New York, NY, USA. doi:10.1145/3613904.3642633

[22] Spyros Kokolakis. 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security* 64 (Jan. 2017), 122–134. doi:10.1016/j.cose.2015.07.002

[23] Thomas Kosch and Sebastian Feger. 2024. Risk or Chance? Large Language Models and Reproducibility in HCI Research. *Interactions* 31, 6 (Oct. 2024), 44–49. doi:10.1145/3695765

[24] Thomas Kosch and Sebastian Feger. 2025. Prompt-Hacking: The New p-Hacking? arXiv:2504.14571 [cs.HC] doi:10.48550/arXiv.2504.14571

[25] Thomas Kosch, Robin Welsch, Lewis Chuang, and Albrecht Schmidt. 2023. The Placebo Effect of Artificial Intelligence in Human–Computer Interaction. *ACM Trans. Comput.-Hum. Interact.* 29, 6, Article 56 (Jan. 2023), 32 pages. doi:10.1145/3529225

[26] Udo Kuckartz. 2012. *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*. Beltz Juventa.

[27] Chinmay Kulkarni, Tongshuang Wu, Kenneth Holstein, Q. Vera Liao, Min Kyung Lee, Mina Lee, and Hariharan Subramonyam. 2023. LLMs and the Infrastructure of CSCW. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing* (Minneapolis, MN, USA) *(CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 408–410. doi:10.1145/3584931.3608438

[28] Yiheng Liu, Tianle Han, Siyuan Ma, Jiayue Zhang, Yuanyuan Yang, Jiaming Tian, Hao He, Antong Li, Mengshen He, Zhengliang Liu, Zihao Wu, Lin Zhao, Dajiang Zhu, Xiang Li, Ning Qiang, Dingang Shen, Tianming Liu, and Bao Ge. 2023. Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology* 1, 2 (Sept. 2023), 100017. doi:10.1016/j.metrad.2023.100017

[29] Chung Kwan Lo. 2023. What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences* 13, 4 (April 2023), 410. doi:10.3390/educsci13040410 Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[30] Tao Long, Katy Ilonka Gero, and Lydia B Chilton. 2024. Not Just Novelty: A Longitudinal Study on Utility and Customization of an AI Workflow. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 782–803. doi:10.1145/3643834.3661587

[31] Brady D. Lund and Ting Wang. 2023. Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries? *Library Hi Tech News* 40, 3 (Jan. 2023), 26–29. doi:10.1108/LHTN-01-2023-0009

[32] Nikola Marangunić and Andrina Granić. 2015. Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society* 14, 1 (March 2015), 81–95. doi:10.1007/s10209-014-0348-1

[33] Devadas Menon and K. Shilpa. 2023. "Chatting with ChatGPT": Analyzing the factors influencing users' intention to Use the Open AI's ChatGPT using the UTAUT model. *Heliyon* 9, 11 (Nov. 2023). doi:10.1016/j.heliyon.2023.e20962 Publisher: Elsevier.

[34] Gabrijela Perković, Antun Drobnjak, and Ivica Botički. 2024. Hallucinations in LLMs: Understanding and Addressing Challenges. In *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*. 2084–2088. doi:10.1109/MIPRO60963.2024.10569238

[35] Partha Pratim Ray. 2023. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems* 3 (Jan. 2023), 121–154. doi:10.1016/j.iotcps.2023.04.003

[36] Mahefa Abel Razafinirina, William Germain Dimbisoa, and Thomas Mahatody. 2024. Pedagogical alignment of large language models (llm) for personalized learning: a survey, trends and challenges. *Journal of Intelligent Learning Systems and Applications* 16, 4 (2024), 448–480. doi:10.4236/jilsa.2024.164023

[37] José-María Romero-Rodríguez, María-Soledad Ramírez-Montoya, Mariana Buenestado-Fernández, and Fernando Lara-Lara. 2023. Use of ChatGPT at University as a Tool for Complex Thinking: Students' Perceived Usefulness. *Journal of New Approaches in Educational Research* 12, 2 (July 2023), 323–339. doi:10.7821/naer.2023.7.1458 Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 2 Publisher: Springer International Publishing.

[38] Jürgen Rudolph, Samson Tan, and Shannon Tan. 2023. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching* 6, 1 (Jan. 2023), 342–363. doi:10.37074/jalt.2023.6.1.9

[39] Artur Strzelecki. 2023. To Use or Not to Use ChatGPT in Higher Education? A Study of Students' Acceptance and Use of Technology. *Interactive Learning Environments* 0, 0 (May 2023), 1–14. doi:10.1080/10494820.2023.2209881

[40] Artur Strzelecki. 2024. Students' Acceptance of ChatGPT in Higher Education: An Extended Unified Theory of Acceptance and Use of Technology. *Innovative Higher Education* 49, 2 (April 2024), 223–245. doi:10.1007/s10755-023-09686-1

[41] Keith S Taber. 2018. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education* 48 (2018), 1273–1296. doi:10.1007/s11165-016-9602-2

[42] Mohsen Tavakol and Reg Dennick. 2011. Making sense of Cronbach's alpha. *International journal of medical education* 2 (2011), 53. doi:10.5116/ijme.4dfb.8dfd

[43] Weiqi Tian, Jingshen Ge, Yu Zhao, and Xu Zheng. 2024. AI Chatbots in Chinese higher education: adoption, perception, and influence among graduate students—an integrated analysis utilizing UTAUT and ECM models. *Frontiers in Psychology* 15 (Feb. 2024). doi:10.3389/fpsyg.2024.1268549 Publisher: Frontiers.

[44] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 27, 3 (2003), 425–478. doi:10.2307/30036540 Publisher: Management Information Systems Research Center, University of Minnesota.

[45] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 27, 3 (2003), 425–478. http://www.jstor.org/stable/30036540

[46] Viswanath Venkatesh, James Y. L. Thong, and Xin Xu. 2012. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly* 36, 1 (2012), 157–178. doi:10.2307/41410412 Publisher: Management Information Systems Research Center, University of Minnesota.

[47] Viswanath Venkatesh, James Y. L. Thong, and Xin Xu. 2016. Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead. https://papers.ssrn.com/abstract=2800121

[48] Steeven Villa, Thomas Kosch, Felix Grelka, Albrecht Schmidt, and Robin Welsch. 2023. The placebo effect of human augmentation: Anticipating cognitive augmentation increases risk-taking behavior. *Computers in Human Behavior* 146 (2023), 107787. doi:10.1016/j.chb.2023.107787

[49] Jörg von Garrel, Jana Mayer, and Markus Mühlfeld. 2023. Künstliche Intelligenz im Studium Eine quantitative Befragung von Studierenden zur Nutzung von ChatGPT & Co. (June 2023). doi:10.48444/h_docs-pub-395

[50] Zijie J. Wang, Chinmay Kulkarni, Lauren Wilcox, Michael Terry, and Michael Madaio. 2024. Farsight: Fostering Responsible AI Awareness During AI Application Prototyping. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 976, 40 pages. doi:10.1145/3613904.3642335

Received October 2024; revised April 2025; accepted August 2025