Proceedings of 2024 AIS SIGED European Conference on Information Systems Education Research

SIGED: IAIM Conference

7-4-2024

EMBARKING ON AN INTERACTIVE LEARNING JOURNEY: EXPLORING THE INTERACTION VALUE OF VOICEBOTS VERSUS CHATBOTS

Ricarda Schlimbach
Heillbronn UAS, Germany, ricarda.schlimbach@hs-heilbronn.de

Janine Richter
TU Braunschweig, Germany

Linda Grogorick
TU Braunschweig, Germany

EMBARKING ON AN INTERACTIVE LEARNING JOURNEY: EXPLORING THE INTERACTION VALUE OF VOICEBOTS VERSUS CHATBOTS

Completed Research Paper

Ricarda Schlimbach, Heillbronn UAS, Germany, ricarda.schlimbach@hs-heilbronn.de
Janine Richter, TU Braunschweig, Germany
Linda Grogorick, TU Braunschweig, Germany

Abstract

This study compares the perceived students' value-in-interaction of a voicebot versus a chatbot that respectively guides through an interactive, gamified learning journey. 32 students participated in this within-subject experiment while exploring learning content on the evolution of computers in a guided either text-based or voice-based interaction with a conversational agent. We measured the value in relatedness, matching, and service value in a follow-up survey, compared the number of correctly answered quiz questions at the end of the experiment, and discussed student's experiences in short interviews afterward to gather further qualitative insights. While our quantitative results do not indicate significant differences between the two interaction modes neither in descriptive statistics nor in t-tests, our qualitative results indicate that both conversational agents are perceived as valuable learning facilitators with individual preferences that are student- and context-specific regarding each interaction mode. Finally, we suggest needed further research.

Keywords: Interactive Learning, Education, Conversational Agent, Gamified Learning.

1 Introduction

Technological advancements have fostered the creation of natural language dialogue systems, known as conversational agents, which are increasingly making their mark in various sectors including education as pedagogical conversational agents (PCAs) (Diederich et al., 2022). These PCAs have the potential to transform education by serving as virtual tutors, organizers, motivators, and companions tailored to individual learning needs (Lehman & Graesser, 2014; Purington et al., 2017); thus, potentially alleviating teachers' workload and enhancing educational accessibility (Winkler, Hobert, & Appius, 2020). Previous studies have highlighted the importance of language style and the interaction mode (text-based or voice-based) as a key design element and social cue influencing valuable interactions with chatbots (Araujo, 2018; Wellnhammer et al., 2020). In addition to the rapidly increasing global usage of ChatGPT, CoPilot and other tools with conversational AI, a study by OMD (2021) indicates that alongside text-based systems, voice-based systems are also becoming increasingly relevant. More than 50% of respondents use a voice assistant at least once a week. In Germany, one-fifth of search queries are now made via voice (Pütz et al., 2021). The availability of voicebots in digital language learning platforms is increasing as well (Wills et al., 2023).

Research-based on social-response theory has demonstrated that users attribute personalities to PCAs based on the way they interact (Ahmad et al., 2022; Hanna & Richards, 2015; Nass & Moon, 2000), thus influencing users' willingness to engage and periodically interact with them (Nißen et al., 2022; Siemon et al., 2022). However, misinterpretation of user input or perceived unfulfilled expectations can lead to frustration and disappointment (Grudin & Jacques, 2019), potentially undermining the perceived benefits of PCA application in education and jeopardizing its acceptance (Janssen et al., 2021; Tsvilodub et al., 2023). Despite the knowledge base on the importance of verbal cues for bot design (Feine et al., 2019; Isbister & Nass, 2000), there is a lack of research on how its interaction mode influences the student's value in interaction (ViI). A value ascribed to the PCA is essential for recurring interactions to then boost learning outcomes (Schlimbach, Lange, et al., 2023). To make a first step towards closing this research gap, we examine students' preference for interaction with either a chatbot or voicebot. Employing Geiger et al.'s (2020a) ViI Model, which evaluates relationship, matching, and service levels, we develop a post-interaction questionnaire and a qualitative follow-up discussion. Our research compares the perceived ViI towards chatbots and voicebots quantitatively and qualitatively after an interactive learning session guided by a conversational agent. We thus intend to contribute insights into valued user-bot interaction to enhance learning outcomes in educational settings. Anticipated contributions include a deeper understanding of ViI levels (i.e., relationship, matching and service value) influencing users' preference for chatbots or voicebots and resulting implications for the preferred interaction mode in digital learning support systems.

2 Research Background

2.1 The Potential of Pedagogical Conversational Agents

Conversational Agents (CAs), also known as virtual assistants, are dialogue systems designed to respond naturally to user queries and provide accurate information (Diederich et al., 2022), facilitating natural language communication with computers (Graesser & McNamara, 2010). They can be broadly categorized into text-based and voice-based agents, with communication occurring via keyboard input for text-based agents and voice commands via microphone for voice-based agents (Kuhail et al., 2023). Their functionality relies on predefined language sequences stored in a database, matched with specific user inputs (Giang et al., 2023), creating an illusion of intelligent conversation (ibid.). Designing these dialogues poses a challenge as the aim is to simulate natural and human-like interactions (Gnewuch et al., 2018). Dialog management in CAs encompasses tasks such as automatic speech recognition, language understanding, dialogue control, database management, natural language generation, and text-to-speech synthesis (Griol et al., 2014). PCAs, a subset of CAs, support learners by providing assistance within virtual learning environments through their natural language capabilities, enabling them to ask questions or access additional materials and exercises, while advancements in educational data mining and learning analytics offer opportunities to create adaptive learning environments tailored to learners' individual needs (Giang et al., 2023).

Students primarily seek to simplify their studies and enhance productivity through PCAs (Brandtzaeg & Følstad, 2018). Unlike traditional exams, which often only reflect the final result of a long learning process, PCA interaction data analysis allows continuous tracking of individual learning processes, offering more accurate insights into progress (Giang et al., 2023). PCAs offer tailored access to resources, customizing content and providing a flexible and personalized learning environment (Clarizia et al., 2018). They facilitate quick resource access, allowing learners to focus on learning tasks, adjust to individual learning purposes (Følstad et al., 2019), and provide immediate responses, enhancing accessibility to knowledge (Diederich et al., 2022). Moreover, they enable time- and location-independent learning, particularly beneficial for working students with limited flexibility and tough schedules (Clarizia et al., 2018). PCAs facilitate the provision of feedback, reinforcing correct responses, encouraging retrying after mistakes, and building trust through empathetic responses (Dennis et al., 2016; Lee et al., 2019). They hold the potential of personalized support, enabling learners to work at their own pace and style, while assisting educators with administrative tasks and providing insights into learner behavior (Kuhail et al., 2023). However, these potentials are still evolving and, in many cases, PCAs are just used for short-term experiments but lack (technical) maturity to excel in real learning scenarios as mid- or long-term facilitators (Janssen et al., 2021; Wollny et al., 2021).

2.2 A PCA's Value-in-Interaction

Any interaction is shaped by the interactors' relationship, matching needs, and provided services (Geiger et al., 2020b). Geiger et al.'s (2021) value-in-interaction model originally designed for physical interactions has been also applied to student-PCA interactions (Schlimbach, Windolf, et al., 2023). Herein, the relationship layer focuses on the emergence of bonding relationships towards the PCA, while the matching layer involves selecting appropriate resources and competencies to facilitate learning. The service layer determines the functional feature set to facilitate learning, either immediately or throughout concurrent activities in value co-creation (Schlimbach, Windolf, et al., 2023). As depicted in Figure 1, these layers are interconnected and influence each other, to exhibit competencies across all three tiers (Geiger et al., 2021).

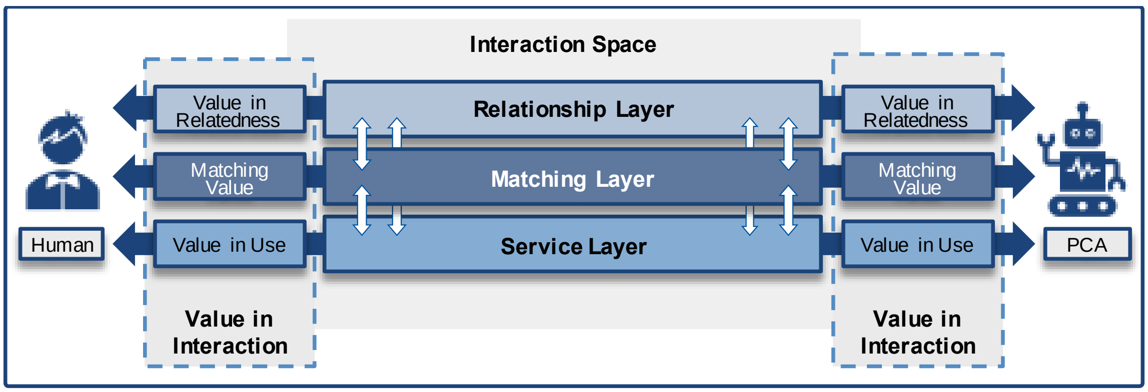

Figure 1. Iterative Artifact Design and Evaluation

People tend to exhibit fewer inhibitions in their interactions with bots compared to human counterparts, feeling more at ease asking questions without fear of judgment (Følstad et al., 2018). This lowered barrier to communication can foster a sense of trust and openness between users and bots, potentially enhancing relationship satisfaction, which is crucial to form a *value in relatedness* (Geiger et al., 2020). Besides, both chatbots and voicebots offer users the convenience of asynchronous communication and are accessible regardless of time constraints (Følstad et al., 2018). This flexibility allows users to engage with the digital assistant at their convenience, potentially contributing positively to their perceived relationship quality with the assistant.

However, disparities exist in the perceived intrusiveness and privacy implications between textual and voice-based interactions. Textual communication is often perceived as less intrusive and offers a transparent record of interaction, enabling users to revisit conversations and clarify any uncertainties (Scholl et al., 2006). In contrast, voice-based communication lacks this traceability and may raise privacy concerns, impacting relationship satisfaction, particularly when users perceive a loss of control over their data (Burbach et al., 2019). Audio technology has been shown to enhance imagination and evoke emotional responses (Liljedahl et al., 2007). Additionally, Kalla & Seiter (2021) noted that users perceive voice-operated interfaces as more elegant, natural, and user-friendly. Verbal communication allows for the expression of empathy and understanding through intonation, potentially leading users to feel that voicebots can better address their needs, thus enhancing a bonding relationship as well as an emerging *matching value*.

Følstad et al. (2018) underscore the criticality of PCAs providing effective assistance for recurring usage, which implies matching services along the interaction. Scholl et al. (2006) delineate factors favoring chatbots over voicebots, including speech comprehension and social familiarity with text-based communication. Challenges with audio quality in public spaces also influence users towards chatbots. Additionally, chatbots serve as convenient tools, offering easier access and fewer comprehension issues compared to audio applications (Kalla & Seiter, 2021). Students appreciate the relaxed pace afforded by chatbots, allowing for comfortable interaction without time pressure and private usage even when surrounded by others (Følstad et al., 2018). However, the preferred communication mode itself might be user-specific and thus contributes itself to the emerging matching value when interacting with PCAs (Schlimbach, Spill, et al., 2023).

Regarding the *service value*, Rummer et al. (2008) highlight the enhanced effectiveness of learning with auditory texts, attributing it to the additional cognitive resources they engage. Davis (1989) emphasizes the paramount importance of technology's utility and user-friendliness, with voice interfaces potentially easing effective usage (Pütz et al., 2021). Interacting with a voicebot is seen as more natural, potentially enhancing user satisfaction, and facilitating better comprehension of complex concepts (Kalla & Seiter, 2021). Chatbots facilitate asynchronous communication, granting users time to reflect on their responses at their own pace (Scholl et al., 2006). Effective operation of voice assistance systems, encompassing clarity with colloquial language but still being accurate in knowledge exchange, is pivotal for widespread acceptance and (effective) usage (Kalla & Seiter, 2021).

3 Research Design

For our experiment, we first developed learning content on the evolution of computers as a mix of content nuggets and corresponding multiple-choice questions to test the acquired knowledge successively. Using gamified elements, the learning journey was guided by a storyline of a dystopia scenario in the year 3025 with a spaceship that needs to be repaired to have the human interactor escape from Earth to Mars. With each correctly answered quiz question, the student earns points and thus gets closer to the desired takeoff. Game-based learning elements are known for enhancing motivation (Benner et al., 2022). Thus, we embedded badges and rankings into the mission as motivational elements. The experiment features two prototypes (coded by two IS students) that use the same learning content, dialogue sequences, and gamified elements but either communicate via voice or text in German language. The chatbot, coded in JavaScript and HTML with CSS defining design elements, follows a script-driven model (Kuhail et al., 2023). Responses are delayed by one second, simulating typing to enhance perceived human likeness (Gnewuch et al., 2018). Its minimalist design aims to maintain user focus. The voicebot uses a speech-to-speech functionality as implemented in e.g., Amazon's Alexa. Students navigate with speech or text commands respectively through the space mission and can thus answer quiz questions, retrieve their current score, ask for repetition, and access the leaderboard.

In randomized order, 32 students from TU Braunschweig interacted successively with both prototypes. A signed privacy statement guaranteed consent, privacy, and confidentiality with the collected data throughout the study to protect participants' rights and increase transparency. After each interaction, participants filled out a follow-up survey that measured the ViI based on the model's underlying constructs as suggested by Geiger et al (2020b;2021) and applied it to PCA interactions in previous studies (Geiger et al., 2022; Schlimbach, Windolf, et al., 2023).

Figure 2 depicts the operationalized ViI and indicates whether a sub-construct is expected to be perceived better for the voicebot (microphone icon) or chatbot (text icon), based on the literature as outlined in section 2.2.

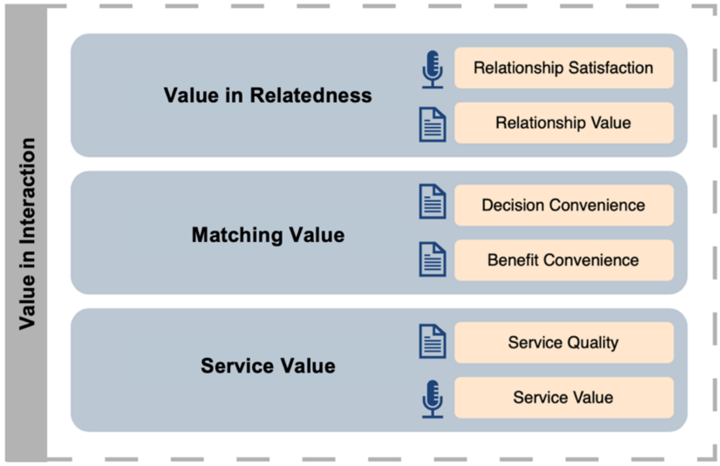

Figure 2. Operationalized Value in Interaction

In addition, we surveyed their overall experience, constructive feedback, and demographic data. We also briefly discussed with each participant at the end of the experiment what they particularly liked *(Q1)* or dislike about each prototype and what should be improved *(Q2)* during the interaction, under which conditions they prefer either PCA as a complement in digital learning contexts *(Q3)*, as well as potential application areas in education *(Q4)*. Their statements were collected in bullet points for each question (author 2) and finally clustered to core topics in the author team. Due to its complementary character and short duration, the qualitative part was rather explorative and its analysis did not follow a strict code plan. The experiment lasted approximately 30 minutes per student and scenario complemented by the beforementioned qualitative discussion (ca. 10 minutes). All participants used the same provided device (PC/speaker) and were exposed to the same lab room to ensure a consistent test experience. Before each test scenario, participants were briefed on the bot's functionalities and the interaction mode for communicating via keyboard or speech. Each participant received a handout detailing the possible commands the bot could react to.

4 Results

4.1 Quantitative Results

A total of 32 students from TU Braunschweig participated in the experiment, with majors ranging from engineering (31%), computer science (31%), and management (10%) to others such as media studies, pedagogics and social sciences (each < 6%). 22 Participants were aged in their 20s (69%), while six were younger (19%) and the remaining four were senior students older than 40 years. Among all participants 59% were male and 41% female. 75% of the participants were enrolled in a bachelor's program, while the remaining 25% studied a master's program.

In the next step, we analyzed the collected survey data. To verify the suitability of the selected items in representing the corresponding value constructs of the ViI (German questionnaire as tailored to CAs by Geiger et al., 2022), we conducted a principal component analysis. This involved examining the Bartlett test's value for sphericity and the Kaiser-Meyer-Olkin (KMO) criterion for the appropriateness of the sample (Dziuban & Shirkey, 1974). The Bartlett test value should be below the selected significance level of 5 %. With a result of $p <.001$, the correlation matrix was found to be significantly different from the random data matrix (Wolff & Bacher, 2010). The KMO criterion also achieved good to very good values for all constructs. For the reliability analysis, we looked at Cronbach's alpha, which indicates high consistency among all constructs and can thus be interpreted as very good (Gliem & Gliem, 2003). As desired, no factor loadings below .06 were found (see Table 1).

Principal Component Analysis					**Reliability Analysis**
Construct		Bartlett Test	KMO	loading < 0.6	Cronbach's Alpha
Value in Relatedness	Relationship Satisfaction	<.001	0.673	no	0.796
	Relationship Value	<.001	0.722	no	0.882
Matching Value	Decision Convenience	<.001	0.731	no	0.851
	Benefit Convenience	<.001	0.710	no	0.791
Service Value	Service Quality	<.001	0.837	no	0.919
	Service Value	<.001	0.701	no	0.794

Table 1. Pre-Analysis

For the statistical analysis in Jamovi, we conducted a dependent samples t-test after validating that the pre-conditions (paired samples, no outliers, normal distributed data) were met (Messer & Schneider, 2019).

Table 2 on the next page summarizes the results of the t-test for the dependent samples of the surveyed 32 participants.

t-Test for paired samples

Voice- vs. Chatbot			**Statistic (t)**	**p**	**Mean Difference**	**Std.-Error**		**Effect Size**
AVR_RS_AUDIO	AVR_RS_TEXT	Student's t	0.870	0.391	0.1562	0.180	Cohens d	0.1538
AVR_RV_AUDIO	AVR_RV_TEXT		0.990	0.330	0.1874	0.189		0.1750
AVR_DC_AUDIO	AVR_DC_TEXT		-1.146	0.260	-0.1719	0.150		-0.2026
AVR_BC_AUDIO	AVR_BC_TEXT		-0.317	0.753	-0.0625	0.197		-0.0561
AVR_SQ_AUDIO	AVR_SQ_TEXT		-0.317	0.753	-0.0625	0.197		-0.0561
AVR_SV_AUDIO	AVR_SV_TEXT		0.359	0.722	0.0625	0.174		0.0634

Abbreviations: Value in Relatedness*: Relationship Satisfaction (RS); Relationship Value (RV);* ***Matching Value****: Decision Convenience (DC), Benefit Convenience (BC);* ***Service Value:*** *Service Quality (SQ); Satisfaction with Service (SV).*

Table 2. Dependent t-Tests

Within the quantitative analysis of the experiment, no statistically significant difference (p >0.05) regarding the ViI and its respective sub-constructs between a chatbot and a voicebot could be identified (see Table 2 above). In line herewith, the descriptive statistical results indicate that all constructs show similar mean values (differences all below 0.25). The standard deviation, measuring data dispersion around the mean, ranged between 0.789 and 1.095, suggesting substantial variation in participant responses relative to the Likert scale's limited range of 1 to 5. Notably, the standard deviation of 1.095 for SQ covers a significant portion of the potential value range, indicating considerable participant response variability. Minimum and maximum values reflect the range of responses on the Likert scale, spanning from 1.33 to 5.00 in this study. This variability demonstrates that participants heavily differ in their subjective perceptions, contributing both low and high values, which appear to balance out in the univariate analysis based on an average value. This seems important to consider, as despite non-significant differences in the t-test comparison, the individual data sets per student demonstrate that there is indeed an individual difference and thus interaction mode preference when comparing both PCAs. The standard error, ranging from 0.140 (DC) to 0.193 (BC), suggests a relatively accurate estimation of mean values due to sample size; a smaller standard error implies greater representativeness of the sample and less susceptibility of estimated means to random fluctuations. Moreover, the measured knowledge gain operationalized by the earned points in the embedded quiz was very similar between both prototypes, as the average number of correctly answered quiz questions was 6 out of 10, regardless of the scenario (always only counting the collected points in the first prototype tested, therefore the sample size was too small for univariate tests).

4.2 Qualitative Findings

To leverage the advantages of both quantitative and qualitative approaches (Kelle, 2022), we briefly interviewed all participants after finishing the experiment. Our focus was on the students' evaluations of questions *Q1-Q4* and the importance they assign, from their user perspective to the measured constructs in PCA interactions across various learning contexts. We explore a potential qualitative preference for either of the two interaction modes.

Q01: What did you particularly like about interacting with the bot? Overall, participants valued the bots' thorough formulation of questions, abundant factual content, and inclusion of *"playfully embedded technical terms"*. They appreciated the gradual presentation of information followed by immediate queries. Gamified elements like earning points motivated them (as expected), and they positively received learning through storytelling. The questions were deemed appropriate and easy to answer. Concerning the voicebot, participants praised its pleasant voice, refreshing sound effects, and positive reinforcement for correct answers. They claimed the sound effects enriched the learning experience and increased engagement. Listening required higher concentration, aiding in better retention. Participants enjoyed not relying solely on visual learning, thus avoiding the effort of reading and recognizing important information. The voicebot's clarity was also commended. Feedback for the chatbot highlighted its flexibility for self-paced learning, allowing *"a lot more time for reflection on questions"*. Interacting with the chatbot was perceived as practical, straightforward, and easy to understand. Participants appreciated the sleek response system and the ability to review information if needed. Compared to the voicebot, the chatbot was perceived as faster, which was perceived as *"great to get information but stressful when being under pressure to solve the quiz"*. Some participants found it beneficial that text was displayed gradually, facilitating comprehension. They particularly valued the opportunity to reread the text, enhancing the learning experience. Text input was considered more precise compared to the voicebot, although the wording was actually exactly the same.

Q02: What did you not like during interaction with the bot? How could it be improved? For the voicebot, participants noted occasional issues with speech recognition, hindering clear communication. Some found the bot's speaking speed too fast and additional information distracting. The bot's pronunciation was criticized for lacking human-like qualities. Participants desired text for reading along in case of missed information. One participant found the voicebot stressful, preferring text input over voice commands. Overall, the voicebot was deemed impractical. Regarding the chatbot, respondents felt that placing answers above questions was unnecessary, given the simplicity of questions. Typing speed was criticized for being either too slow or too fast, with a desire for customization. Continuous text appearing proved distracting for some users, suggesting an option to pause typing and display the entire text. Attention retention was perceived to be quicker with the voicebot than with the chatbot. However, students assumed that in field usage, they would probably interact more likely with a voicebot *"when doing something else like cooking or cleaning"* in a non-laboratory situation, whereas they would *"more likely just sit down for learning as a primary activity with a chatbot"*. Suggestions for improvement included incorporating visualizations into the text for clarity and responding with the correct answer instead of *"answer number X"*. Some also criticized the verbosity of the content.

Q03: How did the bot assist you in learning? Q04: Where else could you envision using such a bot in an educational context? Both bots were generally perceived as supportive. They provided a helpful overview of topics and were particularly useful in introducing a new section or for post-lecture review. They aided in summarizing information and encouraged independent thinking. Students recommend the voicebot for college education or for informal learning accompanying leisure activities, as it could engage their attention, especially for knowledge recall (like flashcards). Some users found it helpful for memorization and language learning. While it may pique interest, it may not be suitable for formal learning. Auditory learning was valued for reinforcing information differently. Potential uses included car rides or situations where *"you would otherwise just listen to music"*, or benefiting visually impaired individuals and those with literacy challenges. It could also combat low motivation for reading. Regarding the chatbot, applications included memorizing factual content or language learning, along with preparing for driving, fishing, or hunting exams, and historical study. It offered interactive and motivational learning opportunities for classroom and individual exam preparation. It served as an alternative learning method, adding variety. Travel settings like commuting by train were suggested. One participant proposed using the chatbot to delve into a new topic and the voicebot for collaborative assessments with peers.

Overall, participants appreciated the thoroughness of bot interactions, though they noted individual and situation-specific preferences as well as issues such as misunderstandings and the risk of distraction. Both bots were generally considered helpful for learning, with potential applications ranging from informal learning to exam preparation, although some users found them less suitable for certain subjects (i.e., mathematics).

5 Discussion

Our study explored the perceived value in interaction of a voicebot in comparison to a chatbot. Although we could not find significant differences in the perceived ViI in our quantitative analysis, our qualitative findings revealed that participants presented compelling arguments both for and against either of the two interaction modes. Some participants favored the chatbot for its written interaction and the ability to review text, while others preferred the voicebot for its more natural communication through a human voice and intonation. These diverse preferences were also reflected in the descriptive statistics, where participants utilized the full range of the Likert scale. The variation in participants' responses, as indicated by the standard deviations (ranging from 0.791 to 1.029), suggests individual differences in perception of interaction. Factors such as personal learning preferences, background experience, or technical affinity may influence the choice of interaction type. These results underscore the importance of personalizing educational applications to accommodate users' diverse needs and preferences. Similar mean values on all interaction value levels indicate that both interaction modes potentially result in effective tools in education to support learners in acquiring knowledge and skills.

However, the decision to use a particular type of interaction should consider the individual preferences and needs of users. In this course, it is notable that adaptation is predominantly discussed as a minor aspect in existing design knowledge concerning PCAs (Biancardi et al., 2021; Schlimbach et al., 2022), suggesting that their crucial role in fostering relationships and generating value subjectively has been overlooked so far (Schlimbach, Spill, et al., 2023). Nonetheless, numerous research contributions acknowledge the significance of adjusting learning applications to tailor education, offering advantages such as personalized learning, increased inclusivity in academia, and the potential to revolutionize education (Gupta & Chen, 2022; Wollny et al., 2021). Offering various options and allowing learners to choose their preferred mode of interaction is a first step towards that goal. Furthermore, participants' qualitative feedback underscored the importance of interpreting results within specific contexts in line with research (van der Zandt et al., 2021). However, recent literature, as outlined by Schlimbach, Rinn et al. (2022), suggests that this potential has yet to be fully utilized in implemented PCAs. Participants noted that their preference for a PCA's conversational mode varies significantly depending on the learning situation (e.g., time constraints, audience, subject matter) and the educational objectives, aligning with findings from Nißen et al. (2022) and van der Zandt et al. (2021).

Particularly with regards to the evolving roles of PCAs, ranging from hierarchical tutors to motivational coaches and learning companions (Khosrawi-Rad et al., 2022), this suggests that the effects of the implemented interaction mode should not be universally measured and designed, but rather tailored to specific usage contexts. For example, while some students favored a voicebot for *"learning on the fly"* during leisure activities (e.g., incorporating knowledge nuggets into a podcast), they anticipated greater acceptance for formal learning tasks (e.g., exam preparation, and course assignments) with text-based PCAs. Incorporating gamified elements may be essential for fostering ongoing engagement with the bot (Benner et al., 2022). These insights have implications for the design of educational applications, suggesting that a personalized and diverse interaction environment can enhance learning outcomes and user satisfaction as shown in prior studies (Schlimbach, Lange, et al., 2023). From a research perspective, our study further highlights the importance of not only drawing conclusions from statistical averages but also considering individual learner preferences. This ensures that promising learning facilitators like PCAs truly provide a customized learning experience *rather than merely catering to the average*.

Future research might focus on examining specific features and contextual factors to gain a more comprehensive understanding of user preferences and the value of tailored interaction modes. Our study was limited in that the experiment was not part of students' formal curriculum and might have impacted their perception of and seriousness towards the PCA's interaction. It appears crucial to explore which interaction mode aligns best with the diverse roles and learning contexts that PCAs may inhabit. For instance, the conversational style may affect the PCA's relationship-building, such as learning companions aiming to establish long-term, friendship-like connections. Subsequent studies could provide valuable insights into the design considerations for PCAs tailored to an interaction mode that aligns with their language style, personality, and respective role in learning facilitation. Additionally, future investigations could examine the integration of emerging technologies, such as LLMs, and their potential to finetune social cues, styles, and context sensitivity. Exploring the adaptation of verbal cues to foster human-bot collaboration, motivate, and integrate pedagogical aspects (like scaffolding techniques) holds promise for leveraging new potentials in PCA interactions (Isbister & Nass, 2000; Winkler, Hobert, Fischer, et al., 2020).

Moreover, on a meta-level, models like Geiger et al.'s ViI framework could guide future research by focusing on the value generated for the human student in specific learning scenarios, rather than offering generalized design recommendations. This framework's value levels, including relationship, matching, and service, potentially align with desirable features per level linked to an adaptable interaction mode (Schlimbach, Spill, et al., 2023). For instance, factors like trust and empathy may be particularly crucial for relationship-oriented PCAs and easier achievable through a speech-based PCA whose voice usually provokes a warm attitude (Novielli et al., 2010) and thus becomes more familiar over time. Currently, many PCAs fail and do not leverage their full potential in learning facilitation (Wollny et al., 2021), but more research i.e., in designing the interaction mode with purpose might be a further step toward that desired direction.

In conclusion, we advocate for a more value-oriented approach in PCA research and thus encourage future research with larger participant samples, diverse learning contexts, and varied experiment durations to comprehensively explore the nuances of PCA interaction modes.

References

Ahmad, R., Siemon, D., Gnewuch, U., & Robra-Bissantz, S. (2022). Designing Personality-Adaptive Conversational Agents for Mental Health Care. Information Systems Frontiers, 1–21. https://doi.org/10.1007/s10796-022-10254-9

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Computers in Human Behavior, 85, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Benner, D., Schöbel, S., Suess, C., Baechle, V., & Janson, A. (2022, February 21). Level-Up your Learning -Introducing a Framework for Gamified Educational Conversational Agents. Proceedings of WI2022. 17th International Conference on Wirtschaftsinformatik, Nürnberg, Germany.

Biancardi, B., Dermouche, S., & Pelachaud, C. (2021). Adaptation Mechanisms in Human–Agent Interaction: Effects on User's Impressions and Engagement. Frontiers in Computer Science, 3. Scopus. https://doi.org/10.3389/fcomp.2021.696682

Brandtzaeg, P., & Følstad, A. (2018). Chatbots: Changing user needs and motivations. Interactions, 25, 38–43. https://doi.org/10.1145/3236669

Burbach, L., Halbach, P., Plettenberg, N., Nakayama, J., Ziefle, M., & Calero Valdez, A. (2019). “Hey, Siri”, “Ok, Google”, “Alexa”. Acceptance-Relevant Factors of Virtual Voice-Assistants. 2019 IEEE International Professional Communication Conference (ProComm), 101–111. https://doi.org/10.1109/ProComm.2019.00025

Clarizia, F., Colace, F., Lombardi, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An Education Support System for Student. In A. Castiglione, F. Pop, M. Ficco, & F. Palmieri (Eds.), Cyberspace Safety and Security (pp. 291–302). Springer International Publishing. https://doi.org/10.1007/978-3-030-01689-0_23

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 319–340.

Dennis, M., Masthoff, J., & Mellish, C. (2016). Adapting Progress Feedback and Emotional Support to Learner Personality. International Journal of Artificial Intelligence in Education, 26(3), 877–931. https://doi.org/10.1007/s40593-015-0059-7

Diederich, S., Brendel, A., Morana, S., & Kolbe, L. (2022). On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research. Journal of the Association for Information Systems.

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychological Bulletin, 81, 358–361. https://doi.org/10.1037/h0036316

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. International Journal of Human-Computer Studies, 132, 138–161. https://doi.org/10.1016/j.ijhcs.2019.07.009

Følstad, A., Nordheim, C., & Bjørkli, C. (2018). What Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study. 194–208. https://doi.org/10.1007/978-3-030-01437-7_16

Følstad, A., Skjuve, M., & Brandtzaeg, P. (2019). Different Chatbots for Different Purposes: Towards a Typology of Chatbots to Understand Interaction Design (pp. 145–156). https://doi.org/10.1007/978-3-030-17705-8_13

Geiger, M., Ahmad, R., Neumann, E., & Robra-Bissantz, S. (2022). Smart Services – Gestaltung wertvoller Interaktionen mit persönlichkeitsadaptiven Conversational Agents. In M. Bruhn & K. Hadwich (Eds.), Smart Services: Band 3: Kundenperspektive – Mitarbeiterperspektive – Rechtsperspektive (pp. 79–103). Springer Fachmedien. https://doi.org/10.1007/978-3-658-37384-9_3

Geiger, M., Jago, F., & Robra-Bissanzt, S. (2021). Physical vs. Digital Interactions: Value Generation Within CustomerRetailer Interaction. 153–165. https://doi.org/10.18690/978-961-286-485-9.12

Geiger, M., Robra-Bissantz, S., & Meyer, M. (2020a). Wie aus digitalen Services Wert entsteht: Interaktionen richtig gestalten. HMD Praxis der Wirtschaftsinformatik, 57(3), 385–398. https://doi.org/10.1365/s40702-020-00611-0

Geiger, M., Robra-Bissantz, S., & Meyer, M. (2020b, June 30). Focus on Interaction: Applying Service-Centric Theories in IS. https://doi.org/10.18690/978-961-286-362-3.46

Giang, C., Wambsganss, T., & Käser, T. (2023). Maschinelles Lernen zur Förderung von höheren Kompetenzen. Lernen und Lernstörungen, 12(2), 67–81. https://doi.org/10.1024/2235-0977/a000393

Gliem, J. A., & Gliem, R. R. (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. Proceedings. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, Ohio, USA.

Gnewuch, U., Morana, S., Adam, M., & Maedche, A. (2018, June 1). Faster Is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction.

Graesser, & McNamara, D. (2010). Self-Regulated Learning in Learning Environments With Pedagogical Agents That Interact in Natural Language. Educational Psychologist, 45(4), 234–244. https://doi.org/10.1080/00461520.2010.515933

Griol, D., Molina, J. M., & Sanchís De Miguel, A. (2014). Developing multimodal conversational agents for an enhanced e-learning experience. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 3(1), 13–26. https://doi.org/10.14201/ADCAIJ2014381326

Grudin, J., & Jacques, R. (2019). Chatbots, Humbots, and the Quest for Artificial General Intelligence. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–11. https://doi.org/10.1145/3290605.3300439

Gupta, S., & Chen, Y. (2022). Supporting Inclusive Learning Using Chatbots? A Chatbot-Led Interview Study. Journal of Information Systems Education, 33(1), 98–108.

Hanna, N., & Richards, D. (2015). The influence of users' personality on the perception of intelligent virtual agents' personality and the trust within a collaborative context: International Workshop on Collaborative Agents Research and Development (6th : 2015) and International Workshop on Multiagent Foundations of Social Computing (2nd : 2015). Advances in Social Computing and Multiagent Systems, 541, 31–47. https://doi.org/10.1007/978-3-319-24804-2_3

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. International Journal of Human-Computer Studies, 53(2), 251–267. https://doi.org/10.1006/ijhc.2000.0368

Janssen, A., Grützner, L., & Breitner, M. H. (2021, December 12). Why do Chatbots fail? A Critical Success Factors Analysis. Proceedings of the 42th International Conference on Information Systems.

Kalla, M., & Seiter, M. (2021). Einsatzszenarien digitaler Sprachassistenzsysteme im Dienstleistungsmanagement. In M. Bruhn & K. Hadwich (Eds.), Künstliche Intelligenz im Dienstleistungsmanagement (pp. 155–183). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-34326-2_6

Kelle, U. (2022). Mixed Methods. In N. Baur & J. Blasius (Eds.), Handbuch Methoden der empirischen Sozialforschung (pp. 163–177). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-37985-8_9

Khosrawi-Rad, B., Rinn, H., Schlimbach, R., Gebbing, P., Yang, X., Lattemann, C., Markgraf, D., & Robra-Bissantz, S. (2022). Conversational Agents in Education – A Systematic Literature Review. ECIS 2022 Research Papers. https://aisel.aisnet.org/ecis2022_rp/18

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with Educational Chatbots: A Systematic Review. Education and Information Technologies, 28(1), 973–1018. https://doi.org/10.1007/s10639-022-11177-3

Lee, S.-Y., Lee, G., Kim, S., & Lee, J. (2019). Expressing personalities of conversational agents through visual and verbal feedback. Electronics (Switzerland), 8(7). Scopus. https://doi.org/10.3390/electronics8070794

Lehman, B., & Graesser, A. (2014). Impact of agent role on confusion induction and learning: Vol. 8474 LNCS (p. 54). Scopus. https://doi.org/10.1007/978-3-319-07221-0_6

Liljedahl, M., Papworth, N., & Lindberg, S. (2007). Beowulf: An audio mostly game. Proceedings of the International Conference on Advances in Computer Entertainment Technology, 200–203. https://doi.org/10.1145/1255047.1255088

Messer, M., & Schneider, G. (2019). Der t-Test. In M. Messer & G. Schneider (Eds.), Statistik: Theorie und Praxis im Dialog (pp. 111–129). Springer. https://doi.org/10.1007/978-3-662-59339-4_9

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. Journal of Social Issues, 56(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nißen, M., Selimi, D., Janssen, A., Cardona, D. R., Breitner, M. H., Kowatsch, T., & von Wangenheim, F. (2022). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. Computers in Human Behavior, 127, 107043. https://doi.org/10.1016/j.chb.2021.107043

Novielli, N., De Rosis, F., & Mazzotta, I. (2010). User attitude towards an embodied conversational agent: Effects of the interaction mode. Journal of Pragmatics, 42(9), 2385–2397. https://doi.org/10.1016/j.pragma.2009.12.016

OMD. (2021). The Age of Voice 3.0" Sprachsteuerung wird für die Verbraucher zur Routine – doch bezüglich Nutzer- und Markenführung liegen noch erhebliche Potenziale brach [Customer Study]. https://www.omnicommediagroup.de/news/omd-studie-the-age-of-voice-30-sprachsteuerung-wird-fuer-die-verbraucher-zur-routine-doch-bezuegl/?tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Baction%5D=detail&cHash=c6aab0e83180f85b8929d1e2154164af

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). "Alexa is my new BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 2853–2859. https://doi.org/10.1145/3027063.3053246

Pütz, C., Düppre, S., Roth, S., & Weiss, W. (2021). Akzeptanz und Nutzung von Chat-/Voicebots. In M. Bruhn & K. Hadwich (Eds.), Künstliche Intelligenz im Dienstleistungsmanagement (pp. 361–383). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-34326-2_14

Rummer, R., Schweppe, J., Scheiter, K., & Gerjets, P. (2008). Lernen mit Multimedia: Die kognitiven Grundlagen des Modalitätseffekts. Psychologische Rundschau, 59(2), 98–107. https://doi.org/10.1026/0012-1924.59.2.98

Schlimbach, R., Lange, T., & Robra-Bissantz, S. (2023). A Longitudinal Study on Boosting Students' Performance with a Learning Companion. ICIS 2023 Proceedings. https://aisel.aisnet.org/icis2023/learnandiscurricula/learnandiscurricula/4

Schlimbach, R., Rinn, H., Markgraf, D., & Robra-Bissantz, S. (2022). A Literature Review on Pedagogical Conversational Agent Adaptation. PACIS 2022 Proceedings. https://aisel.aisnet.org/pacis2022/20

Schlimbach, R., Spill, M., & Robra-Bissantz, S. (2023). Increasing the Value-in-Interaction with Adaptable Learning Companions. ICIS ER Proceedings.

Schlimbach, R., Windolf, C., & Robra-Bissantz, S. (2023). A Service Perspective on Designing Learning Companions as Bonding and Mindful Time Managers in Further Education. ECIS 2023 Proceedings.

Scholl, J., McCarthy, J., & Harr, R. (2006). A comparison of chat and audio in media rich environments. Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, 323–332. https://doi.org/10.1145/1180875.1180925

Siemon, D., Strohmann, T., Khosrawi-Rad, B., de Vreede, T., Elshan, E., & Meyer, M. (2022). Why Do We Turn to Virtual Companions? A Text Mining Analysis of Replika Reviews. In AMCIS 2022 Proceedings. https://aisel.aisnet.org/amcis2022/sig_hci/sig_hci/10

Tsvilodub, P., Chevalier, E., Klütz, V., Oberbeck, T., Sigetova, K., & Wollatz, F. (2023). Improving a Gamified Language Learning Chatbot Through AI and UX Boosting. Lecture Notes in Networks and Systems, 581 LNNS, 557–569. Scopus. https://doi.org/10.1007/978-3-031-21569-8_52

van der Zandt, L., van der Stappen, E., & Van Turnhout, K. (2021). Towards Real-Life Adoption of Conversational Interfaces: Exploring the Challenges in Designing Chatbots That Live up to User Expectations. Proceedings of the 34th British HCI Conference, 306–311. https://doi.org/10.13140/RG.2.2.35102.25927/1

Wellnhammer, N., Dolata, M., Steigler, S., & Schwabe, G. (2020). Studying with the Help of Digital Tutors: Design Aspects of Conversational Agents that Influence the Learning Process. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2020.019

Wills, S., Bai, Y., Tejedor-Garcia, C., Cucchiarini, C., & Strik, H. (2023). Automatic Speech Recognition of Non-Native Child Speech for Language Learning Applications. https://doi.org/10.48550/ARXIV.2306.16710

Winkler, R., Hobert, S., & Appius, M. (2020). Empowering Educators to Build their Own Conversational Agents in Online Education. ECIS 2020 Research-in-Progress Papers. https://aisel.aisnet.org/ecis2020_rip/53

Winkler, R., Hobert, S., Fischer, T., Salovaara, A., Soellner, M., & Leimeister, J. M. (2020). Engaging Learners in Online Video Lectures with Dynamically Scaffolding Conversational Agents. ECIS 2020 Research Papers. https://aisel.aisnet.org/ecis2020_rp/97

Wolff, H.-G., & Bacher, J. (2010). Hauptkomponentenanalyse und explorative Faktorenanalyse. In C. Wolf & H. Best (Eds.), Handbuch der sozialwissenschaftlichen Datenanalyse (pp. 333–365). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92038-2_15

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are We There Yet? - A Systematic Literature Review on Chatbots in Education. Frontiers in Artificial Intelligence, 4. Scopus. https://doi.org/10.3389/frai.2021.654924