

To Add or Not to Add Game Elements? Exploring the Effects of Different Cognitive Task Designs Using Eye Tracking

Manuel Ninaus, Kristian Kiili, Guilherme Wood, Korbinian Moeller, and Silvia Erika Kober

Abstract—Research on instructional design provides inconsistent results on the use of game elements in cognitive tasks or learning. Cognitive load theory suggests that game elements increase extraneous cognitive load and thus may distract the users. In contrast, from an emotional design perspective, the use of game elements is argued to increase performance by providing a more interesting and motivating task environment. To contribute to this debate, the current study investigated the effect of game elements on behavioral performance, attention, and motivation. We designed two versions of the number line estimation task—one with game elements and one without. Participants completed both versions of the task while their eye-fixation behavior was recorded. Results indicated that participants paid attention to game elements, that is, they fixated them, although they were not necessary to complete the task. However, no difference in estimation accuracy was observed between the two task versions. Moreover, the task version with game elements was rated to be more attractive, stimulating, and novel, and participants reported experiencing greater flow. In sum, these data indicate that game elements seem to capture attention but also increase motivational aspects of learning tasks rather than decreasing performance.

Index Terms—Educational games, emotional design, eye tracking, game elements, number line estimation, seductive details.

I. INTRODUCTION

GAME-BASED learning and assessment have become increasingly popular over the last years due to their potential of increasing performance and motivation [1]. Consequently, researchers and practitioners alike often attempt to augment conventional tasks with game elements to enhance performance of young students [2]. However, designing a task in a way that it becomes gamified or even game-like requires the addition of details, game mechanics, or game elements, respectively, that might not be strictly relevant for achieving the overall objective of a given task (e.g., adding a fictitious narrative or virtual incentives). Plass, Homer, and Kinzer [3]

consider game mechanics as design elements that can be used to influence cognitive, behavioral, affective, and social engagement. Further, core game mechanics are activities that players repeat over and over throughout the game [4]. Pawar, Tam, and Plass [5] distinguish game mechanics from learning mechanics. While game mechanics constitute major elements of play activities, learning mechanics describe major elements of learning activities. Accordingly, to design effective and engaging educational games, the core game mechanics should be well integrated with learning content, learning mechanics, and instructional aspects [5]–[7]. Inadequate integration may lead to gameplay in which player’s interaction with the game mechanics and related game elements does not support learning, but may cause unnecessary extraneous cognitive load undermining possibilities to learn.

According to cognitive theories of game-based learning, player’s information processing capacity is limited and it can be allocated among extraneous processing, essential processing, and generative processing [8]. Extraneous processing refers to cognitive processing that does not support the instructional objective of the game. In contrast, essential processing is needed to represent the visual and verbal material of the game in working memory and generative processing is needed for making sense of the material. More practically put, the poor use of game elements may lead to situations in which players have to use most of their cognitive capacity for extraneous processing and thus they do not have enough cognitive resources for essential and generative processing needed for learning [8]. That is, poor educational game design may actually decrease performance or learning as irrelevant details, often referred to as seductive details in the domain of multimedia learning [9], may distract players’ attention from the relevant features of the task. In contrast, it is argued that game elements aim at increasing cognitive and emotional engagement and thereby increase motivation and performance [3], [10]. The disparity between these two instructional approaches might also explain rather heterogeneous but still overall positive effects found for game-based learning [11], [12]. As a consequence, the use of game elements in

Manuscript received February 19, 2020; revised July 24, 2020; accepted XXX XX, XXX. Date of publication XXXX XX, XXXX. (Corresponding author: Manuel Ninaus.)

M. Ninaus was with the Leibniz-Institut für Wissensmedien, Tübingen, 72076 Germany. He is now with the University of Innsbruck, Innsbruck, 6020 Austria (e-mail: manuel.ninaus@uibk.ac.at).

K. Kiili is with the Tampere University, Tampere, FI-33014 Finland (e-mail: kristian.kiili@tuni.fi).

G. Wood is with the University of Graz, Graz, 8010 Austria (e-mail: guilherme.wood@uni-graz.at).

K. Moeller is with the Loughborough University, Loughborough, E20 3BS United Kingdom (e-mail: K.Moeller@lboro.ac.uk).

S. E. Kober is with the University of Graz, Graz, 8010 Austria (e-mail: silvia.kober@uni-graz.at).

instructional materials and cognitive assessments is still a matter of debate [13] and needs to be investigated further. More importantly, the question remains, in which way game elements influence attentional processes as well as motivational states such as flow experience.

Therefore, in the current exploratory study we evaluated participants' eye fixation behavior to investigate attentional processes in a math task. We employed a number line-based fraction task and its game-based equivalent to experimentally investigate the effects of game elements on three different levels: (i) behavioral performance, (ii) attention distribution as reflected by participants eye fixation behavior as well as (iii) motivational aspects. Based on these levels, three research questions were formulated.

- 1) Does performance in a task with and without game elements differ?
- 2) Do game elements alter selective attention as indicated by eye fixation behavior in a task?
- 3) Does the experienced flow and user experience in a task with and without game elements differ?

Accordingly, in the following, we will first give a brief overview on the different rationales arguing for and against augmenting cognitive tasks to alter performance with game elements. After that, we briefly elaborate on the consideration of eye fixation behavior in multimedia research and discuss studies utilizing eye tracking to investigate attentional effects of game elements or the lack thereof, respectively, before addressing motivational aspects such as flow.

II. RELATED WORK

A. Theoretical Contradictions—Minimalistic or Augmented Learning Materials

The use of games or game-like experiences in education is the opposite of what some researches in the learning sciences would recommend [14]. For instance, the cognitive theory of multimedia learning [15] would postulate that embellishing game elements would impede educational efficacy as they distract users from the actual content or task objective. According to the coherence principle, any illustrations or content that is not fundamentally important or relevant should be removed to facilitate reaching the instructional goal [15], [16]. In this vein, research in the field of seductive details indicated that adding illustrations to instructional material such as textbooks that are not relevant for reaching the instructional goal can lead to poorer performance or outcomes, respectively [9], [17]–[19]. Following this theory, instructional material should be minimalistic to avoid or reduce extraneous cognitive load [20] as these irrelevant details occupy working memory with unnecessary information. In line with this, the distraction hypothesis predicts that seductive details direct selective attention away from relevant information [19] as they are selectively processed and remembered at the expense of more important but less interesting information. Consequently, adding rather irrelevant, but interesting and thus attention capturing elements such as game narrative or visually appealing game graphics, might reduce performance of users as they distract them from more relevant task elements [19], [21], [22].

Recently, the approach on seductive details has been criticized. Critics argued that studies on the seductive details

effect showed rather unconvincing and contradictory evidence. It is discussed that seductive details do not always impede understanding and that they can even motivate users [23]–[26]. Park *et al.* [24] aimed at clarifying this controversy by linking positive/negative effects of seductive details to cognitive load theory. High school students learned about biology using a multimedia environment with and without seductive details while also varying cognitive load (i.e., by altering the modality of verbal information). Results showed that students' performance in the multimedia environment was significantly higher when seductive details were present in a low cognitive load environment as compared with all other conditions [24].

Similarly, beyond a pure cognitive perspective, Schnotz *et al.* [27] further assumed that seductive details may increase interest in the instructional material leading to higher persistence and increased performance. Augmenting instructional material with interesting details might therefore affect users' willingness to invest cognitive effort. Consequently, performance might be affected by the users' perception of the task, respectively. This argument is based on the assumption that cognitive resource allocation is flexible and positively affected by users' motivational state [28]–[30]. In line with this, different frameworks and theories such as the social-cognitive control value theory of achievement emotions [31], the cognitive-affective theory of learning with media [32], or the integrated cognitive affective model of learning with multimedia [33] highlight the interdependency of emotion, motivation, and cognitive performance. Accordingly, in contrast to rather minimalistic instructional design perspectives, the emotional design perspective suggests that seductive details might be helpful for increasing performance in instructional settings by increasing motivation of users [30], [34]. For instance, Um *et al.* [30] observed increased motivation and invested mental effort, when applying emotional design principles to instructional materials as compared to a neutral condition with presumably lower extraneous cognitive load.

Accordingly, it has been hypothesized that game elements such as a narrative are useful for improving motivation and performance [35], [36]. In fact, recent meta-analyses on game-based learning identified increased performance as compared to conventional or non-game-based instructional conditions [11], [12]. Similar effects can be found for game-based cognitive assessment and training [37], [38]. Unfortunately, underlying mechanisms on the integration of game elements in a regular cognitive task have not been researched intensively so far. Hence, it is hard to pinpoint particular effects on why game-based learning seems to be superior to conventional approaches [12] even though it directly opposes assumptions posed by the cognitive load theory or cognitive theory of learning with multimedia.

B. Attention and Eye Tracking

Seductive details, such as pictures or emotional materials have been shown to attract users attention [39]–[41]. In this context, eye tracking offers a direct way of identifying whether seductive details or game elements catch attention as reflected by being looked at in an instructional task—probably diverting attention away from more fundamental content or task objectives [42]. Other ways of assessing attention towards game elements or seductive details are neurofunctional measures,

which can only be used in specific and highly controlled scenarios, or self-reports which can be affected, for instance, by inaccurate recall or memory effects [43], [44]. Eye tracking provides an objective and rather unobtrusive way with high temporal resolution to investigate cognitive processes in learners. Consequently, over the past decade, the number of eye-tracking studies in multimedia learning [45], [46] and educational game research has increased [47]–[53].

Previous research showed that eye fixation behavior, especially fixations representing areas where the eye fixates at on the computer monitor, are valid indicators of visual attention during gaming [46], [54]–[56]. This is in line with the eye-mind hypothesis predicting that the focus of human gaze indicates the focus of attention [55], [57]. Although, it is assumed that information at the point a person is currently looking at is attended and processed cognitively [48], [55], [57]. Hyönä [57] points out that this assumption holds only if the visual information is relevant to the task at hand. Accordingly, fixations are the main measures in eye-tracking research, which allow for a number of useful metrics to be derived, for instance, the number of fixations and (mean) fixation duration [58]. For example, Tsai, Huang, Hou, Hsu, and Chiou [47] employed eye tracking to study users' cognitive processes in a dynamic and highly-interactive game-based learning environment. They showed that fixation patterns may reflect users' metacognitive control over visual attention in game-based environments.

Interestingly, fixation measures seem to be associated with task performance [46], [54], [55]. There is some evidence that high performers or experts show less fixations than low performers [59], [60], which might indicate more efficient and focused processing. Moreover, it was argued that longer fixation durations indicate higher cognitive effort reflecting higher task difficulty and cognitive workload but also deeper processing [61]–[63]. Therefore, eye tracking might be a suitable method to evaluate the underlying attentional demands and affordances of popular game elements, such as narrative, virtual incentives, appealing visual aesthetics, etc. [3].

Empirical studies comparing game-based to non-game-based equivalent tasks are still sparse [64], but studies investigating whether game-elements indeed divert attention away from fundamentally important task elements are even rarer. For instance, Kiili *et al.* [48] used eye-tracking measures to explore the perception of user interfaces of four different educational games, but did not compare them to a non-game-based equivalent. They found that extraneous elements in the educational games—especially animated content displayed concurrently to feedback—indeed distracted users and consequently disturbed the instructional process. Moreover, it was also suggested that low performers and inattentive players might be more distracted by such extraneous gaming elements than high performers and highly attentive players, because the former are probably less able to distinguish between important and irrelevant information [48], [52].

C. Motivation and Flow

Besides effects on attention and performance, game elements reportedly affect motivational states, for instance by including a narrative or game fiction [10], [65]. In fact, several meta-analyses indicated that game elements improve motivation [11], [66]. In the domain of game-based learning, the concept of flow

is one of the most popular motivational constructs assessed [67], [68], but in the context of seductive details the research on flow is sparse. To our knowledge, only one study examined differences in flow experience between an interactive simulation on beer brewing with and without game-like elements [69]. The authors, however, did not identify differences between the groups on flow and performance.

Flow is usually considered to be a positive emotional state [70], [71] and an holistic approach to motivation [72]. However, Landhäußer and Keller [73] noted that flow is not the same as having fun, but rather a state characterized by a combination of aspects such as concentration, a merging of action and awareness, reduced self-consciousness, a sense of control, a transformation of time, and an intrinsically rewarding experience. Moreover, Landhäußer and Keller [73] distinguishes flow preconditions (i.e., a skills-demands fit, clear goals, and clear feedback) that foster involvement and enjoyment in the activity. These preconditions provide meaningful recommendations for augmenting tasks with game elements. Game elements, for instance, might be used to support formation of self-regulated goals, provide feedback about player's progress towards goals, and adapt challenges of the game to player's skill level.

Importantly, flow experience also seems to be positively related with learning outcomes and playing performance in game-based learning [68], [74]–[76] and is often related to positive user experience [68], [77]. As such, good (educational) games are often described as being able to hit the “sweet spot” of flow [3] and users experiencing flow usually are willing to extend the interaction with the game in the future and rate games better [70]. Regarding the latter and in the context of seductive details, it is important to assess whether the added seductive details or game elements, respectively, are experienced by the users in a positive way (e.g., the attractiveness of the design: [78]). This is particularly important as seductive details are supposed to draw users attention towards the instructional material [79]. Hence, the relation between flow, user experience, and seductive details needs to be further investigated.

III. METHODS

A. Participants

Forty-two university students (*mean* age 21.26 years, *SD* = 1.93; 18 male) participated in this study. All participants had normal or corrected-to-normal vision (contact lenses but no glasses). Participants gave written informed consent. The study was approved by the Ethics Committee of the University of Graz, Austria (reference number GZ. 39/86/63 ex 2017/18) and is in accordance with the ethical standards of the Declaration of Helsinki.

B. Task Versions—Game vs. No-Game

We employed the so called number line estimation task, which relies on the concept of a mental number line, an often used metaphor to describe our mental representation of number magnitude [80]. All participants performed number line estimation tasks in both game and no-game conditions. In the number line estimation tasks participants had to locate fractions on a number line ranging from 0 to 1, while their eye-fixation

behavior was recorded. In each estimation condition (game vs. no-game) 48 proper fractions involving single- (e.g., $3/7$) and double-digit (e.g., $19/25$) numerators and denominators were used. In the game condition, number line estimation was embedded in an educational game called Semideus [81], [82], whereas the no-game condition reflected a stripped version of the task without any game elements. Game and no-game conditions were comparable regarding task difficulty as they both involved the same target numbers as well as same number of levels (i.e., 6 levels $\hat{=}$ 8 tasks/fractions). The order of the conditions was counterbalanced across all participants.

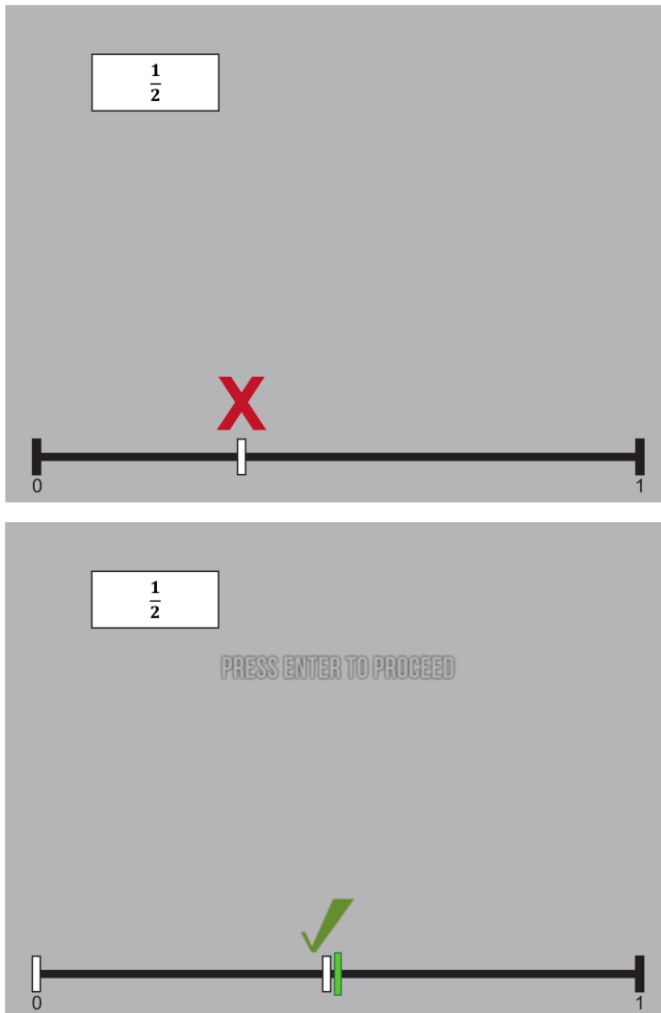


Fig. 1. Exemplary screenshots of the no-game condition; upper chart: negative feedback; lower chart: positive feedback.

The implementation of the no-game version of the estimation task was minimalistic (see Fig. 1) as no game elements were present and thus corresponds to the conventional number line estimation task. A target fraction to-be-located on the number line was displayed in the upper left part of the screen and the number line on bottom of the screen. The participants could move a cursor (white bar) on the number line with the left and right arrow keys and confirm their estimates by pressing the space bar on the keyboard. Feedback was provided after each estimation. Positive feedback was given by showing a green check mark above the cursor if the target fraction was estimated accurately enough (i.e., estimated location no more than $\pm 5\%$

away from the correct location) and at the same time the absolute correct location was shown by a green marker on the number line. Negative feedback was given in the form of a red cross in case the estimated location was more than $\pm 5\%$ away from the correct location. The same negative feedback mechanic was shown if an estimation took too long (>10 seconds per estimate). In case the estimation was incorrect or too slow, participants again had 10 seconds to indicate the location of the fraction. The number of trials on each task was not limited and participants had to estimate the location of the fraction correctly before they could proceed to the next task. After correct estimates, participants were asked to press the Enter-key to proceed to the next task.

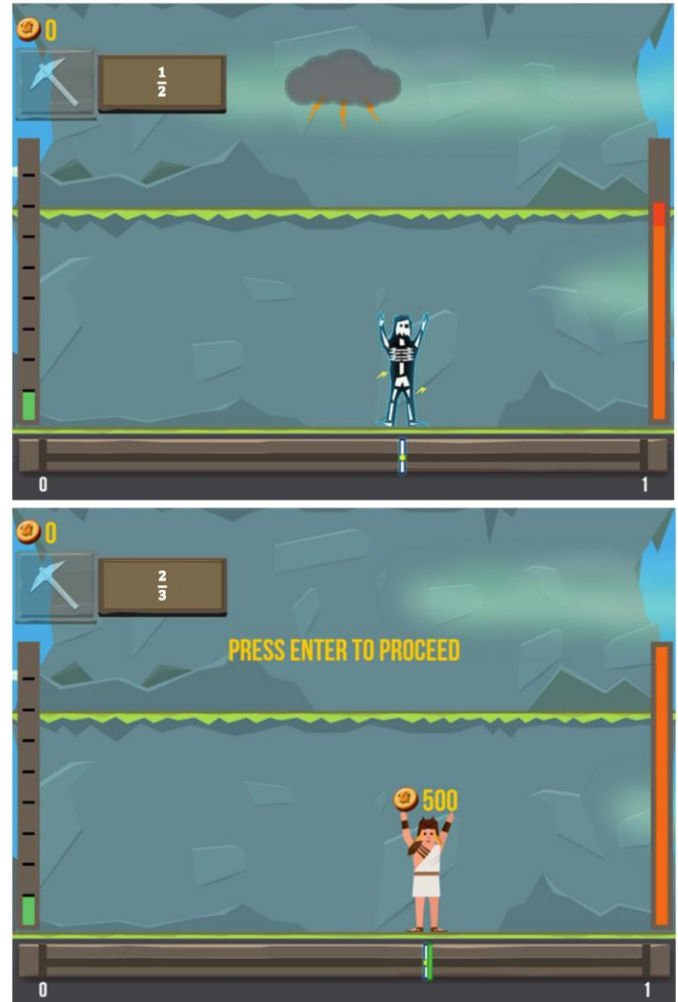


Fig. 2. Exemplary screenshots of the game condition; upper chart: negative feedback; lower chart: positive feedback.

The game version of the task was based on the same core number line estimation mechanic as the no-game version, but it was extended with visual game elements such as a narrative and corresponding scenery, an avatar, coins, progress bar, energy bar, level goals (stars), and gesture-based feedback. As it is argued that game elements disconnected from the actual task seem to distract more from task objectives or learning goals [7], we aimed at integrating the added game elements to the core task in a way that all the elements have a meaningful role. In terms of flow theory, the game elements were included to

facilitate goal formation, sense of control, and interpretation of feedback with respect to goals. With respect to the signaling principle of multimedia learning theory [83], the avatar was walking on the number line highlighting the most essential part of the game, the number line.

Specifically, in the game condition, participants controlled an avatar called Semideus, who tries to collect gold coins, which a goblin had stolen from Zeus. That is, the core mechanic of the game was to locate coins from the trails of the mountain with help of symbolic fraction numbers. The avatar was controlled on the number line in the same way as in the no-game condition (arrow keys of the keyboard). A target fraction, the location of hidden gold coins, was displayed in the upper left part of the game display (Fig. 2).

The task of the participants was to locate the position of this fraction on the number line to dig up the hidden coins. To confirm that Semideus digs the coins up at the estimated location, the participant had to press the space bar on the keyboard and Semideus dug into the ground or number line, respectively. For inaccurate estimates (i.e., estimates more than $\pm 5\%$ away from the correct location) the avatar was struck by lightning (see Fig. 2) and the player lost 10 units of virtual energy (displayed by an orange bar on the right side of the screen) on the first error in a trial, 5 units of energy for the second error on the same trial, and 2.5 units of energy for any further errors on that same trial/item. The same negative feedback mechanic (loss of energy) was used if players took too long to press the space bar (>10 seconds) or respond, respectively. This time limit was visualized by a cloud getting darker with passing time and a numerical countdown within the cloud. Similar to the no-game version, participants again had 10 seconds to indicate the location of the fraction in case the estimation took too long or was incorrect. For correct estimates, the avatar was rewarded with coins added to the reward counter placed at the upper left part of the screen and the correct location was shown by a green marker on the number line. Additionally, positive feedback was provided through gestures of Semideus (e.g., lifting hands up and cheering; see Fig. 2). After correct estimates, participants were asked to press the Enter key to proceed up to the next platform of a mountain, so that the next task/target fraction was presented. Participants could acquire 100 to 500 coins depending on the degree of estimation accuracy (i.e., over 98% = 500 coins; 97%–98% = 300 coins; 95%–96% = 100 coins). When all eight tasks of one level were completed, Semideus reached the top of the mountain where Zeus was waiting for him to bring back the stolen gold coins. After completing a level (reaching the top of the mountain) participants got additional feedback about their overall performance in the level: Stars and earned coins were shown (i.e., one star for completing the level and reaching the mountain top, one star for collecting more than 2000 coins, and one star when more than 80% of virtual energy was left). Although participants could run out of virtual energy in a level (100 units energy in the beginning of each level), they were still able to complete the level. However, at the mountain top participants did not earn the bonus points awarded according to remaining virtual energy.

To sum up, the game condition employed typical characteristics or building blocks, respectively [3], of a game such as narrative elements, appealing visual aesthetics, virtual

incentives in the form of points and stars earned according to the performance of users, as well as positive/negative feedback.

C. Eye Tracking

To objectively assess user behavior during task execution, eye fixation behavior was tracked using a Tobii 1750 Eye Tracker (Tobii Technology). This Eye Tracker is integrated into a 19" computer monitor and was connected to a conventional computer. Above and below the monitor there are infrared light-emitting diodes. The eye tracker collects binocular eye-tracking data at a rate of 50 Hz. Participants were seated with a maximum distance of 55 cm in front of the monitor. With the Tobii 1750 Eye Tracker (Tobii Technology), no fixation of the head was necessary enabling natural behavior of participants. Each eye-tracking measurement started with a short calibration phase. The number and duration of fixations (i.e., eye gaze remaining at the same point on the screen for longer than 100 ms, [84]) were recorded.

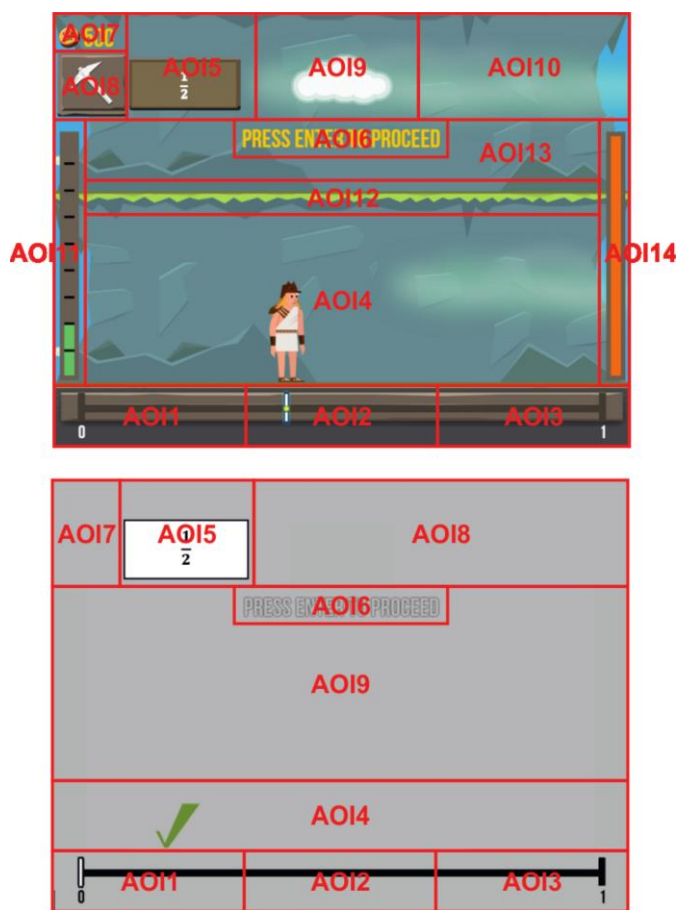


Fig. 3. AOI distribution of game (upper chart) and no-game condition.

For the analysis of the eye-tracking data, the display of the game and the no-game condition was subdivided into different areas of interest (AOI). The screen of the game condition was divided in 14 AOIs, the screen of the no-game condition in 9 AOIs (Fig. 3). The average number of fixations per AOI was analyzed. AOIs 1 to 3 covered the number line in both conditions. AOI1 corresponded to the left third of the number line, AOI2 to the middle third of the number line, and AOI3 to the right third. AOI5 comprised the target fraction, which

should be located on the number line (Fig. 3). AOI4 covered the area in which the Semideus avatar moved along the number line in the game condition and the red cross or green check mark were presented in the no-game condition. AOI6 covered the text “Press Enter To Proceed” after a task was solved correctly.

D. Questionnaires

Questionnaires were used to assess motivational aspects after each condition. Flow experience was measured with the Flow Short Scale (FKS; [85]), which measures components of flow with 10 items. Six of the items measure “fluency of performance” (e.g., “I have no difficulty concentrating”) and four items measure “absorption by activity” (e.g., “I do not notice time passing”) on a 7-point scale (1 = “not at all”–7 = “very much”). The mean of each component was used in the analyses.

User experience was assessed using the User Experience Questionnaire (UEQ; [86]) which has been widely used to assess interaction quality of certain design variations of products [87]. Accordingly, this questionnaire assesses conventional usability aspects (efficiency, transparency, controllability), user experience (originality, stimulation), as well as attractiveness. The rather general attractiveness subscale consists of 6 bipolar ratings, such as “unpleasant–pleasant” or “appealing–repelling.” Efficiency is measured by 4 items, such as “quick–slow” or “efficient–inefficient.” Perspicuity is assessed with 4 items, such as “complicated–simple” or “clear–confusing.” Dependability is assessed with 4 items such as “predictable–unpredictable.” The subscale Novelty is measured by 3 items such as “creative–dull” and, finally, the subscale stimulation is measured by 3 items, such as “boring–exciting.”

E. Procedure

Participants were tested individually. Half of the participants ($n = 21$) started with the game condition and the other half ($n = 21$) started with the no-game condition, which both included practice trials in the respective condition. After completing the first condition (6 levels á 8 tasks), participants completed the FKS and UEQ questionnaires to assess their subjective experience during the previously completed condition. Then, the second condition had to be performed before completing the FKS and UEQ questionnaires again. Overall, the experiment took about one hour.

F. Analyses

1) Behavioral performance level

To test performance differences in both conditions (game vs. no-game) we ran paired t -tests on error rate, mean accuracy, and mean duration per task. The error rate corresponded to the absolute number of errors committed by each participant averaged for each condition. The accuracies of participants’ estimation attempts were pooled to calculate the mean accuracy for each participant. Mean duration was defined by the average time needed to correctly solve a given item or fraction, respectively—from item onset to the last space bar press, i.e., until the estimate was correct. As these effects are crucial to address our primary research questions regarding the seductive detail effect of game elements, we validated null effects by a

Bayesian model selection approach, which investigates whether the null hypothesis or the alternative hypothesis is more supported by the data [88]. Accordingly, we calculated the posterior probability that the data favor the null hypothesis and the alternative hypothesis, respectively.

2) Attentional level

We first analyzed general differences in absolute fixations in both conditions using a paired t -test. We then specifically analyzed differences in the percentage of fixations on task relevant AOIs. Accordingly, we employed a 2 x 4 ANOVA with the within-subject factors condition (game vs. no-game) and task relevant AOIs (AOI1, AOI2, AOI3, and AOI5). Next, we used correlational analyses to investigate associations between the total number of fixations across all AOIs and task accuracy as well as errors.

For analyzing the duration of fixations, we followed a similar approach. We first ran a paired t -test on the mean duration of the fixations averaged over all AOIs per condition and correlational analyses between fixation duration and task accuracy and errors. We then compared mean duration of fixations between the game and no-game condition for task relevant AOIs. Accordingly, mean duration of fixations was analyzed using a 2 x 4 ANOVA with the within-subject factors condition (game vs. no-game) and task relevant AOIs (AOI1, AOI2, AOI3, and AOI5).

3) Motivational level

Due to technical problems ratings for the flow questionnaire (FKS) of 1 participant and ratings for the user experience questionnaire of 20 participants were not recorded in both conditions of the experiment. Questionnaires were analyzed using separate paired t -tests for each subscale of UEQ and FKS.

IV. RESULTS

A. Behavioral Results—Performance Level

Performance in number line estimation was comparable between both game conditions (Table I). Although participants made descriptively more errors in the no-game condition, this difference was statistically not significant. These null effects were further investigated using a Bayesian model selection approach. As regards the errors, the posterior probability for the null hypothesis was 0.63 providing weak evidence for the null hypothesis (i.e., no difference in error rate between game and no-game condition) according to Raftery [89]. As regards the mean accuracy, the posterior probability for the null hypothesis was 0.82 providing positive evidence for the null hypothesis (i.e., no difference in estimation accuracy between game and no-game condition). As regards the mean duration per task, the posterior probability for the null hypothesis was 0.78 providing positive evidence for the null hypothesis (i.e., no difference in mean duration per task between game and no-game condition). We need to note that only 2.42% (game version) and 3.92% (no-game version) of all provided estimates took participants longer than 10 seconds, indicating that the employed time limit per estimate did not particularly pressure participants.

TABLE I
PARTICIPANTS' PERFORMANCE IN GAME AND NO-GAME CONDITION

	Game condition	No-game condition	<i>p</i> -values of <i>t</i> -test	Effect size <i>d</i>
	<i>Mean (SE)</i>	<i>Mean (SE)</i>		
<i>Error rate</i>	24.51 (1.93)	27.31 (2.65)	0.11	0.26
<i>Mean Accuracy (%)</i>	94.70 (0.29)	94.47 (0.30)	0.40	0.13
<i>Mean duration per task (ms)</i>	8530.38 (279.38)	8190.46 (368.98)	0.29	0.17

B. Eye-Tracking Results—Attentional Level

1) Fixations

Overall, participants made significantly more fixations in the game condition (*mean* absolute number of fixations for all AOIs 730.48, *SE* = 69.75) than in the no-game condition (*mean* absolute number of fixations for all AOIs 581.52, *SE* = 61.96; $t(41) = 3.22$, $p < 0.01$, $d = 0.50$). Fig. 4 illustrates the absolute number of fixations per AOI, presented separately for both game conditions.

For the dependent variable percentage of fixations, an ANOVA with the within-subject factors game condition (game vs. no-game) and task relevant AOIs (AOI1, AOI2, AOI3, and AOI5) was calculated. The ANOVA revealed a significant main effect game condition [$F(1,41) = 146.51$, $p < 0.01$, $\eta^2 = 0.78$] and a significant main effect AOI [$F(3,63.34) = 17.30$, $p < 0.01$, $\eta^2 = 0.30$]. Moreover, the interaction of game condition and AOI was also significant [$F(3,41) = 3.96$, $p < 0.05$, $\eta^2 = 0.09$]. Post-hoc *t*-tests (Benjamini-Hochberg corrected for multiple comparisons to adjust the alpha level) revealed that in the no-game condition participants fixated the AOIs 1, 2, 3, and 5, which were the same in both conditions, more frequently than in the game condition (Fig. 5; game: *mean* = 16.07%, *SE* = 0.46; no-game: *mean* = 20.93%, *SE* = 0.44).

In the game condition, the middle and right part of the number line (AOI2, AOI3) were more frequently fixated than the left part (AOI1), and participants fixated more often the task (AOI5) than the number line (AOI1, AOI2, AOI3). In the no-game condition, the middle part of the number line (AOI2) was more frequently fixated than the left and right part (AOI1, AOI3) and the right part (AOI3) was more frequently fixated than the left part (AOI1). The task (AOI5) was more frequently fixated than the left part of the number line (AOI1), but the number of fixations on the task (AOI5) in the no-game condition was comparable with the number of fixations on the middle and right part of the number line (AOI2, AOI3).

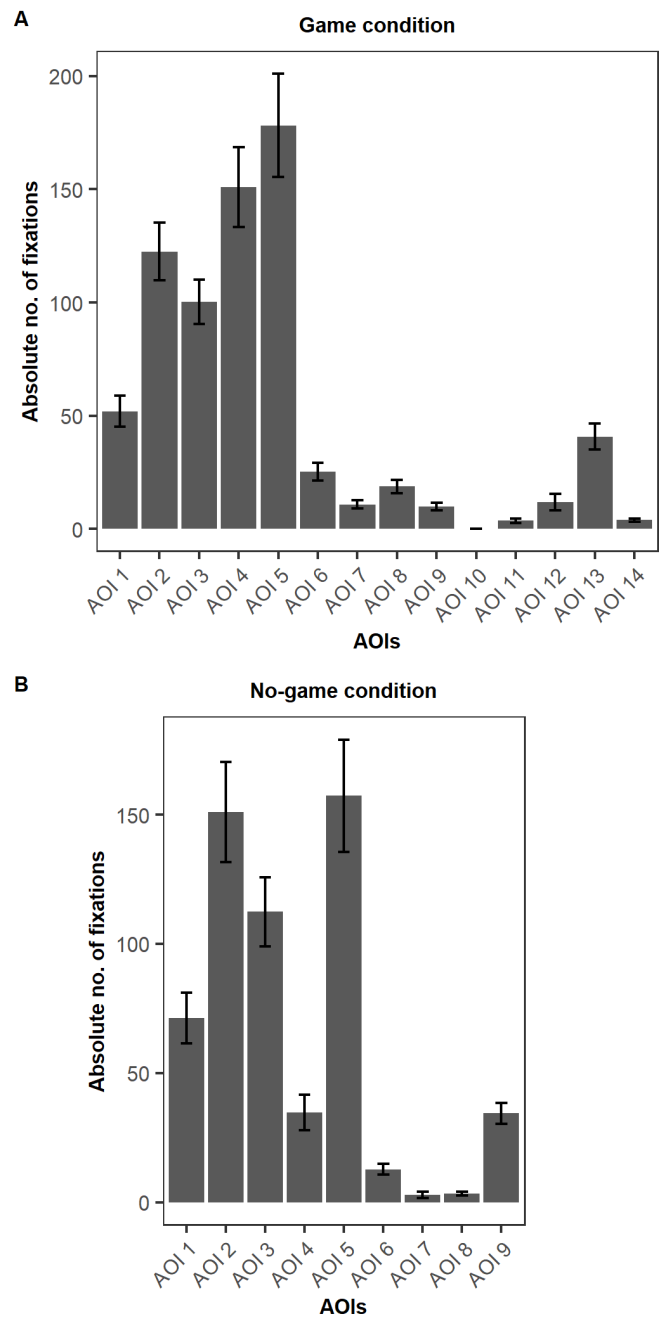


Fig. 4. Number of fixations across all AOIs in game and no-game condition. Error bars depict 1 standard error of the mean.

The patterns of fixations in both, the game and no-game condition, were rather similar and a result of the fraction selection in the current experiment (see density plot in Fig. 6), i.e., correct positions on the number line were more frequently found in the middle (AOI2) and right side (AOI3) of the number line as compared to the left side (AOI1). This could also be identified in a heat map of eye fixations in Fig. 7.

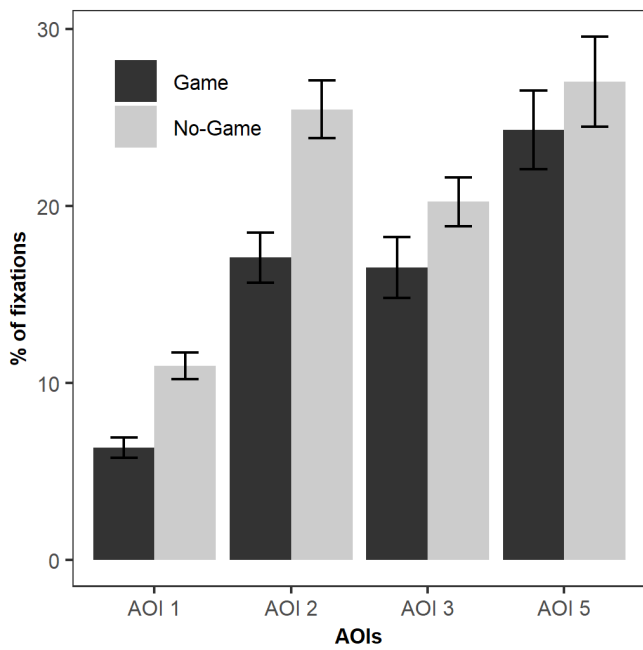


Fig. 5. Percentage of fixations across AOI 1, 2, 3, and 5 for game and no-game condition. Error bars depict 1 standard error of the mean.

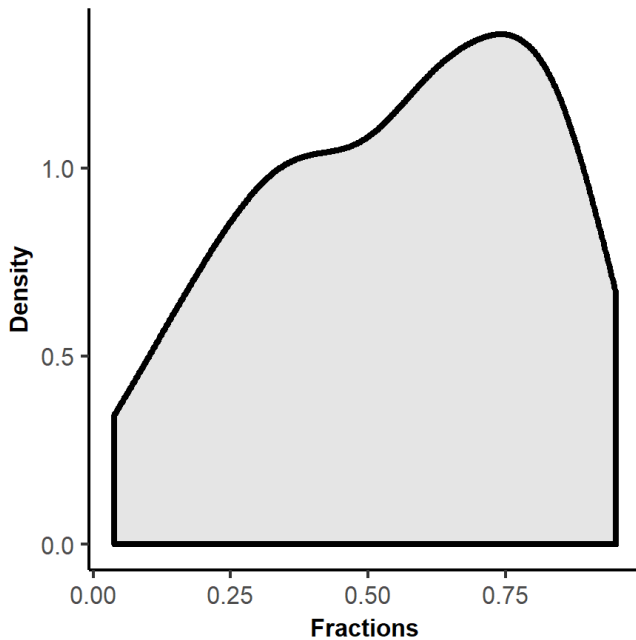


Fig. 6. Percentage of fixations across AOI 1, 2, 3, and 5 for game and no-game condition. Error bars depict 1 standard error of the mean.

Additionally, we found a significant relationship between the total number of fixations across all AOIs (i.e., the whole screen) and performance in the number line estimation task in the no-game condition. The more fixations a participant made in the no-game condition, the higher her/his error rate ($r = 0.50, p < 0.01$) and the worse her/his estimation accuracy ($r = -0.49, p < 0.01$). In the game condition, the correlation results were not significant (absolute number of fixations * errors: $r = 0.20, p = 0.21$; absolute number of fixations * accuracy: $r = -0.22, p = 0.18$).

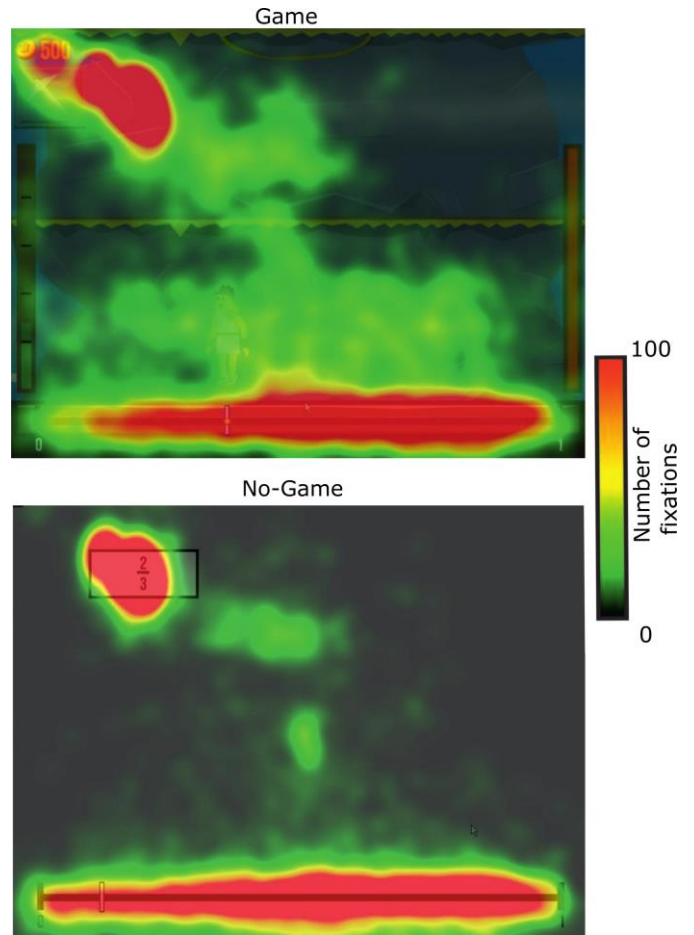


Fig. 7. Average gaze patterns (heat maps of number of fixations) illustrated for the game and the no-game condition.

2) Fixation durations

The mean duration of the fixations averaged over all AOIs per condition was longer in the no-game condition ($mean = 209.30$ ms, $SE = 12.28$) than in the game condition ($mean = 199.14$ ms, $SE = 10.63$; $t(41) = 3.45, p < 0.01, d = 0.53$). However, fixation duration did not correlate with performance (error rate and accuracy) in number line estimation, neither in the game nor in the no-game condition (all $p > 0.71$).

When comparing the mean duration of fixations between the game and no-game condition only for AOI 1, 2, 3, and 5, which were directly comparable between conditions, an ANOVA model with the within-subject factors game condition (game vs. no-game) and AOI (AOI1, AOI2, AOI3, and AOI5) only revealed a significant main effect of AOI [$F(3,54.95) = 13.90, p < 0.01, \eta^2 = 0.26$]. Post-hoc testing indicated that participants fixated the middle and right part of the number line longer than the left part and the task (Fig. 8). All other effects were not significant. Hence, the mean duration of fixations did not differ between the game and no-game condition when analyzing AOI 1, 2, 3, and 5.

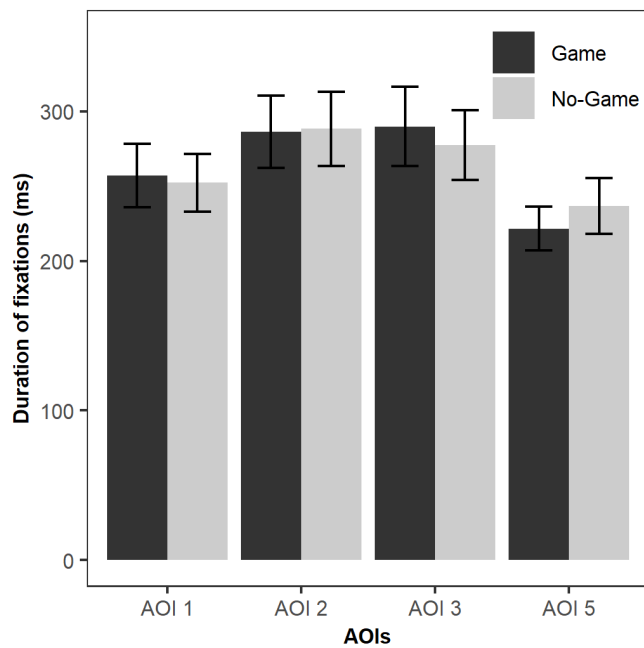


Fig. 8. Duration of fixations in ms across AOI 1, 2, 3, and 5 for game and no-game condition. Error bars depict 1 standard error of the mean.

C. Questionnaire Results—Motivational Level

To examine the hypothesized differences in user experience in both conditions we conducted paired t -tests for each subscale (attractiveness, perspicuity, efficiency, novelty, stimulation, and dependability) of the UEQ with the within-subject factors condition (game vs. no-game). The ratings on each scale were used as dependent variable. Paired t -tests (Benjamini-Hochberg corrected for multiple comparisons) revealed that participants rated the game condition to be more attractive [game: $mean = 0.89$, $SE = 0.25$; no-game: $mean = -0.20$, $SE = 0.34$; $t(21) = 2.86$, $p < 0.05$, $d = 0.61$], more novel [game: $mean = 0.24$, $SE = 0.29$; no-game: $mean = -1.41$, $SE = 0.29$; $t(21) = 4.61$, $p < 0.001$, $d = 0.98$], and more stimulating [game: $mean = 0.58$, $SE = 0.30$; no-game: $mean = -0.33$, $SE = 0.32$; $t(21) = 3.18$, $p < 0.05$, $d = 0.68$] as compared to the no-game condition. No significant differences were found in subscales unrelated to characteristics of the game condition in efficiency [game: $mean = 1.15$, $SE = 0.17$; no-game: $mean = 1.31$, $SE = 0.19$; $t(21) = -1.06$, $p = 0.38$, $d = 0.22$], perspicuity [game: $mean = 2.05$, $SE = 0.17$; no-game: $mean = 2.08$, $SE = 0.17$; $t(21) = -0.24$, $p = 0.81$, $d = 0.05$], and dependability [game: $mean = 0.95$, $SE = 0.18$; no-game: $mean = 1.20$, $SE = 0.31$; $t(21) = -1.02$, $p = 0.38$, $d = 0.22$].

Finally, we examined differences in the perceived flow of participants in both conditions. Therefore, we performed paired t -tests for each FKS subscale (absorption and fluency) with the within-subject factor condition (game vs. no-game). We used participants ratings as dependent variable. Paired t -tests (Benjamini-Hochberg corrected for multiple comparisons) revealed that participants felt the experience in the game condition to be more fluent or automatic [game: $mean = 4.68$, $SE = 0.22$; no-game: $mean = 4.18$, $SE = 0.23$; $t(40) = 2.44$, $p < 0.05$, $d = 0.38$], respectively. Participants did not differ in their experienced absorption in the game and no-game condition [game: $mean = 5.17$, $SE = 0.16$; no-game: $mean = 5.09$, $SE = 0.19$; $t(40) = 0.51$, $p = 0.61$, $d = 0.08$].

V. DISCUSSION

In the current study we investigated the effects of game elements on user behavior in a numerical learning task on three different levels: (i) behavioral performance (performance in the number line estimation task), (ii) attention distribution (eye-fixation behavior), and (iii) motivational aspects (flow and user experience). Therefore, participants performed a number line estimation task with and without game elements. Results indicated differences in selective attention as reflected by eye fixation behavior but no differences in performance between the game and no-game condition. Additionally, flow as well as user experience was higher in the task condition with game elements. In the following, results on the respective levels will be discussed in more detail.

A. Performance Level

Importantly, our results indicated that the game elements in the game condition did not disturb participants' estimation performance as behavioral performance was comparable between both conditions. Neither error rates nor estimation accuracy or time to solve the tasks were affected by the inclusion of game elements. These results were further substantiated by Bayesian analyses. Therefore, our data do not support the assumption that adding game elements, which were not mandatory for solving the task, decreases performance as predicted by the seductive details effect [9]. Although Rey [9] identified overall small to medium-sized negative effects on performance, he acknowledged, amongst others, learning domain as well as the kind of seductive details as important mediating factors. Research on game elements in this regard is particularly sparse emphasizing the importance of the current results. In line with Rey [9] and Habgood & Ainsworth [7], we argue that seductive details or game elements, respectively, that interrupt the coherence of the instructional material, rather than being well (intrinsically) integrated with learning mechanics, are particularly prone to impede processing. As described above, in the present game condition, we took great care to integrate all game elements with the core number line estimation task in a coherent and meaningful way, which may account for the lack of negative effects. However, the eye-tracking results revealed also that participants allocated only very limited amount of attention for the level progress bar and the virtual energy bar in the game condition (Fig. 7). This, together with the relatively low challenge level and the lack of consequences when reaching zero virtual energy might have undermined the meaning of these elements in the game version.

Rey [9] further observed that the presence of time limits led to poorer performance in the seductive detail condition. He argued that users in this condition might need longer for processing larger amounts of instructional material. However, the current study did not find performance differences even though a time limit of 10 seconds per estimate was employed. However, it needs to be mentioned that in both conditions only a very small number of estimates actually exceeded the employed time limit, indicating that participants might not have felt particularly pressured. Moreover, comparable performance can be also explained with the relatively low challenge level of the game. It is possible that even though the game elements induced more extraneous cognitive load, the overall cognitive

load remained within participants' cognitive capacity limits. Thus, processing of game elements did not disturb performance. In line with this, we also did not find differences in time needed to solve tasks between the game and no-game condition. The current results indicate, that participants were able to efficiently differentiate between relevant and the rather "irrelevant" game elements in the game condition. This is consistent with recent research on effects of irrelevant pictorial information indicating that students learned to ignore irrelevant information as they gained experience with the task and thus no longer experienced negative effects [90], [91]. As compared to similar studies in this line of research, our study employed a task with rather simple mechanics and limited relevant information—providing only a target number and the number line. Therefore, users might have got proficient with the task quite quickly and learned to differentiate between the fundamental task and game elements in the game condition.

B. Attentional Level

Although there were no performance differences between game and no-game condition, we observed differential eye movement patterns. Overall, participants made more fixations in the game condition than in the no-game condition, which is in line with other studies investigating user interface attractiveness [92] or seductive details [93], [94]. However, task relevant AOIs (AOI1, AOI2, AOI3 reflecting the number line, and AOI5 representing the to-be-solved fraction) were fixated less often in the game condition as compared to the non-game condition. This might indicate that participants' attention was distracted away from task relevant features more often in the game condition. This is not particularly surprising as more task irrelevant features were present in this condition and is in line with previous studies investigating seductive details [93], [94]. Importantly, however, this did not affect performance indicating users' ability to pay attention to both relevant and rather irrelevant task elements. In the game condition, participants also frequently looked at the Semideus avatar (AOI 4), not present in the no-game condition, as it might have served as visual extension of the cursor indicating the current position on the number line. Nevertheless, the avatar Semideus reflected partly the same information as the cursor visible upon the number line at the bottom of the screen (AOI 1–3). Accordingly, looking at the number line at the bottom of the screen (AOI 1–3) and the target number (AOI 5) was sufficient to solve the estimation tasks adequately in the no-game as well as in the game condition. Other game elements, such as the counter of reward coins (AOI 7), the bars showing the virtual energy and level progression (AOI 11 & 14), and the cloud (AOI 9) did not seem to disturb players, but provided additional information about the state of the game and provided performance feedback. The respective areas of the game screen were even largely ignored as revealed by the eye-tracking data (see Fig. 4).

On the one hand, these results support the notion that users in the game condition might have learned not to pay particular attention on task irrelevant game elements as indicated by low number of fixations on, for instance, the counter of reward coins. However, on the other hand, users in the game condition paid particular attention to the avatar (AOI 4). This might be due to the fact, that (i) the avatar served as a visual extension

indicating the current position on the number line and (ii) the avatar provided gesture-based feedback in terms of jubilant gestures (correct response) or getting hit by lightning (incorrect response) as well as rewarded coins to the user. In contrast, the respective AOI in the no-game condition (AOI 4), which also provided feedback but in a less engaging way (i.e., showing a green check mark for correct responses and a red cross for incorrect responses) received less attention from the learners. Direct comparison of AOI 4 between the game and no-game condition was difficult. Even though, in both conditions AOI 4 provided positive and negative feedback, the size of the respective AOI in the game condition had to be considerably larger due to the avatar and his animations.

Therefore, we only described the results on a descriptive level, which indicated increased attention was being paid to this area in the game condition (*mean* number of fixations = 150.90; *SE* = 17.52) as compared to the non-game condition (*mean* number of fixations = 34.67; *SE* = 6.85). However, this result needs to be treated with caution as the comparability between these two areas is not straightforward due to their difference in size. Nevertheless, AOI 4 including the avatar was one of the most frequently fixated screen areas in the game condition that is in line with previous studies indicating that players tend to focus on their avatar [48]. Consequently, an avatar seems to be an effective channel to provide crucial information, such as performance feedback to the player.

Previous research already showed a relationship between the number of fixations and performance in different tasks (e.g., expert vs. novice problem solving: [59], [60]; puzzle-like task: [54]; reading: [55]; watch a video with geographical content: [95]). In line with this, we observed that performance was lower for users who made more fixations while solving the task. However, we only found this correlation to be significant in the no-game condition, but not in the game condition. To our knowledge, only one other study investigated eye-tracking metrics similar to the current study in a number line estimation task. Schneider *et al.* [96] found eye fixations to be scattered more in worse performers or younger children, respectively. The authors assumed that younger children do not yet have developed appropriate solution strategies and thus are less direct in their approach requiring more orientation and reorientation along the number line in order to locate a specific target number [96]. This general pattern is consistent with our finding, because lower estimation performance was associated with more fixations in the no-game condition. This association was not found in the game condition, which might indicate that inclusion of game elements might undermine the predictive power of fixation behavior [56], [97].

Mean fixation duration over all AOIs was longer in the no-game condition as compared to the game condition. However, the mean fixation duration did not correlate with number line estimation performance (error rate and accuracy), neither in the game nor in the no-game condition. In eye-tracking research, longer mean fixation duration is usually associated with higher processing difficulty and cognitive load, but also deeper processing [46], [98], [99]. The duration of fixations in AOI 1, 2, 3, and 5 (number line at the bottom of the screen and task/fractions), which were the most central and task relevant AOIs for solving the estimation task in both conditions, did not differ between the game and no-game condition. Moreover, in

both conditions, the middle and right part of the number line was fixated longer than the left part and task area, which corresponded to the fraction distribution. Therefore, even though, participants in the game condition made more fixations overall processing difficulty, as well as the level of processing depth of task relevant areas seemed to be comparable in both groups. This further emphasized the notion that users were successful in differentiating between relevant and irrelevant task elements in the game condition.

C. Motivational Level

One of the reasons for including seductive details into a task is to increase interest of the user, which may enhance motivation and flow [25], [26]. We indeed found increased flow experience in the game condition as compared to the no-game condition. However, this was only observed for the subscale “fluency,” but not for “absorption.” We assume that the game elements, especially narrative, level goals, virtual incentives, progress bar, and virtual energy bar contributed to clear goals and sense of control dimensions of flow included in the fluency subscale. Hence, participants felt the task in the game condition to run more smoothly and more fluently, felt more sense of control, and thus had less problems concentrating. We can only speculate as to why absorption did not differ between the two conditions. The used game condition only employed a few game elements, which might not have been enough to affect the immersive dimension of flow in the absorption subscale. Even though flow is often associated with increased performance and learning gains [74]–[76], we did not find performance differences between game and no-game condition. In both conditions performance was rather high (accuracy at about 95%) indicating that the task, even though we used single- and double-digit fractions, was rather easy for the adult participants. Hence, the results might be explained with Yerkes-Dodson Law [100], which postulates an inverted U-shape between arousal and performance. Game elements might have influenced the users arousal level, which helped them to better concentrate on the task and felt the experience to be more fluent. However, since the performance of the task was already rather high no performance difference between the conditions was observable. In line with the results for flow we also identified increased stimulation ratings in the game condition as compared to the no-game condition. Moreover, participants rated the game condition to be more attractive and novel. As such, we might expect, that when users were given the chance to choose between the two task environments, the game-based version of the number line estimation would be favored [78]. Importantly, however, we did not find any difference on the UEQ subscales perspicuity, efficiency, and dependability. This demonstrated that the game elements did not affect the task in a more pragmatic way, indicating that we established two conditions that are comparable, which often renders problems in media comparison studies [101].

D. Limitations and Future Perspective

The main limitation of the current study was the limited interaction time with the actual task of about 15 minutes with each version. Consequently, generalizing the effects of game elements on performance, attention, and motivation on longer lasting interactions is not possible. Moreover, missing transfer

tests and longer retention intervals or follow up-tests, respectively, limit the conclusions that can be drawn from the non-existing performance differences between game and no-game condition. This is particularly important as previous research showed that positive effects of more rewarding instructional experiences might be more pronounced with longer retention intervals [102], [103] due to memory consolidation processes. Moreover, for the university student sample the task might have been rather easy. One might assume that a university student sample might have had higher cognitive resources and wouldn't get as easily overloaded by game elements. Previous research demonstrated that students with lower cognitive capacity were significantly more distracted by seductive details (illustrations). Their attention was drawn to seductive details more often and for longer time intervals [17], [28]. Moreover, potential pre-existing knowledge differences were not evaluated in the current study. Hence, future studies investigating effects of game elements should recruit more heterogeneous samples and employ comprehensive pretests. Lastly, further research is needed to distinguish between a coherent and incoherent augmentation of tasks with game elements. That is, game elements disconnected from the actual learning content are hypothesized to be more distracting and thus hampering learning [7]. In the current study, we paid particular attention to a coherent and meaningful integration of game elements. Accordingly, task mechanic, narrative, feedback, and visual appearance were carefully matched and embodied the learning material within the structure of the game world and the learners' interactions with it. Consequently, in order to be able to generalize current results, the way of integrating game elements needs to be studied systematically with different type of games, in different context, and for various learning domains [9]. Despite the aforementioned limitations, the current study makes a strong case against the negative effects of game elements on performance, at least during rather short interactions with instructional material.

E. Implications and Conclusions

The present findings showed that game elements did not impair performance in number line estimation—an often-used task for assessment and training of number magnitude understanding. Consequently, from a theoretical perspective, we did not observe the seductive details effect elicited by augmenting a task with game elements. We did, however, find different eye-fixation behavior for the game and the no-game condition. Eye-tracking data revealed that participants' selective attention was captured by game elements in the game condition. However, no significant performance difference was observable between the game and the no-game condition. The observed qualitative differences in fixation behavior might also originated from increased user and flow experience. Participants indicated that they felt the game task to be more attractive, stimulating, novel, and had less problems concentrating as the task felt to be smoother and more fluent. On the practical side, we demonstrated that eye tracking is a valuable tool for exploring users attention towards game elements, which would not have been possible or as detailed with conventional post-hoc questionnaires. Accordingly, developers of educational technologies and games can use the information gathered with

eye tracking to enhance the design of learning materials. Finally, we would argue that game elements should be included in tasks rather than excluded, but only when game elements and learning mechanics are appropriately integrated and the game does not overstrain players cognitive resources.

REFERENCES

- [1] E. A. Boyle *et al.*, “An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games,” *Comput. Educ.*, vol. 94, pp. 178–192, Mar. 2016, doi: 10.1016/j.compedu.2015.11.003.
- [2] T. Hainey, T. M. Connolly, E. A. Boyle, A. Wilson, and A. Razak, “A systematic literature review of games-based learning empirical evidence in primary education,” *Comput. Educ.*, vol. 102, pp. 202–223, Nov. 2016, doi: 10.1016/j.compedu.2016.09.001.
- [3] J. L. Plass, B. D. Homer, and C. K. Kinzer, “Foundations of game-based learning,” *Educ. Psychol.*, vol. 50, no. 4, pp. 258–283, 2015, doi: 10.1080/00461520.2015.1122533.
- [4] K. Salen and E. Zimmerman, *Rules of play: Game design fundamentals*. MIT Press, 2004.
- [5] S. Pawar, F. Tam, and J. L. Plass, “Emerging design factors in game-based learning: Emotional design, musical score, and game mechanics design,” in *Handbook of Game-Based Learning*, Cambridge, MA, USA: MIT Press, 2020, pp. 347–365.
- [6] K. Kiili and P. Tuomi, “Teaching educational game design: Expanding the game design mindset with instructional aspects,” in *Games and Learning Alliance Conf.*, 2019, pp. 103–113, doi: 10.1007/978-3-030-34350-7_11.
- [7] M. P. J. Habgood and S. E. Ainsworth, “Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games,” *J. Learn. Sci.*, vol. 20, no. 2, pp. 169–206, Apr. 2011, doi: 10.1080/10508406.2010.508029.
- [8] R. E. Mayer, “Cognitive foundations of game-based learning,” in *Handbook of Game-Based Learning*, Cambridge, MA, USA: MIT Press, 2020, pp. 83–110.
- [9] G. D. Rey, “A review of research and a meta-analysis of the seductive detail effect,” *Educ. Res. Rev.*, vol. 7, no. 3, pp. 216–237, Dec. 2012, doi: 10.1016/j.edurev.2012.05.003.
- [10] M. Ninaus *et al.*, “Increased emotional engagement in game-based learning – A machine learning approach on facial emotion detection data,” *Comput. Educ.*, vol. 142, Jul. 2019, doi: 10.1016/j.compedu.2019.103641.
- [11] D. B. Clark, E. E. Tanner-Smith, and S. S. Killingsworth, “Digital games, design, and learning: A systematic review and meta-analysis,” *Rev. Educ. Res.*, vol. 86, no. 1, pp. 79–122, 2016, doi: 10.3102/0034654315582065.
- [12] P. Wouters, C. van Nimwegen, H. van Oostendorp, and E. D. van Der Spek, “A meta-analysis of the cognitive and motivational effects of serious games,” *J. Educ. Psychol.*, vol. 105, no. 2, pp. 249–265, 2013, doi: 10.1037/a0031311.
- [13] F. Bellotti *et al.*, “Designing serious games for education: From pedagogical principles to game mechanisms,” in *Proc. 5th European Conf. Game-Based Learning*, Athens, Greece, 2011, pp. 26–34.
- [14] S. W. McQuiggan, J. P. Rowe, S. Lee, and J. C. Lester, “Story-based learning: The impact of narrative on learning experiences and outcomes,” in *Intelligent Tutoring Systems*, 2008, pp. 530–539, doi: 10.1007/978-3-540-69132-7_56.
- [15] R. E. Mayer, “Cognitive theory of multimedia learning,” in *The Cambridge Handbook of Multimedia Learning*, R. E. Mayer, Ed. Cambridge, UK: Cambridge University Press, 2005, pp. 31–48.
- [16] R. Moreno and R. E. Mayer, “A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages,” *J. Educ. Psychol.*, vol. 92, no. 1, pp. 117–125, 2000, doi: 10.1037/0022-0663.92.1.117.
- [17] C. A. Sanchez and J. Wiley, “An examination of the seductive details effect in terms of working memory capacity,” *Mem. Cogn.*, vol. 34, no. 2, pp. 344–355, Mar. 2006, doi: 10.3758/BF03193412.
- [18] S. Lehman, G. Schraw, M. T. McCrudden, and K. Hartley, “Processing and recall of seductive details in scientific text,” *Contemp. Educ. Psychol.*, vol. 32, no. 4, pp. 569–587, Oct. 2007, doi: 10.1016/j.cedpsych.2006.07.002.
- [19] S. F. Harp and R. E. Mayer, “How seductive details do their damage: A theory of cognitive interest in science learning,” *J. Educ. Psychol.*, vol. 90, no. 3, pp. 414–434, 1998, doi: 10.1037/0022-0663.90.3.414.
- [20] R. E. Mayer and R. Moreno, “Nine ways to reduce cognitive load in multimedia learning,” *J. Educ. Psychol.*, vol. 38, no. 1, pp. 43–52, 2003.
- [21] S. F. Harp and R. E. Mayer, “The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest,” *J. Educ. Psychol.*, vol. 89, no. 1, pp. 92–102, 1997, doi: 10.1037/0022-0663.89.1.92.
- [22] R. E. Mayer, E. Griffith, I. T. N. Jurkowitz, and D. Rothman, “Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning,” *J. Exp. Psychol. Appl.*, vol. 14, no. 4, pp. 329–339, 2008, doi: 10.1037/a0013835.
- [23] D. Ozdemir and P. Doolittle, “Revisiting the seductive details effect in multimedia learning: Context-dependency of seductive details,” *J. Educ. Multimed. Hypermedia*, vol. 24, no. 2, pp. 101–119, 2015.
- [24] B. Park, R. Moreno, T. Seufert, and R. Brünken, “Does cognitive load moderate the seductive details effect? A multimedia study,” *Comput. Hum. Behav.*, vol. 27, no. 1, pp. 5–10, Jan. 2011, doi: 10.1016/j.chb.2010.05.006.
- [25] R. E. Mayer, *Multimedia Learning*. Cambridge, MA, USA: Cambridge Univ. Press, 2009.
- [26] K. A. Renninger and S. Hidi, *The power of interest for motivation and engagement*. New York, NY, USA: Routledge, 2016.
- [27] W. Schnotz, S. Fries, and H. Horz, “Motivational aspects of cognitive load theory,” in *Contemporary Motivation Research: From Global to Local Perspectives*, M. Wosnitza, S. A. Karabenick, A. Efkliides, and P. Nenniger, Eds., Göttingen, Germany: Hogrefe & Huber Publishers, 2009, pp. 69–96.
- [28] B. Park, L. Knörzer, J. L. Plass, and R. Brünken, “Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms,” *Comput. Educ.*, vol. 86, pp. 30–42, 2015, doi: 10.1016/j.compedu.2015.02.016.
- [29] J. L. Plass, S. Heidig, E. O. Hayward, B. D. Homer, and E. Um, “Emotional design in multimedia learning: Effects of shape and color on affect and learning,” *Learn. Instr.*, vol. 29, pp. 128–140, Feb. 2014, doi: 10.1016/j.learninstruc.2013.02.006.
- [30] E. Um, J. L. Plass, E. O. Hayward, and B. D. Homer, “Emotional design in multimedia learning,” *J. Educ. Psychol.*, vol. 104, no. 2, pp. 485–498, 2012, doi: 10.1037/a0026609.
- [31] R. Pekrun, “A social-cognitive, control-value theory of achievement emotions,” in *Motivational Psychology of Human Develop.: Developing Motivation and Motivating Develop.*, J. Heckhausen, Ed. New York: Elsevier Science Ltd., 2000, pp. 143–163.
- [32] R. Moreno and R. E. Mayer, “Interactive multimodal learning environments,” *Educ. Psychol. Rev.*, vol. 19, no. 3, pp. 309–326, 2007, doi: 10.1007/s10648-007-9047-2.
- [33] J. L. Plass and U. Kaplan, “Emotional design in digital media for learning,” in *Emotions, Technology, Design, and Learning*, S. Y. Tettegah and M. Gartmeier, Eds., Cambridge, MA, USA: Academic Press, 2016, ch. 7, pp. 131–161.
- [34] L. Knörzer, R. Brünken, and B. Park, “Emotions and multimedia learning: the moderating role of learner characteristics,” *J. Comput. Assist. Learn.*, vol. 32, no. 6, pp. 618–631, Dec. 2016, doi: 10.1111/jcal.12158.
- [35] K. M. Lee, N. Park, and S. A. Jin, “Narrative and interactivity in computer games,” in *Playing Video Games: Motives, Responses, and Consequences*, P. Vorderer and J. Bryant, Eds. Mahwah, New Jersey, New Jersey: Lawrence Erlbaum Associates, 2006, pp. 259–274.
- [36] T. Marsh, “Activity-based scenario design, development and assessment in serious games,” in *Gaming and Cognition: Theories and Practice from the Learning Sciences*, R. Van Eck, Ed. Hershey, PA, USA: IGI Global, 2010, pp. 213–226, 2010.
- [37] M. Ninaus *et al.*, “Game elements improve performance in a working memory training task,” *Int. J. Serious Games*, vol. 2, no. 1, pp. 3–16, Feb. 2015, doi: 10.17083/ijsg.v2i1.60.
- [38] P. J. M. Prins, S. Dovis, A. Ponsioen, E. ten Brink, and S. van der Oord, “Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD?,” *Cyberpsychol. Behav. Soc. Netw.*, vol. 14, no. 3, pp. 115–22, Mar. 2011, doi: 10.1089/cyber.2009.0206.
- [39] T. Brosch, D. Sander, G. Pourtois, and K. R. Scherer, “Beyond fear,” *Psychol. Sci.*, vol. 19, no. 4, pp. 362–370, Apr. 2008, doi: 10.1111/j.1467-9280.2008.02094.x.

- [40] M. J. Fenske and J. D. Eastwood, "Modulation of focused attention by faces expressing emotion: Evidence from flanker tasks," *Emotion*, vol. 3, no. 4, pp. 327–343, 2003, doi: 10.1037/1528-3542.3.4.327.
- [41] J. Vogt, J. De Houwer, E. H. W. Koster, S. Van Damme, and G. Crombez, "Allocation of spatial attention to emotional stimuli depends upon arousal and not valence," *Emotion*, vol. 8, no. 6, pp. 880–885, 2008, doi: 10.1037/a0013981.
- [42] G. D. Rey, "Seductive details and attention distraction—An eye tracker experiment," *Comput. Hum. Behav.*, vol. 32, pp. 133–144, 2014, doi: 10.1016/j.chb.2013.11.017.
- [43] M. Ninaus *et al.*, "Neurophysiological methods for monitoring brain activity in serious games and virtual environments: A review," *Int. J. Technol. Enhanc. Learn.*, vol. 6, no. 1, pp. 78–103, 2014, doi: 10.1504/IJTEL.2014.060022.
- [44] S. Nebel and M. Ninaus, "New perspectives on game-based assessment with process data and physiological signals," in *Game-Based Assessment Revisited. Advances in Game-Based Learning*, D. Ifenthaler and Y. Kim, Eds., Cham, Switzerland: Springer, 2019, pp. 141–161.
- [45] B. B. de Koning, H. K. Tabbers, R. M. J. P. Rikers, and F. Paas, "Attention guidance in learning from a complex animation: Seeing is understanding?," *Learn. Instr.*, vol. 20, no. 2, pp. 111–122, Apr. 2010, doi: 10.1016/j.learninstruc.2009.02.010.
- [46] E. Alemdag and K. Cagiltay, "A systematic review of eye tracking research on multimedia learning," *Comput. Educ.*, vol. 125, pp. 413–428, Oct. 2018, doi: 10.1016/j.compedu.2018.06.023.
- [47] M.-J. Tsai, L.-J. Huang, H.-T. Hou, C.-Y. Hsu, and G.-L. Chiou, "Visual behavior, flow and achievement in game-based learning," *Comput. Educ.*, vol. 98, pp. 115–129, Jul. 2016, doi: 10.1016/j.compedu.2016.03.011.
- [48] K. Kiili, H. Ketamo, and M. D. Kickmeier-rust, "Eye tracking in game-based learning research and game design," *Int. J. Serious Games*, vol. 1, no. 2, 2014.
- [49] N. H. Mat Zain, F. H. Abdul Razak, A. Jaafar, and M. F. Zulkipli, "Eye tracking in educational games environment: Evaluating user interface design through eye tracking patterns," in *Visual Informatics: Sustaining Research and Innovation*, 2011, pp. 64–73, doi: 10.1007/978-3-642-25200-6_7.
- [50] S. Deng, J. A. Kirkby, J. Chang, and J. J. Zhang, "Multimodality with eye tracking and haptics: A new horizon for serious games?," *Int. J. Serious Games*, vol. 1, no. 4, Oct. 2014, doi: 10.17083/ijsg.v1i4.24.
- [51] M. S. El-Nasr and S. Yan, "Visual attention in 3D video games," in *Proc. 2006 ACM SIGCHI Int. Conf. Advances in Computer Entertainment Technology (ACE '06)*, 2006, pp. 22–29, doi: 10.1145/1178823.1178849.
- [52] M. D. Kickmeier-Rust, E. Hillemann, and D. Albert, "Tracking the UFO's paths: Using eye-tracking for the evaluation of serious games," in *Virtual and Mixed Reality: New Trends*, 2011, pp. 315–324, doi: 10.1007/978-3-642-22021-0_35.
- [53] S. Alkan and K. Cagiltay, "Studying computer game learning experience through eye tracking," *Br. J. Educ. Technol.*, vol. 38, no. 3, pp. 538–542, May 2007, doi: 10.1111/j.1467-8535.2007.00721.x.
- [54] M. Frutos-Pascual and B. Garcia-Zapirain, "Assessing visual attention using eye tracking sensors in intelligent cognitive therapies based on serious games," *Sensor*, vol. 15, no. 5, pp. 11092–11117, 2015, doi: 10.3390/s150511092.
- [55] M. A. Just and P. A. Carpenter, "A theory of reading: From eye fixations to comprehension," *Psychol. Rev.*, vol. 87, no. 4, pp. 329–354, 1980, doi: 10.1037/0033-295X.87.4.329.
- [56] V. Sundstedt, E. Stavrakis, M. Wimmer, and E. Reinhard, "A psychophysical study of fixation behavior in a computer game," in *Proc. 5th Symp. Applied Perception in Graphics & Visualization (APGV '08)*, 2008, pp. 43–50, doi: 10.1145/1394281.1394288.
- [57] J. Hyönä, "The use of eye movements in the study of multimedia learning," *Learn. Instr.*, vol. 20, no. 2, pp. 172–176, Apr. 2010, doi: 10.1016/j.learninstruc.2009.02.013.
- [58] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychol. Bull.*, vol. 124, no. 3, pp. 372–422, 1998, doi: 10.1037/0033-2909.124.3.372.
- [59] R. H. Tai, J. F. Loehr, and F. J. Brigham, "An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments," *Int. J. Res. Method Educ.*, vol. 29, no. 2, pp. 185–208, 2006, doi: 10.1080/17437270600891614.
- [60] M. T. H. Chi, R. Glaser, and F. J. Rees, "Expertise in problem solving," Pittsburgh Univ PA Learning Research and Develop. Center, Pittsburgh, KS, USA, Tech. Rep. ADA100138, 1981.
- [61] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: Methods and constructs," *Int. J. Ind. Ergon.*, vol. 24, no. 6, pp. 631–645, 1999, doi: 10.1016/S0169-8141(98)00068-7.
- [62] M. Nakayama, K. Takahashi, and Y. Shimizu, "The act of task difficulty and eye-movement frequency for the oculomotor indices," in *Proc. 2nd ACM Symp. Eye Tracking Research & Applications (ETRA '02)*, Mar. 2002, pp. 37–42, doi: 10.1145/507072.507080.
- [63] B. Pan, H. A. Hembrooke, G. K. Gay, L. A. Granka, M. K. Feusner, and J. K. Newman, "The determinants of web page viewing behavior," in *Proc. 3rd ACM Symp. Eye Tracking Research & Applications (ETRA '04)*, Mar. 2004, pp. 147–154, doi: 10.1145/968363.968391.
- [64] C. Girard, J. Ecalle, and A. Magnan, "Serious games as new educational tools: how effective are they? A meta-analysis of recent studies," *J. Comput. Assist. Learn.*, vol. 29, no. 3, pp. 207–219, Jun. 2013, doi: 10.1111/j.1365-2729.2012.00489.x.
- [65] W. L. Bedwell, D. Pavlas, K. Heyne, E. H. Lazzara, and E. Salas, "Toward a Taxonomy Linking Game Attributes to Learning," *Simul. Gaming*, vol. 43, no. 6, pp. 729–760, Dec. 2012, doi: 10.1177/1046878112439444.
- [66] M. Sailer and L. Hommer, "The Gamification of learning: A meta-analysis," *Educ. Psychol. Rev.*, vol. 32, no. 1, pp. 77–112, Mar. 2020, doi: 10.1007/s10648-019-09498-w.
- [67] K. Procci, A. R. Singer, K. R. Levy, and C. Bowers, "Measuring the flow experience of gamers: An evaluation of the DFS-2," *Comput. Hum. Behav.*, vol. 28, no. 6, pp. 2306–2312, Nov. 2012, doi: 10.1016/j.chb.2012.06.039.
- [68] A. Perttula, K. Kiili, A. Lindstedt, and P. Tuomi, "Flow experience in game based learning – a systematic literature review," *Int. J. Serious Games*, vol. 4, no. 1, pp. 57–72, Mar. 2017, doi: 10.17083/ijsg.v4i1.151.
- [69] C. Brom, T. Stárková, E. Bromová, and F. Děchtěrenko, "Gamifying a simulation: Do a game goal, choice, points, and praise enhance learning?," *J. Educ. Comput. Res.*, vol. 57, no. 6, pp. 1575–1613, Oct. 2019, doi: 10.1177/0735633118797330.
- [70] K. Kiili, A. Lindstedt, and M. Ninaus, "Exploring characteristics of students' emotions, flow and motivation in a math game competition," in *Proc. 2nd Int. GamiFIN Conf.*, 2018, pp. 20–29. [Online]. Available: <http://ceur-ws.org/Vol-2186/paper3.pdf>
- [71] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York, NY, USA: Harper & Row, 1990.
- [72] D. C. Moos and E. Marroquin, "Multimedia, hypermedia, and hypertext: Motivation considered and reconsidered," *Comput. Hum. Behav.*, vol. 26, no. 3, pp. 265–276, May 2010, doi: 10.1016/j.chb.2009.11.004.
- [73] A. Landhäuser and J. Keller, "Flow and Its Affective, Cognitive, and Performance-Related Consequences," in *Advances in Flow Research*, S. Engeser, Ed., New York, NY: Springer, 2012, pp. 65–85.
- [74] K. Kiili, "Content creation challenges and flow experience in educational games: The IT-Emperor case," *Internet Higher Educ.*, vol. 8, no. 3, pp. 183–198, Jul. 2005, doi: 10.1016/j.iheduc.2005.06.001.
- [75] K. Kiili and T. Lainema, "Foundation for measuring engagement in educational games," *J. Interact. Learn. Res.*, vol. 19, no. 3, pp. 469–488, 2008.
- [76] S. Barzilai and I. Blau, "Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences," *Comput. Educ.*, vol. 70, pp. 65–79, Jan. 2014, doi: 10.1016/j.compedu.2013.08.003.
- [77] V. Nagalingam and R. Ibrahim, "User experience of educational games: A review of the elements," *Procedia Comput. Sci.*, vol. 72, pp. 423–433, 2015, doi: 10.1016/j.procs.2015.12.123.
- [78] O. Javora, T. Hannemann, T. Stárková, K. Volná, and C. Brom, "Children like it more but don't learn more: Effects of esthetic visual design in educational games," *Br. J. Educ. Technol.*, vol. 50, no. 4, pp. 1942–1960, Jul. 2019, doi: 10.1111/bjet.12701.
- [79] S. Schneider, S. Nebel, and G. D. Rey, "Decorative pictures and emotional design in multimedia learning," *Learn. Instr.*, vol. 44, pp. 65–73, Aug. 2016, doi: 10.1016/j.learninstruc.2016.03.002.
- [80] W. Fias and M. H. Fischer, "Spatial representation of numbers," in *Handbook of Mathematical Cognition*, J. Campbell, Ed., New York, NY, USA: Psychology Press, 2005.
- [81] K. Kiili, K. Moeller, and M. Ninaus, "Evaluating the effectiveness of a game-based rational number training - In-game metrics as learning indicators," *Comput. Educ.*, vol. 120, pp. 13–28, May 2018, doi: 10.1016/j.compedu.2018.01.012.
- [82] M. Ninaus, K. Kiili, J. McMullen, and K. Moeller, "Assessing fraction knowledge by a digital game," *Comput. Hum. Behav.*, vol. 70, pp. 197–206, May 2017, doi: 10.1016/j.chb.2017.01.004.

- [83] R. E. Mayer and L. Fiorella, "Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles," in *The Cambridge Handbook of Multimedia Learning*, R. Mayer, Ed., Cambridge, UK: Cambridge Univ. Press, 2014, pp. 279–315.
- [84] D. D. Salvucci and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proc. 1st ACM Symp. Eye Tracking Research & Applications (ETRA '00)*, 2000, pp. 71–78, doi: 10.1145/355017.355028.
- [85] F. Rheinberg, R. Vollmeyer, and S. Engeser, "Measuring flow-experience," in *Diagnostik von Motivation und Selbstkonzept*, Göttingen, Germany: Hogrefe (in German), 2003, pp. 261–279.
- [86] B. Laugwitz, M. Schrepp, and T. Held, "Construction of a questionnaire to measure user experience of software products," in *Mensch und Computer: Mensch und Computer im Strukturwandel*, München, Germany: Oldenbourg (in German), 2016, pp. 125–134.
- [87] M. Schrepp, A. Hinderks, and J. J. Thomaschewski, "Design and evaluation of a short version of the user experience questionnaire (UEQ-S)," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 4, no. 6, pp. 103–108, 2017, doi: 10.9781/ijimai.2017.09.001.
- [88] M. E. J. Masson, "A tutorial on a practical bayesian alternative to null-hypothesis significance testing," *Behav. Res. Methods*, vol. 43, no. 3, pp. 679–690, 2011, doi: 10.3758/s13428-011-0126-4.
- [89] A. E. Raftery, "Bayesian model selection in social research," *Sociol. Methodol.*, vol. 25, pp. 111–163, 1995, doi: 10.2307/271063.
- [90] G. Rop, A. Schüller, P. P. J. L. Verkoijen, K. Scheiter, and T. van Gog, "Effects of task experience and layout on learning from text and pictures with or without unnecessary picture descriptions," *J. Comput. Assist. Learn.*, vol. 34, no. 4, pp. 458–470, Aug. 2018, doi: 10.1111/jcal.12287.
- [91] G. Rop, M. van Wermeskerken, J. A. de Nooijer, P. P. J. L. Verkoijen, and T. van Gog, "Task experience as a boundary condition for the negative effects of irrelevant information on learning," *Educ. Psychol. Rev.*, vol. 30, no. 1, pp. 229–253, 2018, doi: 10.1007/s10648-016-9388-9.
- [92] F. Guo, Y. Ding, W. Liu, C. Liu, and X. Zhang, "Can eye-tracking data be measured to assess product design?: Visual attention mechanism should be considered," *Int. J. Ind. Ergon.*, vol. 53, pp. 229–235, 2016, doi: 10.1016/j.ergon.2015.12.001.
- [93] B. Strobel, S. Grund, and M. A. Lindner, "Do seductive details do their damage in the context of graph comprehension? Insights from eye movements," *Appl. Cogn. Psychol.*, vol. 33, no. 1, pp. 95–108, 2019, doi: 10.1002/acp.3491.
- [94] M. J. Tsai, A. H. Wu, and Y. Chen, "Static and dynamic seductive illustration effects on text-and-graphic learning processes, perceptions, and outcomes: Evidence from eye tracking," *Appl. Cogn. Psychol.*, vol. 33, no. 1, pp. 109–123, 2019, doi: 10.1002/acp.3514.
- [95] E. L. Law, E. E. Mattheiss, M. D. Kickmeier-Rust, and D. Albert, "Vicarious learning with a digital educational game: Eye-tracking and survey-based evaluation approaches," in *HCI in Work and Learning, Life and Leisure: USAB 2010*, G. Leitner and A. Holzinger, Eds. 2010, pp. 471–488.
- [96] M. Schneider *et al.*, "A validation of eye movements as a measure of elementary school children's developing number sense," *Cogn. Dev.*, vol. 23, no. 3, pp. 409–422, Jul. 2008, doi: 10.1016/j.cogdev.2008.07.002.
- [97] C. Jennett *et al.*, "Measuring and defining the experience of immersion in games," *Int. J. Hum. Comput. Stud.*, vol. 66, no. 9, pp. 641–661, Sep. 2008, doi: 10.1016/j.ijhcs.2008.04.004.
- [98] J.-L. L. Kruger and S. Doherty, "Measuring cognitive load in the presence of educational video: Towards a multimodal methodology," *Australas. J. Educ. Technol.*, vol. 32, no. 6, pp. 19–31, Dec. 2016, doi: 10.14742/ajet.3084.
- [99] H.-C. Liu and H.-H. Chuang, "An examination of cognitive processing of multimedia information based on viewers' eye movements," *Interact. Learn. Environ.*, vol. 19, no. 5, pp. 503–517, Dec. 2011, doi: 10.1080/10494820903520123.
- [100] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *J. Comp. Neurol. Psychol.*, vol. 18, no. 5, pp. 459–482, Nov. 1908, doi: 10.1002/cne.920180503.
- [101] R. E. Mayer, "Computer games in education," *Annu. Rev. Psychol.*, vol. 70, no. 1, pp. 531–549, Jan. 2019, doi: 10.1146/annurev-psych-010418-102744.
- [102] R. M. Krebs, B. H. Schott, and E. Düzel, "Personality Traits Are Differentially Associated with Patterns of Reward and Novelty Processing in the Human Substantia Nigra/Ventral Tegmental Area," *Biol. Psychiatry*, vol. 65, no. 2, pp. 103–110, 2009, doi: 10.1016/j.biopsych.2008.08.019.
- [103] B. C. Wittmann, R. J. Dolan, and E. Düzel, "Behavioral specifications of reward-associated long-term memory enhancement in humans," *Learn. Mem.*, vol. 18, no. 5, pp. 296–300, 2011, doi: 10.1101/lm.1996811.

Manuel Ninaus received his Ph.D. in Psychology from University of Graz, Austria, in 2015. Currently he is PostDoc at the University of Innsbruck, Austria. He is elected board member of the Serious Games Society. His research interests include neuroscience and educational psychology in the context of educational games and learning analytics.

Kristian Kiili received his Ph.D. from Tampere University of Technology, Finland, in 2006. He is a full professor in game-based learning at Tampere University and holds the title of docent in user centered design. He has been a visiting scholar at Stanford University (20 months). His research interests are related to learning sciences, game-based learning, user experience, and learning analytics.

Guilherme Wood received his Ph.D. in Psychology from RWTH Aachen, Germany, in 2005. He is a full professor in neuropsychology at University of Graz, Austria. His research interests include neuroscience in the context of neuropsychological rehabilitation and numerical cognition.

Korbinian Moeller received his Ph.D. in Psychology from University of Tübingen, Germany, in 2010. Currently he is a full professor at the Loughborough University, United Kingdom. He investigates the development / acquisition of numerical competencies in formal as well as informal educational contexts to evaluate the cognitive but also neural correlates of numerical learning.

Silvia Erika Kober received her Ph.D. in Psychology from University of Graz, Austria, in 2012. Currently she is Senior Scientist at University of Graz, Austria. Her research interests include neuroscience in the context of developing and evaluating new and innovative rehabilitation tools.