Generative AI in Higher Education: Adoption and Impact of Custom GPT Chatbots on Student Learning in Management Accounting and Finance

Prof. Dr. Isabel Lausberg

Ruhr West University of Applied Sciences, Mülheim an der Ruhr
Isabel.Lausberg@hs-ruhrwest.de

Prof. Dr. Alexander Bönner

Ruhr West University of Applied Sciences, Mülheim an der Ruhr
Alexander.Boenner@hs-ruhrwest.de

Timur Gümüs

Ruhr West University of Applied Sciences, Mülheim an der Ruhr
Timur.Guemues@hs-ruhrwest.de

Abstract: Generative AI has become an integral part of students' everyday learning. A majority of students already use commercial chatbots as support tools. Against this backdrop, Horus and FinPro were created, two Custom GPT Chatbots for teaching support in the subjects of Management Accounting and Finance at Hochschule Ruhr West (Germany). They are designed to direct students to quality-assured bots that have been specially trained in the respective topics. Initial findings show that the bots are well received – two-thirds of students now use them regularly despite already being accustomed to using commercially available chatbots. Furthermore, Horus and FinPro are utilised in a variety of ways for learning purposes, with traditional features such as answering questions being employed more frequently than options like generating exercises. In our study, the use of generative AI has surpassed classic internet research and the use of YouTube videos. However, students are also aware that the chatbot's answers cannot be trusted completely due to hallucinations and incorrect source references. Supporting students in developing a critical and reflective approach to generative AI should therefore be understood as a key responsibility of higher education teaching.

Keywords: Custom GPT Chatbots, AI-assisted learning, Student Learning Behavior, Critical Thinking, Management Accounting, Finance

1 Introduction

As a key technology of digital transformation, generative artificial intelligence (GenAI) is increasingly reshaping how teaching and learning are designed and experienced in higher education. Recent review studies show that Large Language Models (LLMs) are now being applied across a wide range of academic contexts, including tutoring, formative feedback, assessment support, content generation, and student assistance. At the same time, the literature emphasizes that their educational value is not uniform: while some studies report gains in engagement, accessibility, and learning support, others highlight important limitations, including over-reliance, uneven effectiveness across disciplines, and unresolved questions concerning pedagogical impact [1]–[3].

This debate takes place against the backdrop that GenAI has become firmly embedded in students' everyday study practices. Recent evidence from Germany shows that AI is already used frequently in higher education: 65% of surveyed students reported using AI weekly or daily in their studies [6]. The most common purposes include general research and gaining an overview of a topic, brainstorming, using AI as a learning partner or tutor, and translation [6, 7]. At the same time, usage patterns differ across disciplines, indicating that AI adoption is now widespread but still shaped by subject-specific learning cultures [6]. Beyond the German context, current international evidence likewise points to very high levels of adoption, with more than 90% of respondents in a recent university study reporting that they use AI to support academic activities [8]. Together, these findings suggest that the use of GenAI is no longer a marginal phenomenon, but a structural feature of contemporary student learning [6], [8], [9].

At the same time, recent studies indicate that institutional support has not yet fully caught up with students' rapid uptake of these tools. In the German CHE survey, students rated their university's existing provision for acquiring AI-related competencies at only 2.7 out of 5 on average [6]. They also expressed strong expectations regarding clearer ethical guidelines, data protection, and more systematic learning opportunities for AI literacy [6]. This suggests that higher education institutions are under growing pressure not only to regulate AI use, but also to integrate it pedagogically in ways that are transparent, discipline-sensitive, and educationally meaningful [6]. Closely linked to this is the question of whether GenAI is ultimately helpful or harmful for student learning [10, 11]. Research on self-regulated learning (SRL), understood as learners' active regulation of goals, strategies, motivation, and reflection throughout the learning process, shows that students vary considerably in how strategically and critically they use AI tools:

some employ AI in ways that support deeper engagement and metacognitive regulation, whereas others rely on it more superficially or uncritically [12, 13, 14]. Current evidence therefore points less to a simple positive-or-negative assessment than to the importance of usage quality, strategic awareness, and pedagogical framing [12]. In particular, recent work argues that higher education should foster students' metacognitive and ethical awareness in AI-supported learning and cultivate forms of use that promote reflection rather than shortcut thinking [12, 13]. A systematic review of AI applications in supporting SRL confirms that while AI tools can facilitate all three phases of the SRL cycle – forethought, performance, and self-reflection – their impact remains uneven across phases, with the performance phase receiving the most support. The review further cautions that over-reliance on AI-generated guidance can undermine learners' autonomous regulation and self-efficacy if not accompanied by deliberate pedagogical framing [14]. Overall, reviews of ChatGPT and related systems conclude that GenAI can provide substantial benefits for feedback, explanation, and academic support, but only when embedded in clearly structured educational settings and accompanied by critical guidance [1]. This underlines the need for institutionally supported, course-aligned AI applications that do not merely offer technical convenience, but contribute to academically sound and pedagogically responsible learning processes [12], [1], [13].

Within this broader landscape, virtual assistants based on LLMs have emerged as a particularly relevant application area for course-aligned learning support. Current evidence suggests that such systems can support individualized learning by providing immediate explanations, adaptive responses, and dialogue-based assistance tailored to learners' needs [1], [2], [4]. Importantly, the educational potential of these assistants increases when they are not deployed as generic chatbots, but are embedded in course-specific environments and linked to curated instructional materials [1], [5]. A key enabling concept in this context is Retrieval-Augmented Generation (RAG), which improves factual grounding and contextual relevance by connecting the model to an external knowledge base [5]. Accordingly, RAG is widely regarded as a promising architecture for educational applications, especially when assistants are grounded in course materials rather than relying solely on the model's general knowledge [4], [5]. At the same time, the literature identifies important limitations, including hallucinations, incomplete retrieval, outdated source material, and limited multimodal capabilities [5, 15]. These challenges are particularly relevant in higher education, where students require not only quick answers, but also reliable, transparent, and pedagogically appropriate support [1], [2], [4].

Empirical work on virtual teaching assistants further points to promising results under these conditions. A large-scale real-world evaluation of an LLM-based virtual teaching assistant found that such systems can provide instant support at scale, facilitate multi-turn interaction, and complement traditional instructor–student communication patterns [4]. However, student acceptance and educational value depended strongly on perceived usefulness, response quality, and the extent to which the assistant was aligned with the learning context [4], underscoring that success is shaped not only by technical performance, but also by didactic design and institutional integration [1], [4]. A prominent example of a virtual teaching assistant ecosystem is ETHEL, developed at ETH Zurich. ETHEL has been used across multiple semesters and comprises several production-ready tools, including a reference-based chatbot, homework feedback, and short polling functions [16]. Further research shows that ETHEL serves as a platform for extending conversational support into additional learning functions, such as the generation of validated practice problems on demand [17]. In parallel, work on AI-supported grading in physics demonstrates that LLM-based assessment systems can reach a comparatively high level of agreement with human grading while still requiring human oversight, especially in high-stakes contexts [18]. Taken together, these findings indicate that virtual assistants can provide pedagogically valuable support when instructional control remains with the instructor and the AI system is confined to clearly defined educational functions [16]–[18]. In summary, recent research suggests that GenAI-based virtual assistants can become a meaningful component of higher education teaching when they are grounded in course materials, embedded in clear pedagogical structures, and accompanied by ongoing human oversight – a principle that appears highly transferable to other applied subject areas, including management accounting and finance [1], [5], [16].

Against this background, the present paper focuses on virtual learning assistants in the fields of management accounting and finance. The aim is to show how AI-supported assistants can provide academically sound, evidence-based, and didactically meaningful support when they are embedded in a clearly defined course context and constrained by an approved source base. More specifically, the paper presents the conception, implementation, and evaluation of the two pilot projects HORUS and FinPro at Ruhr West University of Applied Sciences. Both assistants were designed to support students in a dialogue-based manner, guide them in working with subject-specific materials, and provide structured assistance in understanding concepts and solving course-related tasks. This paper therefore contributes to the current debate on how GenAI can be integrated into higher education not merely as a general-purpose tool, but as a course-aligned learning support system with clear pedagogical guidance and boundaries [1], [4], [5].

2 Methods

2.1 Conceptual Design of HORUS and FinPro

HORUS and FinPro are so-called custom GPTs, that is, specially configured versions of ChatGPT that can be tailored to particular target groups or tasks without requiring programming skills [19], [20], [21]. In this process, instructors define the assistant's behavior through simple natural-language instructions ("custom instructions") and upload course-relevant materials such as PDF documents, which the system can draw upon when answering questions [20]. The objective is to provide learners with academically sound and comprehensible support, with a particular emphasis on reliable information provision and transparent source citation. This makes it possible to control the model's behavior precisely with regard to language style, technical terminology, and response strategy [20].

Technically, HORUS and FinPro are – like ETHEL – based on a Retrieval-Augmented Generation (RAG) approach that enables answers to be generated exclusively on the basis of predefined sources [5]. The materials used stem from the Management Accounting and Finance modules at Ruhr West University of Applied Sciences, which have been taught in a flipped-classroom format since 2020. A broad range of documents is therefore available, including lecture notes, exercise materials with solutions, as well as instructional videos and recorded lectures by the instructors. For HORUS and FinPro, these audiovisual materials were transcribed with AI support using the OpenAI tool Whisper and stored as structured text in PDF format. HORUS and FinPro answer only those questions that can clearly be assigned to the subject areas of Management Accounting and Finance. They are instructed not to draw on their own inferences or on sources outside the predefined set. In addition, HORUS uses selected authorized external sources, such as the Controlling-Wiki, the ControllingPortal, and specialist articles that are likewise made available in the Moodle course environment. Independent inferences or external data are explicitly excluded in order to ensure content reliability.

Unlike ETHEL, no university-owned server infrastructure is available. The system therefore runs through a standard ChatGPT Plus account. Students access HORUS and FinPro via a link provided in the Moodle course environment.

The custom GPT instructs HORUS and FinPro to include a source reference with every answer, either by indicating the name of the PDF used, including the relevant page number, or by providing the URL of one of the approved websites. The PDFs constitute the primary source of information. If no relevant information can be found there, an additional search of the authorized websites is carried out automatically and without any further action required on the part of the users.

If a question can be answered neither from the PDFs nor from the approved websites, or if it does not fall within the thematic scope of Management Accounting or Finance, the GPT is instructed to inform students politely. According to its configuration, responses are factual, precise, and easy to understand, and they are intended to support clear and structured knowledge transfer for students. Complex content is simplified where necessary. The consistency of source references and response structure, both in terms of content and form, is defined as mandatory within the custom GPT configuration [20].

The system serves several didactic and technical purposes. It answers subject-specific questions automatically on the basis of the provided materials, generates standardized source references including page numbers, for example "script page: X; source: Y," and restricts the scope of answers to course-related content. New documents can be added easily and integrated systematically, with each source being referenced automatically [20]. In addition, HORUS and FinPro are capable of generating practice tasks and even reviewing students' handwritten solutions. HORUS can furthermore be extended with functions such as web search, data analysis, or code interpretation, which offers particular advantages for academic advising, research-oriented work, or data-based analytical tasks [20], [22]. To illustrate Fig. 1 shows the start page of HORUS.

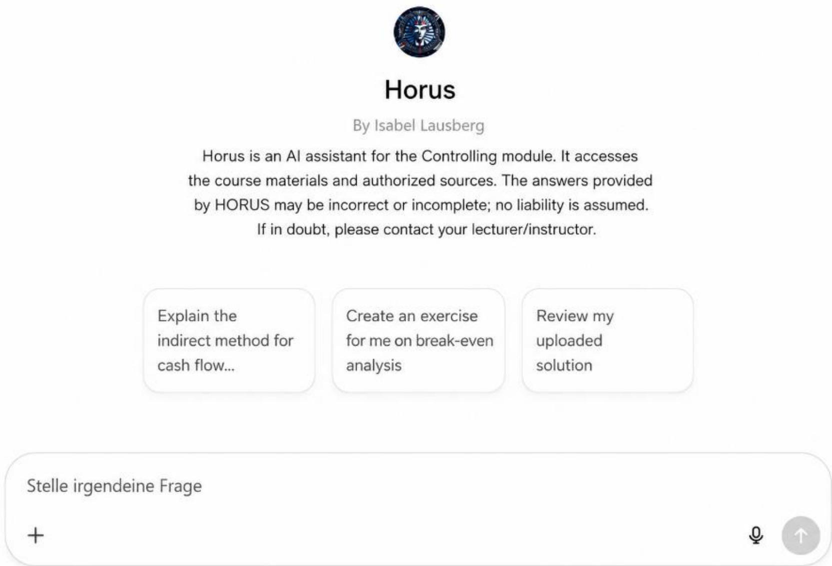

Figure 1
Start page of HORUS, Source: own edition

Similar in design to HORUS, Fig. 2 shows the FinPro start page. Here, too, following a brief introduction to the learning assistant, there are three suggested starting questions to help users get started.

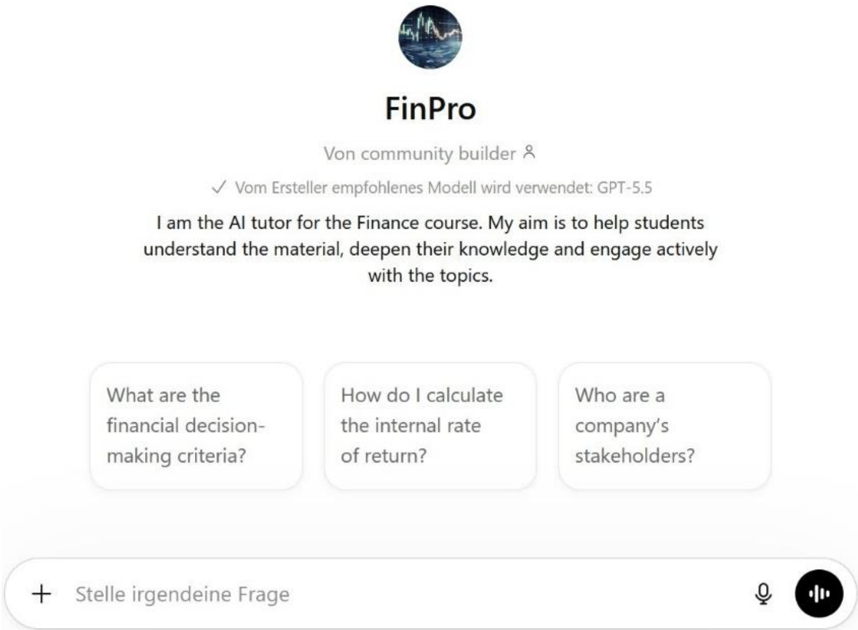

Figure 2
Start page of FinPro, Source: own edition

During development, extensive tests were conducted in order to examine the functionality and reliability of HORUS and FinPro. These included verifying correct citation behavior, testing responses to questions outside the disciplinary scope, integrating new documents, and analyzing potential sources of error. During the test phase, several PDFs were used in order to broaden the information base and simulate typical usage scenarios. Additional functions such as code interpretation and internet search were activated selectively in order to represent more complex use cases in the field of management accounting.

Overall, HORUS and FinPro represent specifically tailored AI applications which, through their combination of clearly defined content boundaries, technical extensibility, and didactic controllability, provide effective support for teaching and learning in the fields of management accounting and finance.

2.2 Survey Design and Analysis

After HORUS and FinPro had been introduced to the students as part of their respective courses, a survey was designed to evaluate the actual frequency of use, patterns of use, and students' perceived reliability of the virtual assistants HORUS and FinPro. The survey period extended from 22 January to 10 April 2026 and comprised 29 main questions, in addition to items on basic demographic and study-related characteristics such as age, gender, and number of examination attempts. These background variables, however, are not central to the findings reported here. The survey sample comprised n = 127 students enrolled in the Management Accounting and Finance courses.

3 Results

One of the key questions in the survey was: "How often do participants use AI tools for learning purposes?" As noted in the introduction, recent studies indicate that the huge majority of students in Germany and abroad use AI applications [6], [7]. The current survey on HORUS and FinPro shows an even higher level of adoption: 98% of participants stated that they use AI tools for learning purposes.

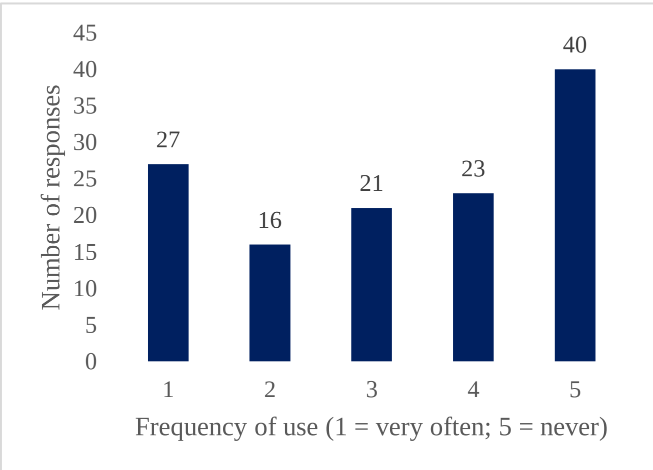

Figure 3
Frequency of use of the virtual assistans HORUS and FinPro (n = 127), Source: own data and edition

In addition, 69% reported that they use the course-specific AI assistants. As Fig. 3 depicts, the frequency distribution among those who use the module chatbot is relatively even. Around 21% of students (27 participants) report that they use the respective module chatbot "very often." By contrast, 32% of students state that they "never" use the module chatbot. The most frequently cited reasons for non-use are, above all, the use of a general-purpose chatbot (17 responses) and "I have not gotten around to it yet" (7 responses).

Another survey item, illustrated in Fig. 4, aimed to examine how important the use of the course-specific AI assistants is in comparison with other learning resources ("How central is the use of generative AI compared with traditional learning aids and tools?"). The results indicate a moderate level of importance for the AI assistants HORUS and FinPro. On average, the most frequently used learning resources are the exercise script and the instructional videos available in the Moodle course environment (vodcasts / recorded lectures). General-purpose chatbots are used more frequently than the chatbot assigned to the respective module (HORUS or FinPro). However, it is remarkable that the use of generative AI has already surpassed classic internet research and the use of YouTube videos.

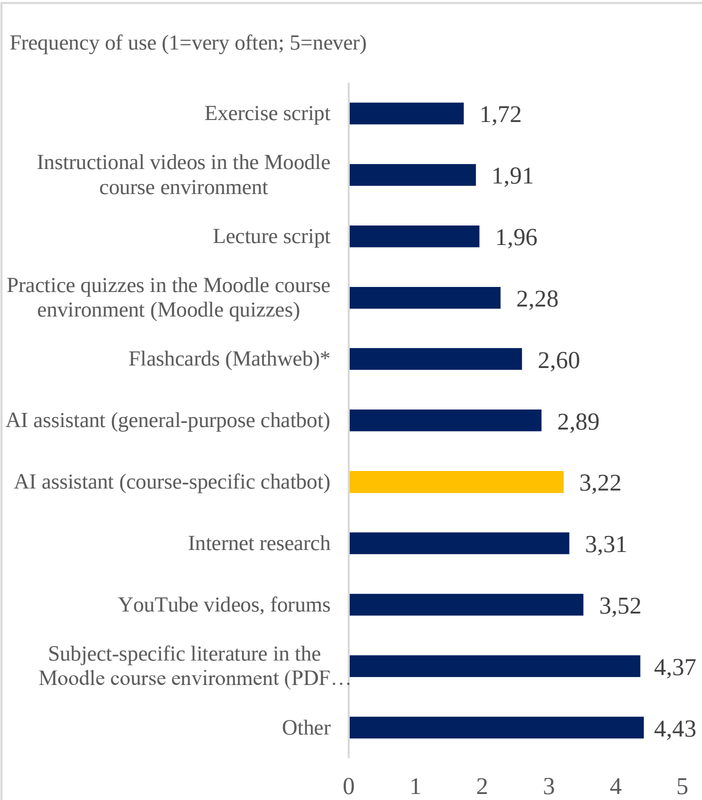

Figure 4

Use of different learning materials (arithmetic mean) (n = 127), Source: own data and edition

The least frequently used category was other learning materials (4.43). This category included, for example, mock exams, support from students in higher semesters, self-prepared summaries based on the available learning materials, and practice exercises from additional sources. As no predefined response options were provided for this category, its lower weight is likely due at least in part to this open format. In substantive terms, however, some of these resources – particularly mock exams – are likely to be of greater importance than the overall mean value may suggest.

Furthermore, the survey examined the purposes for which the course-specific chatbots were used. For this purpose, three categories were distinguished: clarification of subject-specific questions, review of students' own solutions, and generation of additional practice tasks. The results, which can be seen in Fig. 5, show that the classic question-and-answer function is used most frequently, although it is followed closely by the upload of students' own solutions combined with requests for feedback or correction. The generation of additional practice tasks is used somewhat less frequently.

Twenty-two students reported that they had also used the course-specific chatbot for other application areas. These included using it as an active sparring partner that quizzes students on specific topics, support in solving randomized online math problems (Mathweb), as well as providing explanations and summaries ("flashcards").

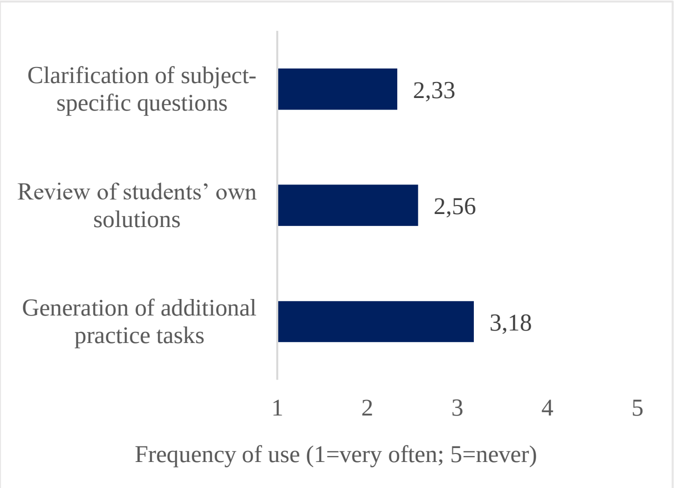

Figure 5
Frequency of use of the course-specific chatbot by application area (arithmetic mean) (n = 87)[2],
Source: own data and edition

2 The following figures include only responses of participants who had previously indicated that they use the respective course-specific chatbot.

While 63 out of 83 students stated that they did not wish for any additional chatbot functions, 20 respondents indicated that they would welcome further features. These include the generation of additional learning materials by the respective chatbot, for example in the form of flashcards and podcasts, the expansion of the source base to include, for instance, subject-specific literature, and the generation of tasks at examination level.

After examining frequency of use, intensity of use, and patterns of use, a further section of the survey addressed the extent to which the responses provided by the respective chatbot were perceived as reliable. As can be seen in Fig. 6, 16 students resp. 19% of participants answered the question “How often do you perceive the chatbot’s statements to be incorrect?” with “often” or “very often.”

According to the students, the most frequently reported errors concerned incorrect calculation results in quantitative tasks. Respondents also noted that the chatbot sometimes had difficulties reading and reproducing formulas correctly and that its outputs did not always correspond exactly to the lecture script. More broadly, RAG-based systems have been shown to improve factual grounding through external evidence retrieval, yet residual hallucination risks persist even in well-configured architectures [15]. This is consistent with recent findings suggesting that chatbots tend to process HTML-based content more reliably than PDF documents and are more prone to errors when working with the latter [23].

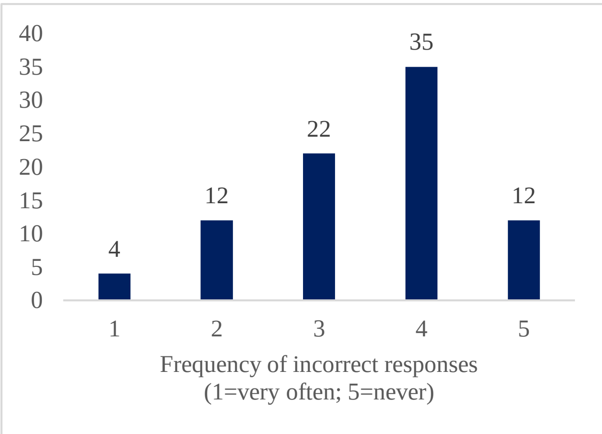

Figure 6:
In your perception, how often are the chatbot’s responses incorrect? (n = 85), Source: own data and edition

According to 66 out of 78 students, the respective chatbot generally refers to the correct source or sources. Fig. 7 below illustrates how often, in the students’ assessment, the course-specific chatbot cites sources incorrectly. Only the responses of participants who answered "No" to the question "Are sources cited correctly?" were taken into account (12 out of 78). Only one respondent stated that the chatbot "very frequently" provides incorrect source references.

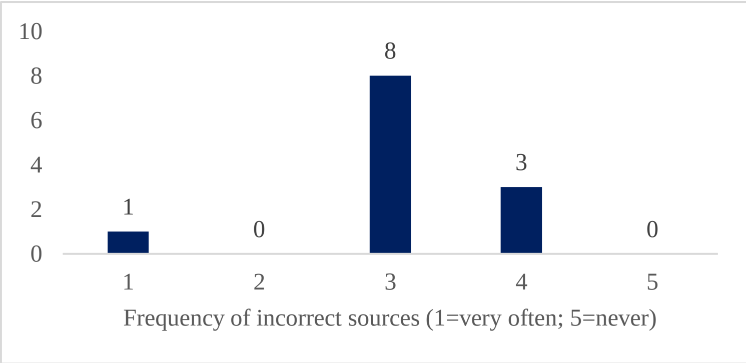

Figure 7
How often are the cited sources incorrect? (n = 12), Source: own data and edition

The majority of survey participants (73%) stated that, in their perception, the respective chatbot does not tend to hallucinate. Nevertheless, as shown in Fig. 8, 21 out of 79 students answered "Yes" to the question: "In your perception, does the chatbot tend to hallucinate, that is, to invent information not contained in the sources?"

Combined with the findings in Fig. 7 it shows that students are aware that the chatbot's answers cannot be trusted completely.

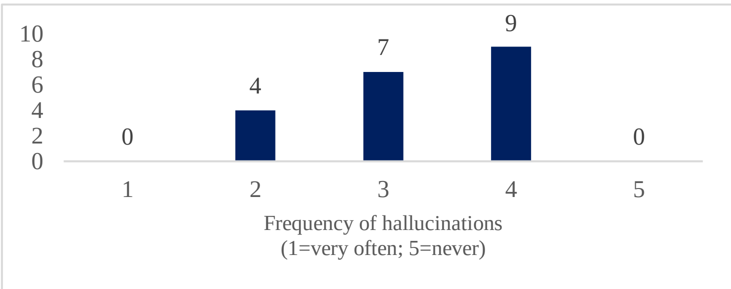

Figure 8
Frequency of hallucinations (n = 20)[3], Source: own data and edition

3 Figure 8 includes only the responses of those participants who answered "Yes." One response was excluded because, despite answering "Yes," the respondent selected the category "never" (5), resulting in a final subsample of n = 20.

The students’ self-efficacy in relation to chatbots was assessed on the basis of three predefined statements. On average, students feel confident using AI assistants, see added value in their learning, and trust them, although the latter point is only valid to a limited extent as can be derived from Fig. 9.

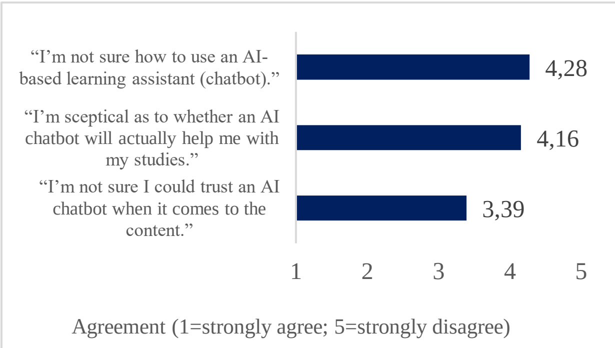

Figure 9

Self-efficacy and barriers to learning (arithmetic mean) (n = 127), Source: own data and edition

Discussions

This paper presented the design and initial evaluation of HORUS and FinPro as two course-specific, RAG-based virtual learning assistants for management accounting and finance. The findings suggest that generative AI has become a structural element of students’ study practices. The survey results show both very high overall adoption of AI tools and substantial uptake of the course-specific assistants.

The overall AI adoption rate of 98% notably exceeds levels reported in comparable surveys [6]. This difference may partly reflect the effect of explicit course-level AI integration. Students primarily used HORUS and FinPro for subject-specific questions and feedback on their own solutions, whereas the generation of additional practice tasks played a secondary role. Although students perceived the assistants as useful and the source references as mostly correct, reported errors in calculations, formula handling, and occasional hallucinations indicate that course grounding improves reliability but does not fully guarantee it.

Overall, the findings suggest that course-specific virtual assistants can provide scalable and pedagogically meaningful learning support when embedded in a clearly defined course context and accompanied by explicit guidance on verification, critical evaluation, and reflective use. At the same time, the study is limited by its reliance on self-reported perceptions from a single institutional setting and by the absence of direct measures of actual learning outcomes. Initial longitudinal evidence from the German higher education context suggests that the relationship between AI tool use and domain-specific reasoning performance is complex and requires careful empirical examination [24]. Future research should therefore examine how course-specific AI assistants can be didactically integrated in meaningful ways and how they affect SRL and learning performance.

Building on these results, the next step for HORUS and FinPro is to scale the approach by transitioning from a pilot setting to a university-wide implementation. This requires not only technical expansion but also organizational alignment to ensure consistent quality, governance, and sustainable adoption across faculties and degree programs. In parallel, continued tool development is needed to enhance system effectiveness and user experience. Key priorities include greater personalization, improved reliability through refined custom instructions, reduced errors, incorrect citations, and hallucinations, as well as the testing and integration of additional AI tools that extend functionality.

Beyond scaling and robustness, a central objective is to integrate the virtual learning assistants more deeply into the teaching and learning process. Rather than remaining supplemental tools, they should be embedded more systematically into course design so that their use becomes pedagogically purposeful. In particular, they can support individualized and self-regulated learning by providing adaptive guidance, timely feedback, and structured learning prompts, such as goal-setting prompts, progress monitoring, and reflection cues, that foster planning, monitoring, and reflection [25]. This is especially relevant because current evidence suggests that AI tools primarily support the performance phase of the SRL cycle, while their contribution to planning and self-reflection remains comparatively underdeveloped [14]. Accordingly, future development and research should explicitly examine whether and how course-specific assistants can strengthen these under-supported SRL phases through the implementation and evaluation of SRL-oriented design features.

At the same time, integration should explicitly promote AI literacy by making students' interactions with AI an object of learning. This can be achieved, for example, through guided prompt practices, critical evaluation of AI outputs, and transparency about limitations, bias, and appropriate academic use. Initial approaches to implementing such activities with groups of students have already been tested and appear promising. Taken together, responsible institutional scaling, deeper didactic embedding, SRL-oriented design, AI literacy integration, and continuous technical improvement can help increase both the educational impact and the scalability of course-specific virtual learning assistants.

References

[1] Qian Liu, Anjin Hu, Tehmina Gladman, Steve Gallagher: Eight Months into Reality: A Scoping Review of the Application of ChatGPT in Higher Education Teaching and Learning, Innovative Higher Education, vol. 50, 2025, pp. 1677-1700.

[2] Yuhong Shi, Kun Yu, Yifei Dong, Fang Chen: Large language models in education: a systematic review of empirical applications, benefits, and challenges, Computers and Education: Artificial Intelligence, vol. 10, 2026, Art. 100529.

[3] Kellie (Kai Lai) Hon: Generative AI in Higher Education: A Systematic Review of Its Effects on Learning Outcomes and Academic Performance, Journal of Educational Technology Systems, vol. 54, no. 3, 2026, pp. 537-560.

[4] Sunjun Kweon, Sooyohn Nam, Hyunseung Lim, Hwajung Hong, Edward Choi: A Large-Scale Real-World Evaluation of an LLM-Based Virtual Teaching Assistant, in Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 6: Industry Track), Vienna, Austria, July 2025, pp. 850-864.

[5] Zongxi Li, Zijian Wang, Weiming Wang, Kevin Hung, Haoran Xie, Fu Lee Wang: Retrieval-augmented generation for educational application: A systematic survey, Computers and Education: Artificial Intelligence, vol. 8, 2025, Art. 100417.

[6] M. Hüsch, N. Horstmann, and A. Breiter: Künstliche Intelligenz im Studium – die Sicht von Studierenden im Wintersemester 2024/25, CHE DatenCHECK 6/2025, 2025.

[7] J. von Garrel and J. Mayer: Künstliche Intelligenz im Studium – Eine quantitative Längsschnittstudie zur Nutzung KI-basierter Tools durch Studierende, Preprint, Hochschule Darmstadt, Darmstadt, 2025. https://doi.org/10.48444/h_docs-pub-533

[8] S. N. Archana, V. R. Renjith, P. K. Padmakumar, S. C., N. Aboobaker et al.: AI assisted learning and research: an exploratory study among university students and scholars, Discover Education, vol. 4, 2025.

[9] J. Freeman: Student Generative AI Survey 2025, HEPI Policy Note 61, Higher Education Policy Institute (HEPI) / Kortext, February 2025.

[10] M. Abbas, F. A. Jam, and T. I. Khan: Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students, International Journal of Educational Technology in Higher Education, vol. 21, Art. 10, 2024, doi: 10.1186/s41239-024-00444-7.

[11] D. Saleh and A. ElSayary: Generative AI use and self-learning in higher education: the role of learning difficulties, Frontiers in Education, vol. 10, Frontiers, 2026.

[12] X. Liu, Y. Xiao, and D. Li, "Assessing strategic use of artificial intelligence in self-regulated learning: Instrument development and evidence from

Chinese university students," International Journal of Educational Technology in Higher Education, vol. 22, Art. 69, 2025.

[13] I. Lausberg, J. Tosic, and S. Feldermann: Studentisches Schreiben mit generativer KI: Inspiration oder intellektuelle Aneignung?, Zeitschrift für Hochschulentwicklung, Sonderausgabe KI, no. 02/25, 2025, pp. 101-122, doi: 10.21240/zfhe/SH-KI-2/06.

[14] M. Lan and X. Zhou: A qualitative systematic review on AI empowered self-regulated learning in higher education, npj Science of Learning, vol. 10, Art. 21, Nature Publishing Group, 2025.

[15] M. Klesel and H. F. Wittmann: Retrieval-Augmented Generation (RAG), Business & Information Systems Engineering, vol. 67, no. 4, Springer, 2025, pp. 551–561.

[16] ETH Zürich: Ethel. Virtuelle Lehrassistenz, updated Dec. 5, 2025.

[17] T. Geisler and G. Kortemeyer: When AI Evaluates Its Own Work: Validating Learner-Initiated, AI-Generated Physics Practice Problems, arXiv:2508.03085, 2025.

[18] G. Kortemeyer and A. Yaroshchuk: Assessing Confidence in AI-Assisted Grading of Physics Exams through Psychometrics: An Exploratory Study, Physical Review Physics Education Research, vol. 21, Art. 010136, Apr. 2025. DOI: 10.1103/PhysRevPhysEducRes.21.010136.

[19] OpenAI: GPTs in ChatGPT, OpenAI Help Center, 2026.

[20] OpenAI: Creating and editing GPTs, OpenAI Help Center, 2026.

[21] I. Lausberg and B. Meyer, Intelligente Assistenten in der Controlling-Lehre – das Pilotprojekt HORUS, in: Behringer, S. (Hrsg.): CARF Controlling. Accounting. Risiko. Finanzen, Schriften aus dem Institut für Finanzdienstleistungen Zug IFZ, Konferenzband, Hochschule Luzern, Luzern, 2025.

[22] OpenAI: Data analysis with ChatGPT, OpenAI Help Center, 2026.

[23] M. Kovarova-Simecek, E. Barrantes, A. Horváthová, H. Zülch, L. Kirschbaum, K. Klammer, and C. Schilling: GenAI as a Reader: How ChatGPT & Co. use Annual Reports, USTP – University of Applied Sciences St. Pölten, HHL Leipzig Graduate School of Management, and nexxar, 2026.

[24] D. Molerov, D. Federiakin, O. Zlatkin-Troitschanskaia, K. Shenavai, L. Trierweiler, and M.-T. Nagel: The relationship between AI-chatbots use, student assessment performance and learning outcomes in higher education, Unterrichtswissenschaft, Springer, 2026, pp. 1–32. DOI: 10.1007/s42010-026-00242-2.

[25] X.-Y. Wu, J. D. Radloff, I. H. Yeter, L. Wang, and T. K. F. Chiu: Designing artificial intelligence chatbots for self-regulated learning from a systematic review based on Habermas's three interests, Interactive Learning Environments, Taylor & Francis, 2025. DOI: 10.1080/10494820.2025.2563086.