Engaging Minds – How Gamified Chatbots can Support and Motivate Learners in Digital Education

Dennis Benner
University of Kassel
benner@uni-kassel.de

Sofia Schöbel
University of Osnabrück
sofia.schoebel@uni-osnabrueck.de

Andreas Janson
University of St. Gallen
andreas.janson@unisg.ch

Jan Marco Leimeister
University of Kassel
leimeister@uni-kassel.de

Abstract

Blended and online learning is growing, and self-regulated learning is becoming more relevant. Most often, students struggle with organizing their own learning processes, lose focus or procrastinate. Keeping learners motivated and engaged can be a real challenge. Therefore, we present gamified chatbots as a potential solution. On the one hand, chatbots can provide a more engaging learning experience. On the other hand, gamification can provide motivational incentives to keep learners engaged and motivated. So far, not many studies have elaborated on how gamification can be effectively used to make a chatbot interaction more engaging or improve the learning experience. This study uses an experimental approach to distinguish how a combination of badges and a progress bar can support and motivate learners to stay engaged with their learning activities. We elaborate on the effects of gamified chatbots and support practitioners with guidance on how to design gamified chatbots in education.

Keywords: Gamification, Chatbot, Motivation, Engagement, Education

1. Introduction

Digital learning allows for more flexibility and facilitates better access to learning resources (Broadbent & Poon, 2015). With the increase in digital and self-regulated learning, it has become more important to effectively motivate learners to train with digital tools (Janson et al., 2020; Schlegel et al., 2023). In digital education, chatbots are becoming more relevant in teaching and have received more attention in research (Hobert & Meyer von Wolff, 2019; Schöbel et al., 2021). Chatbots support all kinds of learning processes by communicating with users in natural language via a user interface (UI) that is similar to typical messenger services (Schöbel et al., 2023; Weber et al., 2021). Chatbots promise individual as well as time- and location-independent support of learners in a 1:1 relationship (Hobert & Meyer von Wolff, 2019). The dialogic and interactive learning form of chatbots can lead to higher motivation of users, which in turn can have a positive effect on their learning outcomes (Ruan et al., 2020; Winkler & Söllner, 2020).

However, the design of chatbots in education is challenging. Chatbots can fail because of their poor design and wrong embeddingin the learning context, resulting in a poor learning experience (Wellnhammer et al., 2020). Furthermore, learning with chatbots requires advanced metacognitive skills from users so that they can self-regulate their learning processes to achieve their learning goals (Broadbent & Poon, 2015). In this regard, there is a lack of studies that consider contextual and motivational factors in the design of chatbots in digital education (Hobert & Meyer von Wolff, 2019; Wellnhammer et al., 2020). A promising solution approach to make the UX of chatbots in education more effective and to increase learner motivation and satisfaction is gamification. Gamification describes the use of elements from games, such as badges, progress bars, or levels, in non-entertainment contexts (Deterding, Dixon, et al., 2011; Deterding, Sicart, et al., 2011).

In digital learning, gamification has already been extensively researched, supporting its effectiveness on learning motivation and learning outcomes (Koivisto & Hamari, 2019). However, the scientific investigation of gamified chatbots is not yet very advanced (Diederich et al., 2022; Janson et al., 2023), so it is to be discussed if game elements are able to optimize the UX of chatbots in educational contexts. Therefore, with this study, we answer the following research question:

RQ: How can a gamified chatbot in education support the motivation and learning behavior?

URI: https://hdl.handle.net/10125/106383
978-0-9981331-7-1

We conducted a 2x2 between-subjects experiment with three groups of gamified chatbots and a control group of non-gamified chatbots, collecting 164 datasets to answer our research question.. Based on our findings, we present design recommendations and the end of our paper. Researchers and practitioners are guided by our study results in several ways. By identifying how gamified chatbots can influence motivation, cognitive absorption, and learner behavior, our research contributes to theory. Additionally, we consolidate our results to provide implications for future research. From a practitioner's perspective, our research provides guidance on how to gamify chatbots for educational contexts to better self-regulate learning processes.

The remainder of this paper is structured as follows, after motivating our study, we provide an overview of related work and specify the terms gamification and chatbots. In the next step, we provide an overview of our method and discuss our results by presenting propositions. Lastly, we outline limitations, future research ideas, and a conclusion.

2. Related Research

Chatbots are so-called text-based conversational agents (Araujo, 2018). In other words, they are software programs that communicate with users using either artificial intelligence (AI) or rule-based natural language dialog. They assist users in performing various tasks and provide contextual and personalized information and assistance (Diederich et al., 2022). Unlike voice assistants such as Amazon's Alexa, chatbots are characterized by text-based communication and therefore have a UI that resembles that of a messaging service.By providing personalized support and making learning content more engaging or interesting, chatbots in digital learning can provide structure to self-directed learning. However, chatbots often fail to motivate learners to use them and to provide them with a pleasant user or learner experience (Benner, Schöbel, & Süess, 2022; Benner, Schöbel, Süess, et al., 2022; Schöbel et al., 2023).

Gamification can be used to design an interaction with a chatbot that is more motivating and engaging to users (Hamari & Koivisto, 2013). The primary goal of gamification is to support meaningful engagement (Liu et al., 2017). Through the integration of game design elements, a gameful experience, which encompasses the enjoyment, immersion, and sense of competence that users experience while playing a game, users are intrinsically motivated leading to desired behaviors (Hamari et al., 2016; Xi & Hamari, 2019).

Table 1. Related research.

Game design elements / **Source**	Points	Badges	Virtual Goods	Ranking	Levels	Progress	Feedback	History	User avatar	System avatar	Goal setting	Story telling
Alaimi et al. (2020)										X	X	
Allameh and Zaman (2021)							X					
Benner, Schöbel, and Süess (2022)		X			X		X				X	
Fadhil and Villafiorita (2017)	X		X	X						X		
Fidan and Gencel (2022)							X			X		
Giboney et al. (2021)										X	X	X
Gupta and Chen (2022)						X	X			X	X	
Hobert (2019)							X				X	
Katchapakirin et al. (2022)	X			X	X					X	X	
Lee et al. (2021)		X				X			X	X	X	X
Pereira (2016)							X	X				
Pereira and Díaz (2021)								X			X	
Pham et al. (2018)									X	X		
Ruan et al. (2019)							X			X		
Ruan et al. (2020)	X				X	X	X			X	X	X
Tegos and Demetriadis (2017)									X	X	X	
Tian et al. (2021)							X			X	X	
Waldner et al. (2022)										X		
Wambsganss et al. (2021)	X						X	X		X	X	
Winkler et al. (2020)							X					
Winkler and Söllner (2020)							X				X	
Yin et al. (2021)							X			X		
Sum (*n* = 22)	4	2	1	2	3	3	14	3	3	15	13	3

Chatbots in digital learning can increase, satisfaction, fun, and motivation of learners, which in turn leads to users learning more often with the help of a chatbot (Koivisto & Hamari, 2019; Liu et al., 2017).

To analyze related work on gamified chatbots, we conducted a systematic literature review following Vom Brocke et al. (2015). We used the keywords (learn* OR educat* OR pedagogical OR tutor*) AND (agent OR chatbot OR assistant) and searched in AISeL, ACM DL, and IEEE Xplore. We ended up with 491 articles. After excluding studies that did not focus on chatbots in education, 22 articles remained, that we analyzed in more detail by looking at the design of the chatbot and the game elements that were used (see Table 1 above). We based our analysis on the taxonomy of game design elements by Schöbel et al. (2020). We find that the majority of research on gamifying chatbots focuses on feedback, avatars and goals while rewards and progress are under researched. There is research investigating the combination of rewards such as badges and progress bars in a general application (e.g., Mazarakis & Bräuer, 2020); however, these studies were not conducted in the context of chatbots.

3. Gamified Chatbot Design to Support and Motivate Learners

For our educational use case, we chose a topic that most participants would be able to understand and relate to. Therefore, we chose the theoretical driving test as the topic, based on official multiple-choice questions of the driver's license exam of our country. From a pedagogical point of view all questions represent the lower learning goal dimensions of testing knowledge and comprehension of the material (Krathwohl, 2002). We have organized the questions into six separate lessons that will be unlocked sequentially, challenging our participants to continue with the training. For the experimental design, we decided to make the first two lessons mandatory for successful participation and the other four lessons optional. Based on the interplay between the gamification design and the intensity of chatbot use, we tracked the number of completed lessons to derive potential implications.. Regarding the design of the game elements we focus on elements that respect the learners' basic psychological needs such as autonomy (Benner, Schöbel, Süess, et al., 2022; Sailer et al., 2017). We decided to exclude competitive game elements (e.g., rankings or leader boards), as research has shown that such designs, especially in digital learning, can have negative effects such as loss of motivation or decrease in learning performance (Benner, Schöbel, Süess, et al., 2022; Santhanam et al., 2016)., Competitive elements can encroach on users' need or rights regarding ethical concerns (Benner, Schöbel, Janson, & Leimeister, 2022). Based on the logic of self-determination theory (Deci & Ryan, 2000; Ryan & Deci, 2000a) and aligned with the results of our literature review, we implement two game elements based on the gamification mechanisms of collection (i.e., badges) and progression (i.e., progress bar). These two game design elements have been shown to work alone and in combination in a non-conversational setting (Mazarakis & Bräuer, 2020).

The first game element - badges (see Figure 1), can be earned by learners forfor completing lessons (quantity) and answering questions correctly (quality). Badges have the potential to increase user activity (Hamari, 2017), which in our context is reflected in the number of lessons completed. In addition, badges can enhance the learning experience, make learning more interesting or help to motivate learners (Davis & Singh, 2015). In relation to our experiment, we raise Hypothesis 1 (H1): *a chatbot gamified with badges has a positive impact on motivating learners*.

Figure 1. Badge collection.[1]

We have chosen a selection of six badges for the badge design. Badges can be earned by completing lectures or answering questions correctly. At the beginning, the badges are hidden and only the criteria are revealed. This badge design is a common design choice for many video games or gaming platforms, such as World of Warcraft or Steam. The second game element - progress bar (see Figure 2), visualizes the learning progress within each lesson based on the

[1] Translated from German in English

number of questions answered correctly. Similar to badges, the progress bar can positively influence learning (Mazarakis & Bräuer, 2020). Our approach to the design of the progress bar is a simple bar that indicates the progress of the learner within a lesson. In terms of our experimental evaluation, we argue that our gamified progress bar implemented in our chatbot can have a positive influence on learning as suggested by the literature on non-conversational gamification. Thus, we propose Hypothesis 2 (H2): *a gamified chatbot with a progress bar has a positive impact on motivating learners.*

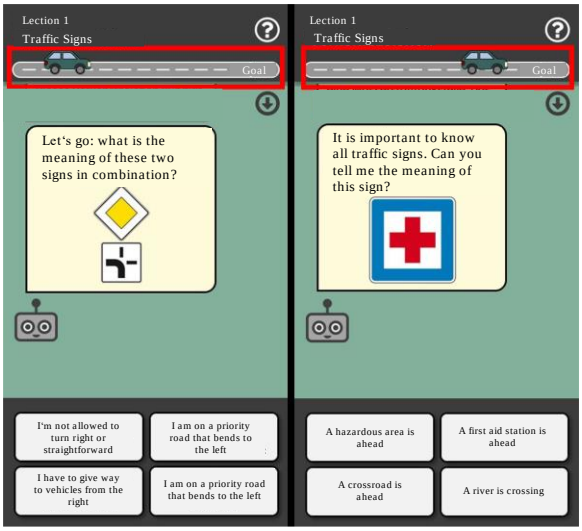

Figure 2. Progress bar (see red marking).[1]

4. Method & Data Collection

To analyze our gamified chatbot, we conducted a 2x2 between-subjects experiment with a quantitative questionnaire. All subjects (n = 164) participated on a voluntary basis, and we incentivized participation with a gift card lottery. We recruited our participants in a university context, mostly including students, tutors, and neighboring people, we did not specifically choose any particular demographic for our participants (see Table 2 for demographics). Overall, our sample is fairly balanced, but we do have a slight bias in certain areas such as a bias towards young to medium aged female participants. Additionally, because of our driving test context, we control for this factor as well. All but 10 participants already have a driver's license and all, but three participants are currently not learning for a driving examination.

We have four treatment groups: T0 representing the control group with no design, T1 representing the badge treatment, T2 representing the progress bar treatment and T3 representing the combined treatment of badges and progress bar design (see table 3).

Table 2. Participant demographics.

Characteristic	Attribute	Value
Gender	Male	33.54 %
	Female	65.85 %
	Diverse or prefer not to disclose	0.61 %
Age	<18	0.61 %
	18-24	23.78 %
	25-34	52.44 %
	35-44	9.76 %
	45-54	7.32 %
	55-64	6.10 %
	>64	0 %

Each participant was randomly assigned to one of our four groups at the beginning of the experiment. Then, each participant was given an explanation and instructions for the experiment, e.g., what the task was, how the experiment would be conducted. All participants had unlimited time to familiarize themselves with the explanation and instructions. Participants started the interaction from the menu screen of our chatbot application, where they could view and select the lessons. As mentioned above, two lessons were mandatory for successful participation. However, after completing the mandatory lessons, participants had the option to select additional lessons from the main menu. The additional lessons could then be unlocked up to a maximum of six lessons. The gamification design followed a similar approach. With each unlocked and successfully completed lesson, participants received a badge while being able to track their progress within each lesson. Once our participants decided they wanted to stop or had completed all lessons, they were directed to a questionnaire. We used a 7-point Likert scale (Likert, 1932) ranging from 1 = low/disagree to 7 = high/agree, for all of our questions. The questionnaire included an instructional manipulation check (Oppenheimer et al., 2009) that participants had to pass. Participants who failed the manipulation check were excluded, resulting in a total of 164 successful participants divided into four groups as shown in Table 3.

Table 3. Experimental setup.

Treatment	T0	T1	T2	T3
Badge		X		X
Progress bar			X	X
Participants	33	53	35	43

To evaluate the results of our experiment, we used a set of widely known and accepted constructs and scales. First, we measured motivation as it is the primary concept for gamification design (Koivisto & Hamari, 2019; Majuri et al., 2018) and a key aspect of our research.We included the scale of Ryan et al. (1983) and its constructs of interest/enjoyment,

pressure/tension as well as effort/importance. In addition, we include Self-Determination Theory (SDT) to account for the basic psychological needs satisfaction in terms of competence, autonomy, and relatedness (Deci & Ryan, 2000; Ryan & Deci, 2000b). We additionally include the concept of cognitive absorption in our questionnaire. Cognitive absorption describes how individuals perceive the interaction with an artifact and the effect of "time flies" when users are engaged (Agarwal & Karahanna, 2000). For cognitive absorption, we rely on the joy, control, focused immersion, temporal dissociation, and curiosity (Agarwal & Karahanna, 2000). Additionally, we decided to include the Technology Acceptance Model (TAM) because we are implementing a fairly novel artifact and it is thus relevant to know how users perceive its usefulness (PU), ease of use (PEOU), and whether they would intend to use (BITU) the artifact (Venkatesh & Davis, 2000).

5. Evaluation & Results

For the evaluation, we first examined the reliability of the constructs we measured in our study (see Table 4). We calculated Cronbach's alpha (α) for each construct (Cronbach, 1951). Overall, we can report that all but one of our constructs passed the goodness criterion of $\alpha > 0.7$ (Cortina, 1993). The only construct that did not pass the reliability test is effort/importance with a score of 0.67 just below the threshold of 0.7 (Ryan et al., 1983). We can only speculate as to the reason for this, but it may have been due to our experimental setting where participants had no real stake in the game and thus perceived the interaction as less important.

The goal of our research is to investigate how learners can be supported and motivated by a gamified chatbot. In the context of our experimental setting, we used the number of lessons completed as a proxy variable for the intensity of the interaction and thus the motivation of the participants. We argue that the more the participants perceive our artifact as supportive and the more motivated they are, the more optional lessons they will complete. Therefore, we conduct a regression analysis with lessons completed as the target variable for all our constructs. Overall, we can report a result of F (16, 147) = 2.70 with $p < 0.001$ and $R^2 = 0.23$ for our regression analysis for the number of lessons completed. In addition, we have run a regression analysis for the constructs of **PU** and **BITU**. For **PU** we can report a result of F (14, 149) = 38.50 with $p < 0.001$ and $R^2 = 0.78$ which represents a slight positive correlation between our gamified chatbot design and the perceived usefulness of interacting with it.

Table 4. Constructs and reliability.

Construct (source)		**α**
Interest/enjoyment (Ryan et al., 1983)	IE	0.89
Pressure/tension (Ryan et al., 1983)	PT	0.73
Effort/importance (Ryan et al., 1983)	EI	0.67
Competence (Ryan & Deci, 2000b)	CT	0.85
Autonomy (Ryan & Deci, 2000b)	AT	0.71
Relatedness (Ryan & Deci, 2000b)	RT	0.83
Joy (Agarwal & Karahanna, 2000)	JY	0.89
Control (Agarwal & Karahanna, 2000)	CR	0.88
Focused immersion (Agarwal & Karahanna, 2000)	FI	0.81
Temporal dissociation (Agarwal & Karahanna, 2000)	TD	0.91
Curiosity (Agarwal & Karahanna, 2000)	CU	0.87
BITU (Venkatesh & Davis, 2000)	IU	0.93
PU (Venkatesh & Davis, 2000)	PU	0.95
PEOU (Venkatesh & Davis, 2000)	EU	0.80

Furthermore, we found that the subconstructs of **joy (p = 0.004)** and **curiosity (p = 0.016)**, which underlie the construct of cognitive absorption, have a significant positive influence on PU. Although technically not statistically significant according to strict criteria, we can report a decent loading for focused immersion (p = 0.079) which is also part of cognitive absorption. We found similar results with our regression analysis for BITU. We can report a result of F (15,148) = 39.81 with $p < 0.001$ and R^2 = 0.80, which translates to a strong significant effect on BITU. Similar to PU, we find a significant influence for the construct of cognitive absorption and it's subconstructs of **joy (p = 0.016)** and **focused immersion (p = 0.003)**. In addition, interest/enjoyment (p = 0.098) shows a decent significance, although it does not meet the strict 0.05 significance criterion. Overall, we cannot report a sufficiently significant correlation between gamification design for chatbots in education and lessons taught. However, we can report a significant influence of gamified chatbots on the cognitive absorption of the participants in the interaction, especially in relation to PU and BITU.

5. Discussion of Findings, Contributions, and Future Research

The goal of our research was to analyze the effects of a gamified chatbot on learner motivation and behavior. To summarize our findings and guide future research, we provide design recommendations (DR#) and discuss the results of our work.

Based on our statistical analysis, we can conclude that there is some correlation between our gamified chatbot, the number of lessons completed and thus the motivation and engagement of learners. Overall, only

28.04% of users abandoned our bot after completing the mandatory lessons while 71.96% of users chose to continue the interaction. Based on our results the gamified design of the educational chatbot may have a particularly relevant impact on the cognitive absorption of learners. While this may not directly translate into motivation and doing more lessons, it does translate to PU and a BI to use the chatbot. The significant result for particularly focused immersion is reflected in our participants' decision to voluntarily stay engaged with the gamified chatbot and continue with the optional lessons.. This suggests that gamified chatbots help learners to stay immersed and focused on learning, rather than losing attention or motivation. This could help to combat procrastination of learners and support learners in their self-regulated learning (Klassen et al., 2008; Senécal et al., 1995).Thus, resulting in an improved support for learners and potentially increased motivation to learn, which could then be reflected in the overall performance of learners. Accordingly, we propose the following design recommendation:

DR1: *Gamified chatbots can increase learner engagement and motivate them to do more self-directed learning.*

In addition, we need to discuss the statistical reality of our analysis and think critically about its implications. For results that do not meet a higher level of significance or for the number of lessons completed, we cannot provide a clear explanation at this time.. However, we did ask our participants two additional questions about why they stopped or continued interacting with our gamified learning chatbot. We asked our participants why they stopped interacting with the chatbot after completing the two mandatory lessons. 27% of our participants said that they weren't interested in the chatbot or weren't having enough fun. Given that we only implemented two game elements, we may need to revise this design to be more diverse, as it may lack variety in gamification designs (Groening & Binnewies, 2021).

Similar to the uncanny valley, there may be a point where too many game elements in a dialog-based interaction with a chatbot are experienced as distracting (Groening & Binnewies, 2021). This could, for example, lead to higher cognitive load and less effective outcomes (Liao et al., 2019). Research has already discussed the number of elements that are useful in combination (Faiella & Ricciardi, 2015); however, in a more interactive and different type of learning process – game elements may be experienced differently. Moreover, some learners might prefer some game design elements over others, while other learners might prefer other elements. This highlights the potential research opportunity to explore more gamification designs for chatbots. Accordingly, we can make the following recommendation.

DR2: *Gamification design for chatbots should include a variety of game design elements that respond positively to as many learners as possible.*

Thus, the number and composition can therefore always be a limiting factor but also serve as foundation for future research. However, due to the number of possible elements and combinations, researchers need to focus on the most promising ones. Too many elements can be harmful and have negative effects because they are experienced as too overwhelming. We argue that such a choice needs to be made and suggest some options in this regard. For example, gamification could be tailored to the personality, learning style or player type, e.g. preferred game elements according to the personality of the learners (Leung et al., 2022; Passalacqua et al., 2021; Tondello et al., 2017). Consequently, it may be a worthwhile effort for future research to figure out how many game elements in what combination are useful. Additionally, an investigation into competitive elements, such as rankings or social comparison may be worthwhile since we have not implemented such game design elements.

Furthermore, our participant responses indicate that 21% of participants did not find enough individual use or benefit from the interaction. This could be explained by the very high number of participants who had already completed their driving examination and were not taking another class of driving license. This finding may be reflected in the results of our regression analysis where competency was negatively correlated with the number of lessons completed. A possible explanation could be that participants either felt insufficiently competent or perceived the difficulty of the questions as too hard or too easy. To better inform the design of gamified chatbots, research could ground the design of elements on self-determination theory (Deci & Ryan, 2012). We grounded our design in autonomy. Due to our experimental setting, an additional 20% of participants reported time pressure or constraints i.e., they did not want to spend more time than necessary to complete the experiment. Although we did not specifically subject our participants to any time pressure, this should be kept in mind as it can have a significant negative impact on learners (Ackerman & Lauterman, 2012). With this in mind, future research would need to re-examine our approach in the field with an actual user group that has an actual stake in the game. Beyond these reasons, the remaining 32% of participants reported that the questions were too easy, they felt they did not need to familiarize themselves more with the chatbot or they experienced technical

difficulties. While these findings may largely be a result of our experimental setting, they may have potential implications as well. For instance, our applied lessons may not have provided enough cognitive stimulation for our participants. However, overly difficult tasks should also be avoided as they may overwhelm learners. Additionally, this may be further influenced by the application of gamification and the chatbots themselves (Liao et al., 2019). Therefore, we suggest the following recommendation for further analysis:

DR3: *To support the effectiveness of educational gamified chatbots, it is necessary to identify the right balance between the difficulty of the questions asked and the complexity of the gamification applied to chatbots.*

The second question, we asked our participants was why they actively chose to continue interacting with the chatbot. The vast majority, 34% of participants, said they were curious and had fun interacting with the chatbot. In second place, 30% of participants who continued the interaction said they were interested in testing their knowledge, even though they did not need to learn theory for a driving test. In third place, 19% of participants said that they liked the functionality of the chatbot, especially the gamification mechanisms. The remaining 17% of participants said that they simply forgot to end the interaction or did not find the option to do so. Given these results, the nearly 20% of users who essentially did not want to continue the interaction may have had a negative impact on our analysis. Nevertheless, the majority of participants continued the lessons voluntarily. This is reflected in our findings regarding joy which is related to participants having fun interacting with our gamified chatbot and curiosity which is related to our participants expressing interest in our gamified chatbot. Thus, we can conclude that our gamified chatbot promotes interest and a joyful learning experience. Therefore, we propose the following recommendation:

DR4: *Interacting with a gamified chatbot makes learning more engaging, helps learners stay focused, and can improve learning outcomes.*

Overall, when looking at the results of our regression analysis, we find that the statistical numbers do not paint the whole picture. According to our two additional questions, we find that the majority of users actually chose to stay engaged with our chatbot, and that the majority of these users found the interaction to be supportive and motivating. Therefore, we find support for our DR; however, we also like to emphasize the potential for future research to expand on our findings.

6. Limitation

Our research is not without limitations. First, we based our evaluation on a 2x2 between-subjects design, which allows for a limited number of design choices for treatment combinations. While we based our choices on the relevant literature, other authors may have come up with a different design and thus different results. Therefore, investigating different approaches to implementing gamification may be interesting for fellow researchers.

Second, while experimental settings such as ours may have strong internal validity, they may lack external validity (Bagozzi, 1978). Because of our selected sample and use case, participants may have no stake in the interaction and thus negatively influence the outcome of the experiment. In this regard, we have presented our two additional questions that can help shed light on this issue. We have shown and explained why participants chose to continue or terminate the interaction, but the limitation remains. Furthermore, although our sample is not exclusively university students, we have a bias towards younger people with a university background, which may have influenced our results. A sample with bias towards younger people in the context of driving examination may arguably not be that biased after all, since most people do their licenses under 30 years of age. However, considering the overall generalization of our research this remains a limitation.

Overall, we have to acknowledge these limitations but also want to emphasize on the previously discussed potentials for future research. Our study may present a foundation for fellow researchers to expand into various directions that could cover our existing limitations, such as conducting a field study under real-world conditions or including a more diverse sample.

7. Conclusion

Educational chatbots have the potential to motivate and support learners in a variety of ways, thereby contributing to learner success (Janson et al., 2023; Schöbel et al., 2021; Yin et al., 2021). In order to realize this potential, a variety of design decisions must be made when developing educational chatbots, including their gamification design. Therefore, in this work, we investigated how gamified chatbots can support and motivate learners through the use of badges and progress bars. Thus, we presented a gamification design for educational chatbots in the existing literature and developed a prototype. Then, we conducted an online experiment and questionnaire to evaluate our gamified bot, for which we performed

a regression analysis. Our results show that there is a tendency in favor of our gamified chatbot approach to be supportive and motivating for learners. Regarding H1 and H2, we can conclude that gamified chatbots can have a positive impact on supporting and motivating learners. In particular, we can report an impact on cognitive absorption, i.e., supporting learners to stay motivated to engage more intensely in learning. Based on these findings, we have made four recommendations for the design of gamified chatbots, which have potential for future research.

However, we must also to acknowledge that our results are not as statistically significant as we would have hoped for. In this context, we have discussed potential shortcomings and reasons for these results in detail. Considering that gamification in general has already been demonstrated to work (e.g., Koivisto & Hamari, 2019; Majuri et al., 2018), we conclude that more research is needed to explain why results may vary in a conversational context. Therefore, we argue that more research is needed on the gamification of conversational artifacts such as chatbots to shed light on potential differences and pitfalls where gamification is otherwise expected to yield significant results. The gamification of chatbots in education remains an interesting and worthwhile research topic that has not yet been systematically applied (reference removed for review). Nevertheless, we observed that game elements can enrich the interactive learning experience with a chatbot, as over 70% of participants chose to continue the interaction, and over 50% of those who did so had fun or found the bot supportive of their learning, i.e., were motivated to interact. Thus, gamification of chatbots in education can still offer great potential to increase learners' motivation and support them in their learning efforts.

Acknowledgements

The research presented in this paper is partially funded by the *Stiftung Innovation* in der Hochschule within the project "Universität Kassel digital: Universitäre Lehre neu gestalten" and the *Flagship by Innosuisse* project (Flagship PFFS-21-29).

We also want to thank Felix Hupfeld for his effort implementing the prototype and collecting data.

References

Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior*, *28*(5), 1816–1828.

Agarwal, R., & Karahanna, E. (2000). Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. *MIS Quarterly*, *24*(4), 665.

Alaimi, M., Law, E., Pantasdo, K. D., Oudeyer, P.-Y., & Sauzeon, H. (2020). Pedagogical Agents for Fostering Question-Asking Skills in Children. *CHI Conference on Human Factors in Computing Systems*.

Allameh, M., & Zaman, L. (2021). Jessy: A Conversational Assistant for Tutoring Digital Board Games. *Annual Symposium on Computer-Human Interaction in Play*, 168–173.

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.

Bagozzi, R. P. (1978). The Construct Validity Of The Affective, Behavioral, And Cognitive Components Of Attitude By Analysis Of Covariance Structures. *Multivariate Behavioral Research*, *13*(1), 9–31.

Benner, D., Schöbel, S., Janson, A., & Leimeister, J. M. (2022). How to Achieve Ethical Persuasive Design: A Review and Theoretical Propositions for Information Systems. *AIS Transactions on Human-Computer Interaction*, *14*(4), 548–577.

Benner, D., Schöbel, S., & Süess, C. (2022). Towards Gamified Conversational Agents for Self-Regulated Learning in Digital Education. *International Conference on Information Systems*.

Benner, D., Schöbel, S., Süess, C., Baechle, V., & Janson, A. (2022). Level-Up your Learning - Introducing a Framework for Gamified Educational Conversational Agents. *International Conference on Wirtschaftsinformatik*.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Davis, K., & Singh, S. (2015). Digital badges in afterschool learning: Documenting the perspectives and experiences of students and educators. *Computers & Education*, *88*, 72–83.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227–268.

Deci, E. L., & Ryan, R. M. (2012). Self-Determination Theory. In P. van Lange, A. Kruglanski, E. Higgins, & P. A. van Lange (Eds.), *Handbook of theories of social psychology* (pp. 416–437). SAGE.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification". *International Academic MindTrek Conference Envisioning Future Media Environments*, 9–15.

Deterding, S., Sicart, M., O'Hara, K., Dixon, D., & Nacke, L. (2011). *Gamification: Using Game Design Elements in Non-Gaming Contexts*. ACM.

Diederich, S., Brendel, A. B., 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*, *23*(1), 96–138.

Fadhil, A., & Villafiorita, A. (2017). An Adaptive Learning with Gamification & Conversational UIs. *Conference on User Modeling, Adaptation and Personalization*, 408–412.

Faiella, F., & Ricciardi, M. (2015). Gamification and learning: A review of issues and research. *Journal of E-Learning and Knowledge Society*, *11*(3), 13–21.

Fidan, M., & Gencel, N. (2022). Supporting the Instructional Videos With Chatbot and Peer Feedback Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Motivation. *Journal of Educational Computing Research*, 1–26.

Giboney, J., Schuetzler, R., Grimes, M., & Dayley, I. (2021). A Social Engineering Victim Chatbot to Teach Adversarial Thinking. *Americas Conference on Information Systems*.

Groening, C., & Binnewies, C. (2021). The More, the Merrier? - How Adding and Removing Game Design Elements Impact Motivation and Performance in a Gamification Environment. *International Journal of Human–Computer Interaction*, *37*(12), 1130–1150.

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

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, *71*, 469–478.

Hamari, J., & Koivisto, J. (2013). Social Motivations To Use Gamification: An Empirical Study Of Gamifying Exercise. *European Conference on Information Systems*.

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, *54*, 170–179.

Hobert, S. (2019). Say Hello to 'CodingTutor'! Design and Evaluation of a Chatbot-based Learning System Supporting Students to Learn to Program. *International Conference on Information Systems*.

Hobert, S., & Meyer von Wolff, R. (2019). Say Hello to Your New Automated Tutor – A Structured Literature Review on Pedagogical Conversational Agents. *International Conference on Wirtschaftsinformatik*, 301–314.

Janson, A., Schmidt-Kraepelin, M., Schöbel, S., & Sunyaev, A. (2023). Special Issue Editorial: Adaptive and Intelligent Gamification Design. *AIS Transactions on Human-Computer Interaction*, *15*(2), 136–145.

Janson, A., Söllner, M., & Leimeister, J. M. (2020). Ladders for Learning: Is Scaffolding the Key to Teaching Problem-Solving in Technology-Mediated Learning Contexts? *Academy of Management Learning & Education*, *19*(4), 439–468.

Katchapakirin, K., Anutariya, C., & Supnithi, T. (2022). ScratchThAI: A conversation-based learning support framework for computational thinking development. *Education and Information Technologies*, 1–28.

Klassen, R. M., Krawchuk, L. L., & Rajani, S. (2008). Academic procrastination of undergraduates: Low self-efficacy to self-regulate predicts higher levels of procrastination. *Contemporary Educational Psychology*, *33*(4), 915–931.

Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, *45*, 191–210.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice*, *41*(4), 212–218.

Lee, M., Noortman, R., Zaga, C., Starke, A., Huisman, G., & Andersen, K. (2021). Conversational Futures: Emancipating Conversational Interactions for Futures Worth Wanting. *CHI Conference on Human Factors in Computing Systems*.

Leung, A. C. M., Santhanam, R., Kwok, R. C.-W., & Yue, W. T. (2022). Could Gamification Designs Enhance Online Learning Through Personalization? Lessons from a Field Experiment. *Information Systems Research*.

Liao, C.-W., Chen, C.-H., & Shih, S.-J. (2019). The interactivity of video and collaboration for learning achievement, intrinsic motivation, cognitive load, and behavior patterns in a digital game-based learning environment. *Computers & Education*, *133*, 43–55.

Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*(140), 5–55.

Liu, D., Santhanam, R., & Webster, J. (2017). Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. *MIS Quarterly*, *41*(4), 1011–1034.

Majuri, J., Koivisto, J., & Hamari, J. (2018). Gamification of education and learning: A review of empirical literature. *GamiFIN*, 11–19.

Mazarakis, A., & Bräuer, P. (2020). Gamification of an open access quiz with badges and progress bars: An experimental study with scientists. *GamiFIN*.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872.

Passalacqua, M., Sénécal, S., Frédette, M., Nacke, L. E., Pellerin, R., & Léger, P.-M. (2021). Should Gamification be Personalized? A Self- deterministic Approach. *AIS Transactions on Human-Computer Interaction*, *13*(3), 265–286.

Pereira, J. (2016). Leveraging chatbots to improve self-guided learning through conversational quizzes. *International Conference on Technological Ecosystems for Enhancing*, 911–918.

Pereira, J., & Díaz, Ó. (2021). Struggling to Keep Tabs on Capstone Projects: A Chatbot to Tackle Student

Procrastination. *ACM Transactions on Computing Education*, *22*(1), 1–22.

Pham, X. L., Pham, T., Nguyen, Q. M., Nguyen, T. H., & Cao, T. T. H. (2018). Chatbot as an Intelligent Personal Assistant for Mobile Language Learning. *International Conference on Education and E-Learning*, 16–21.

Ruan, S., He, J., Ying, R., Burkle, J., Hakim, D., Wang, A., Yin, Y., Zhou, L., Xu, Q., AbuHashem, A., Dietz, G., Murnane, E. L., Brunskill, E., & Landay, J. A. (2020). Supporting Children's Math Learning with Feedback-Augmented Narrative Technology. *Proceedings of the Interaction Design and Children Conference*, 567–580.

Ruan, S., Jiang, L., Xu, J., Tham, B. J.-K., Qiu, Z., Zhu, Y., Murnane, E. L., Brunskill, E., & Landay, J. A. (2019). QuizBot: A Dialogue-based Adaptive Learning System for Factual Knowledge. *CHI Conference on Human Factors in Computing Systems*, Article 357.

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67.

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, *45*(4), 736–750.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380.

Santhanam, R., Liu, & Shen, W.-C. M. (2016). Research Note—Gamification of Technology-Mediated Training: Not All Competitions Are the Same. *Information Systems Research*, *27*(2), 1–13.

Schlegel, L., Schöbel, S., & Söllner, M. (2023). Nudging Digital Learning - An Experimental Analysis of Social Nudges to Manage Self-Regulated Learning and Online Learning Success. *Hawaii International Conference on System Sciences*.

Schöbel, S., Janson, A., & Söllner, M. (2020). Capturing the complexity of gamification elements: a holistic approach for analysing existing and deriving novel gamification designs. *European Journal of Information Systems*, *29*(6), 641–668.

Schöbel, S., Saqr, M., & Janson, A. (2021). Two decades of game concepts in digital learning environments – A bibliometric study and research agenda. *Computers & Education*, *173*, 104296.

Schöbel, S., Schmitt, A., Benner, D., Saqr, M., Janson, A., & Leimeister, J. M. (2023). Charting the Evolution and Future of Conversational Agents: A Research Agenda Along Five Waves and New Frontiers. *Information Systems Frontiers*.

Senécal, C., Koestner, R., & Vallerand, R. J. (1995). Self-Regulation and Academic Procrastination. *The Journal of Social Psychology*, *135*(5), 607–619.

Tegos, S., & Demetriadis, S. (2017). Conversational Agents Improve Peer Learning through Building on Prior Knowledge. *Journal of Educational Technology & Society*, *20*(1), 99–111.

Tian, X., Risha, Z., Ahmed, I., Lekshmi Narayanan, A. B., & Biehl, J. (2021). Let's Talk It Out: A Chatbot for Effective Study Habit Behavioral Change. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), Article 97, 1–32.

Tondello, G. F., Orji, R., & Nacke, L. E. (2017). Recommender Systems for Personalized Gamification. *Conference on User Modeling, Adaptation and Personalization*, 425–430.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., & Plattfaut, R. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems*, *37*, Article 9.

Waldner, S., Seeber, I., Waizenegger, L., & Maier, R. (2022). How do Pedagogical Conversational Agents affect Learning Outcomes among High School Pupils: Insights from a Field Experiment. *Hawaii International Conference on System Sciences*, 408–417.

Wambsganss, T., Kueng, T., Söllner, M., & Leimeister, J. M. (2021). ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills. *CHI Conference on Human Factors in Computing Systems.* Advance online publication.

Weber, F., Wambsganss, T., Rüttimann, D., & Söllner, M. (2021). Pedagogical Agents for Interactive Learning: A Taxonomy of Conversational Agents in Education. *International Conference on Information Systems*.

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*, 146–156.

Winkler, R., Hobert, S., Fischer, T., Salovaara, A., Söllner, M., & Leimeister, J. M. (2020). Engaging Learners in Online Video Lectures with Dynamically Scaffolding Conversational Agents. *European Conference on Information Systems*.

Winkler, R., & Söllner, M. (2020). Towards Empowering Educators to Create their own Smart Personal Assistants. *Hawaii International Conference on System Sciences*, 22–31.

Xi, N., & Hamari, J. (2019). Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management*, *46*, 210–221.

Yin, J., Goh, T.-T., Yang, B., & Xiaobin, Y. (2021). Conversation Technology With Micro-Learning: The Impact of Chatbot-Based Learning on Students' Learning Motivation and Performance. *Journal of Educational Computing Research*, *59*(1), 154–177.