Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge

Miia Ronimus, Janne Kujala, Asko Tolvanen, Heikki Lyytinen

Agora Center, P.O. Box 35, University of Jyväskylä, FI-40014 Jyväskylä, Finland
Department of Mathematical Information Technology, P.O. Box 35, University of Jyväskylä, FI-40014 Jyväskylä, Finland
Department of Psychology, P.O. Box 35, FI-40014 University of Jyväskylä, FI-40014 Jyväskylä, Finland

Abstract
This study investigated the effects of two game features (the level of challenge and the reward system) on first and second graders’ engagement during digital game-based learning of reading. We were particularly interested in determining how well these features managed to maintain children’s engagement over the 8-week training period. The children (N = 138) used GraphoGame, a web-based game training letter–sound connections, at home under the supervision of parents. Data regarding the children’s gaming and engagement were stored on the GraphoGame online server. A 2 × 2 factorial design was used to investigate the effects of the level of challenge (high challenge vs. high success) and the presence of the reward system (present vs. absent). Children’s engagement was measured by session frequency and duration and through an in-game self-report survey that was presented at the end of the each session. According to the results, the children enjoyed GraphoGame but used it less frequently than expected. The reward system seemed to encourage the children to play longer sessions at the beginning of the training period, but this effect vanished after a few sessions. The level of challenge had no significant effect on children’s engagement. The results suggest a need to investigate further the effectiveness of various game features in maintaining learner’s engagement until the goals set for learning are achieved.

1. Introduction
Playing computer games is a popular leisure-time activity among young children. In Finland, 84 percent of first graders play computer games at least sometimes, 31 percent every day (Hirvonen, 2012). Computer games can hold children’s attention for hours a day, so it is not surprising that many parents and teachers are interested in their potential as educational and motivational tools. Several studies suggest that children enjoy computer-based learning tasks more than traditional learning tasks (Barrera, Rule, & Diemart, 2001; Rosas et al., 2003; Seymour, Sullivan, Story, & Mosley, 1987; Tüzün, Yilmaz-Soylu, Karakus, İnal, & Kizilkaya, 2009; Wrzesien & Raya, 2010). Also, children seem to concentrate better while engaged in computer-based learning than in traditional school tasks (Clarfield & Stoner, 2005; MacArthur, Haynes, & Malouf, 1986; Mautone, DuPaul, & Jitendra, 2005; Ota & DuPaul, 2002).

Despite the apparent motivational appeal of digital learning, little experimental research concerning the long-term development of engagement during digital game-based learning has been conducted. When computer-based learning activities or games are introduced to young learners, they typically trigger curiosity and interest (Mitchell, 1993; Seymour et al., 1987). However, this type of interest is situational, triggered by environmental stimuli and it may or may not last over time (Hidi & Renninger, 2006). Some earlier studies imply that interest triggered by educational software may be short-lived (Goodwin, Goodwin, Nansel, & Helm, 1986; Kerawalla & Crook, 2005), but also relatively long-lasting interest has been observed (Rosas et al., 2003). If the situational interest triggered by a novel learning tool can be maintained for a longer period of time, it may eventually develop into what Hidi and Renninger (2006) called individual interest, a relatively enduring predisposition to seek repeated reengagement with particular content over time. Interest often has a positive impact on learning. For example, interest in recreational reading has been found to predict the development of reading ability during the first years of school.
In contrast, children who spend little time engaged in reading activities may develop a reading disability at school age, even when they have no cognitive risk factors in their background (Eklund, Torpps, & Lyytinen, 2013).

1.1. Engaging the learner

Intrinsic motivation refers to a situation where actions are performed in the absence of any apparent external contingency, that is, an intrinsically motivated person finds the activity rewarding in itself and does not expect to gain anything particular, such as extrinsic rewards, from it (Deci & Ryan, 1980). The concept of flow is related to intrinsic motivation because, when in the state of flow, a person’s activity is autotelic, rewarding in itself, and extrinsic outcomes of the activity have little personal significance (Nakamura & Csikszentmihalyi, 2002). The concept of flow emphasizes the subjective experience of this state, which is characterized by intense and focused concentration, the merging of action and self-awareness, a sense of control, a loss of reflective self-consciousness, and a distortion of temporal experience (Nakamura & Csikszentmihalyi, 2002).

According to the theory by Malone (1980), computer games that provide challenge, that include elements of fantasy, and that rouse the curiosity of the learners are intrinsically motivating. Lepper and Malone (1987) also include control, suggesting that young learners should have opportunities for making choices about instructionally irrelevant aspects of the activity, such as types and names of characters and fantasy elements. Providing choice within a digital game potentially enhances the player’s perception of autonomy, which in turn has been shown to increase intrinsic motivation (Ryan, Rigby, & Przybylski, 2006). The results of experimental studies support these design principles. For example, Parker and Lepper (1992) found that embedding computer-based instruction in a fantasy context had positive effects both on children’s learning and interest in the learning task. In another study, Cordova and Lepper (1996) found that in addition to fantasy embellishments, the provision of task-irrelevant choices, such as a choice of type of spaceship in a computer game, improved children’s intrinsic motivation and learning of the subject matter. More recently, Sweetser and Wyeth (2005) have developed a series of criteria for designing engaging computer games on the basis of the flow theory. According to their GameFlow model, the elements that produce flow in computer games include concentration, challenge, control, clear goals, feedback, immersion, and social interaction. This model has been used in the evaluation of computer games for different user groups. For example, Inal and Cagiltay (2007) examined the flow experiences of 7- to 9-year-old children in a social game environment and found that the majority of the children emphasized challenge as the most important flow element. Games that included challenge levels and clear, immediate feedback seemed to produce flow experiences most frequently. Ke and Abras (2013) found that students with special learning needs seemed to enjoy games that comprise simple fantasy with instant rewards, and the students also benefited from visual cuing, or feedback. Ke and Abras (2013) also emphasize that it is important that the academic content is integrated in the game-play, not just added to the fantasy context, and that game challenges match the diverse skill levels of the students.

1.2. GraphoGame

GraphoGame (in Finnish “Ekapeli”) is a web-based learning game developed at the University of Jyväskylä in Finland. Its development originated from the findings of the Jyväskylä Longitudinal Study of Dyslexia (see e.g., Lyytinen et al., 2006). In this extensive study, the language and reading development of 200 children, half of whom have a familial risk for dyslexia, has been followed from birth to the age of 15 so far. The findings suggest that difficulties in perceptual differentiation of acoustically close phonemes and, at least partly consequential to that, difficulties in the learning of the connections between letters and sounds are predictors of future reading problems (e.g., Lyytinen, Erskine, Kujala, Ojanen, & Richardson, 2009). GraphoGame was developed to help children overcome these difficulties.

In GraphoGame, the player hears a single phoneme or a longer phonemic unit while a number of graphemes (target and distracters) appear on the screen (Fig. 1). The player’s task is to find the grapheme matching the spoken phoneme and to click it with the mouse. This is followed by immediate feedback that informs the player whether the choice was correct or incorrect. After an incorrect answer, the correct alternative is highlighted before the next trial is presented. When the child has learned the connections between sounds and letters, the

![Fig. 1. An example of a GraphoGame training task in which the player hears a letter sound and is expected to find the matching letter from the alternatives shown on the screen.](image-url)
Several versions of GraphoGame have been developed for different user groups, such as for immigrant children who are learning to read in Finnish and for native-speaking children who need remedial training in reading fluency. GraphoGame is not intended for replacing the regular reading instruction children receive at school, but to work as a supplementary tool. In the context of remedial instruction, a recommended way of usage and its validation has been illustrated in a recent study by Saine, Lerkkanen, Ahonen, Tolvanen, and Lyytinen (2011). To date, GraphoGame has been available free of charge for all children in Finland. Other language versions of GraphoGame have been developed for research purposes. Demos of the English version are available at the GraphoGame web site (http://info.graphogame.com/). The Grapho Learning Initiative (see http://info.graphogame.com/partners) aims to make GraphoGame available to children throughout the world.

If, as a registered GraphoGame user, the child’s guardian or teacher signs into the game, the child’s usage data are stored on the GraphoGame web server. The usage data consist of detailed logs that include information about session times and durations and player’s responses to the learning tasks, tests, and surveys. Earlier studies suggest that GraphoGame has a positive impact on children’s reading acquisition (Lyytinen et al., 2009; Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007). In a longitudinal intervention study by Saine et al. (2011), Finnish first graders (7-year-old children) with low reading skills (n = 50) were randomly assigned either to traditional remedial training or remedial training that included sessions with GraphoGame. The remedial training took place at school and consisted of four weekly sessions of 45 min, held over a period of 28 weeks. The results suggested that the children receiving training that included GraphoGame sessions made more gains in reading and spelling than the children receiving traditional small-group-based remedial training, and they were still improving in the follow-ups 12 and 16 months later. In the third grade, the GraphoGame group reached the level of normally developing peers, whereas those who received the traditional remedial training were lagging behind. In another study Hintikka, Aro, and Lyytinen (2005) found that a 6-week training period with GraphoGame improved 7-year-old children’s ability to name letters and that this intervention was more effective than traditional instruction for the children with low phoneme awareness skills and poor attention. It has also been found that a brief training (10 sessions of 5–10 min) with GraphoGame can improve the child’s reading speed of syllables practiced in the game (Heikkilä, Aro, Näärhi, Westerholm, & Ahonen, 2013).

The effects of two English versions of the GraphoGame have been studied by Kyle, Kujala, Richardson, Lyytinen, and Goswami (2013). The children were 6- to 7-year-old poor readers. One of the game versions (GG Phoneme) trained connections between letters and sounds and the other (GG Rime) used rhyming word families to introduce and reinforce grapheme–phoneme connections. The children received approximately 11 h of training over 12 weeks as a supplement to the instruction provided by their teachers. Both versions turned out to be more effective than typical instruction alone, but players of GG Rime showed greater improvement. It is important to note that Finnish and English differ in the transparency of the writing system (see Aro, 2006). Finnish has a regular grapheme–phoneme correspondence, and the connections can be trained by matching single sounds to letters. However, practicing larger units may be more effective in languages that are more inconsistent at the level of single grapheme. It has also been found that a brief training (10 sessions of 5–10 min) with GraphoGame can improve the child’s reading speed of syllables practiced in the game (Heikkilä, Aro, Näärhi, Westerholm, & Ahonen, 2013).

The effects of GraphoGame on children’s neural processing have been studied by electroencephalography and functional magnetic resonance imaging. Brem et al. (2010) found that an 8-week GraphoGame intervention for nonreading 6-year-olds resulted in an initial sensitization to print on specific areas within the occipito-temporal cortex, which are often engaged in print processing in readers. Lovio, Halttunen, Lyytinen, Näättänen, and Kujala (2012) found that a short, 3-week intervention improved 6-year-old children’s reading related skills such as phonological processing and writing of words and pseudowords. Furthermore, functional changes in the brain, reflected in increased mismatch negativity amplitudes for the vowel and vowel-duration changes, were also observed in the intervention group, suggesting that playing GraphoGame improved children’s phonological discrimination.

1.3. The current study

In this study, we aim to find out how certain features of GraphoGame affect children’s engagement during game play in the home environment. Maintaining children’s motivation to use GraphoGame is critically important, because regular and intensive training is often required for significant learning gains to occur (Saine et al., 2011). We were interested in the children’s motivation to play the game at home, because school days in Finland are short and many teachers find it challenging to include regular GraphoGame sessions into the daily schedule. Also, relatively little is known about educational gaming at home, as most research is done in formal learning environments such as schools.

The data were collected via the Internet and saved on the GraphoGame server. This method enabled us to investigate GraphoGame usage at home with minimal intrusion. The children were able to use the game as a part of their daily activities, without knowing they are participants in a study. This improves the ecological validity of the study, which is often a problem in laboratory-based experiments. This does, however, limit our chances of controlling certain extraneous variables, such as the quality of parental support, which may influence children’s motivation to use GraphoGame, and weaken the relationship between the game features and engagement. There is variation in the learning conditions at children’s homes, which will add error to the results. However, we determined that it is worth studying whether the effects of the game features would be apparent despite the uncontrolled aspects of the children’s home environments. This research method is a cost-effective way to study the impacts of various game design features, especially their long-term impacts, with the benefit of generalizability to real-life.

We used a 2 × 2 factorial design to compare the effects of two game features, the presence of a reward system and the level of challenge on children’s engagement. We used both objective (usage data) and subjective (ratings of enjoyment and interest) measures of engagement, which allowed us to assess the development of engagement during the training period.

All game versions used in the study consisted of letter–sound connection tasks that were presented in sets of 20. However, if the reward system was present, the player received a game token after completing each set of 20. After collecting five tokens, the player gained access to
the new reward game levels. Playing a reward game cost five tokens. After spending all the game tokens, the player needed to return to the basic game and collect more tokens to regain access to the reward games. Initially, the player could choose from only two new reward games, but the number of available reward games increased as the total playing time increased. The reward games taught the same content as the main game, but in different format and context. For example, in one of the reward games the player tried to hit the target letters by navigating a spaceship (Fig. 2). Using rewards to motivate children has been criticized, because extrinsic rewards can undermine intrinsic motivation if the receiver feels that his or her behavior is being controlled by the rewarder (Deci, Koestner, & Ryan, 1999). However, it is unclear whether these findings also apply to virtual environments. According to Deci et al. (1999), tangible rewards, such as candy or money, have the most detrimental effect on intrinsic motivation. Also, the reward games in GraphoGame were not solely rewards, but they also added elements of choice and fantasy to the game, which according to earlier studies are beneficial to intrinsic motivation (e.g. Cordova & Lepper, 1996).

The second game feature we investigated was the level of difficulty of the game. We were interested in determining if a high success rate would increase children's motivation to play, and if high challenge would result in more effective learning. All game versions had an adaptive level of difficulty, which means that correct answers were followed by more difficult trials and incorrect answers by easier trials, but the configuration of the adaptation logic varied between the game versions, resulting in differing levels of overall challenge. The educational content itself was identical among the game versions.

The educational content of the subsequent trials in the game was determined by a Bayesian-probability-model-based adaptation technique developed by Kujala, Richardson, and Lyytinen (2010). All educational content of the game was categorized by level of difficulty, the easiest categories containing single letters and the most difficult categories comprising sets of long and complex pseudowords.

In the versions set to high challenge adaptation, the presentation of the trials was based on Kujala’s (2010, 2011) goal of maximizing the measurement information about the child’s skill level. Thus, the target difficulty level of every subsequent trial was placed at approximately the middle of the range of uncertainty of the child’s level given the results of the previous trials. For the first few trials, this appeared like a binary search: the trials started from the middle of the list (category 43, see Fig. 3), and in the case of a correct answer, the next trial was selected from around the midpoint of the upper half of the list; whereas, in the case of an incorrect answer, the next trial was selected from around the midpoint of the lower half of the list. The later trials deviate from this pattern due to the inherent randomness in the child’s results for content near her level (i.e., a correct or incorrect answer does not completely exclude the lower or upper part of the list). Over several sessions, this adaptation resulted in a mean success rate of 60.

In the high success adaptive versions, the goal of maximizing the measurement information was balanced with a goal of high success rate (Fig. 3; see Kujala, 2010, 2011 for the theoretical details). This resulted in presentation of trials closer to the easier end of the range of uncertainty of the child’s level and, consequently, a consistently higher level of correct answers. Over several sessions, this adaptation resulted in the mean success rate of 80.

1.4. Research questions

We investigate whether certain game features (the reward system and the level of challenge) have an effect on children’s engagement in playing GraphoGame. Specifically we try to answer following questions:

- Can children’s engagement be triggered and maintained by a novel reward system that adds more choice, variety, and fantasy elements in the learning game?
- What is the appropriate level of challenge for children playing GraphoGame? Does high success rate and positive feedback maintain children’s engagement better, or could high challenge and potentially faster progress in learning have a more positive impact?
- Is there an interaction between these two game features? For example, would the children enjoy the high challenge more when the new reward system is on, because a more fun experience might make the challenges more enjoyable as well?

![Fig. 2. An example of a reward game in GraphoGame.](image-url)
2. Methods

2.1. Participants

The participants were recruited by sending an e-mail to all registered GraphoGame users. The study was aimed at first graders who were unable to read and in need of support in their reading acquisition. The GraphoGame users who responded with interest in participating were sent further information about the study and a consent form to be signed by the child’s parent or guardian. The consent form also included questions about the child’s education, development, and previous experience of GraphoGame and other digital games.

All children whose parents returned an affirmative consent form and who played the game at least once at home during the training period were included in the study (N = 138). Typically there was one child participating from each family, but three families had two participants. There were more boys (n = 82) than girls, most likely because boys typically have more problems with reading than girls (e.g. Rutter et al., 2004). The mean age of the participants was 7.36 years (range 6.82–8.01). Most of the children (94.2%) were in the first grade, but some second graders were also accepted in the study because they were still struggling at reading acquisition. Some of the children (14.5%) attended a special education class or received mainstream instruction in a small teaching group. Because the parents were already registered GraphoGame users, it is not surprising that the majority of the children (88.4%) had previous experience playing GraphoGame. Nearly all of the children (98.6%) played digital games in their leisure time. A full 30.4 percent of parents told that their child had experienced some learning or developmental problems. Most common were problems with language development (14.5%), motor skills (10.1%), and attention (7.2%). Some of these problems were suspected by parents, but not diagnosed by a professional. Generally, the participants were a representative sample of the children that typically use GraphoGame in Finland.

2.2. Measures

The data concerning children’s gaming were automatically recorded and stored on the GraphoGame online server while the child played the game at home. The usage data were later retrieved from the server for the analysis.

Children’s enjoyment of GraphoGame was measured by an in-game survey that was presented at the end of each play session, when the child was exiting the game. A female voice asked the child to rate how much he or she had enjoyed playing the game that time, and five faces, ranging from having a big smile to a big frown, appeared on the screen. The child rated his or her enjoyment by clicking one of the faces. The child’s interest in reading-related activities was measured by a similar in-game survey that was presented twice: at the beginning and end of the training period. The second survey appeared automatically 8 weeks after the first play session. In the task, a female voice asked the child to tell how much he or she liked school tasks that involved reading, and the child answered by clicking one of the five faces on the screen.

Parents were asked to evaluate children’s motivation and concentration during GraphoGame play via two questions on an online questionnaire sent to the parents at the end of the study. Motivation was measured by the question “How eagerly did the child play

Fig. 3. The adaptation logics that determined the level of challenge in the GraphoGame versions used in the study. The high success version is presented above and the high challenge version below. Each circle shows the category number of the subsequent trial to be presented after a certain sequence of correct (1) or incorrect (0) answers as indicated by the arrows.
GraphoGame during the study?” and concentration by “How well did the child concentrate while playing GraphoGame?” The parents answered using a 5-point scale, ranging from “very eagerly/well” to “very reluctantly/poorly”.

Reading performance was measured by two in-game tasks that were presented at the beginning and the end of the training period. In the first task, the player heard 44 increasingly difficult words, one at a time, and was asked to select the corresponding written word from the six alternatives shown on the screen. No feedback was given. The words were presented in sets of 10, and if the child made fewer than two correct choices in the 10 words presented, the task was discontinued. The second task was similar, but included meaningless pseudowords. The mean of the task scores was used as the child’s reading performance score in the analysis. An earlier study conducted in the GraphoGame project indicates that the results of in-game reading tests such as these have a strong correlation with those of traditional reading tests administered by a researcher (Heinola, Latvala, Heikkilä, & Lyytinen, 2010).

### 2.3. Procedure

Each parent who returned the signed consent form was sent a download link to GraphoGame and instructions for the training period by electronic mail. When the game was used for the first time, the user (parent) was randomly assigned to one of four game groups. Eventually, 36 children were playing with the high challenge plus reward system version; 33 children with the high challenge, no reward system version; 32 children with the high success plus reward system version; and 37 children with the high success, no reward system version. The group sizes are slightly unequal, because there were three families with two participating children, and these families were assigned in the first and fourth groups.

The children were expected to play the game for 8 weeks. The parents were told that the children should play at least twice a week during this time, preferably daily, approximately 10–15 min at a time. The parents also were reminded that using the game should be voluntary to the child and that a reluctant child should not be pushed into playing the game. Ten weeks after the first play session, the parents were sent an online questionnaire in which they were asked to evaluate their experiences with GraphoGame and the children’s attitude toward playing the game. The response rate to the questionnaire was 83 percent (112 out of 135 parents filled in the questionnaire).

### 2.4. Statistical analyses

The data about children’s enjoyment of GraphoGame and session durations were hierarchical in nature and comprised multiple within-individual measures about enjoyment and duration of each play session. All children who played more than one session (n = 127) and for whom, therefore, the within-individual change could be estimated, were included in the analysis. The maximum number of play sessions varied between individuals, and because of this variation, the serial number of the play session was recoded to vary from zero (first play session) to one (last play session) for each individual. This transformation preserved the proportional intervals between measurements, and it enabled a linear measurement of the amount of change between the first and last play session.

To explain within-individual and between-individual variation in enjoyment we used a two-level random coefficient linear growth model with a maximum likelihood with robust standard errors estimation method (Muthén & Muthén, 1998–2007). In the within-individual level, the transformed serial number of the play session was used for predicting the child’s rating of enjoyment, producing the regression coefficient that varied randomly from individual to individual. In the between-individual level analysis, the mean value and variance of the initial level of enjoyment as well as the mean value and variance of the linear change of enjoyment were estimated. The variance of the initial level of enjoyment and the variance of the linear change of enjoyment was predicted by two game features (high/low level of challenge and the presence/absence of the reward system). The same method was used to explain the within-individual and between-individual variation in the durations of the play sessions. The analyses were generated using Mplus Version 6 (Muthén & Muthén, 1998–2007).

To analyze the effects of game features on parent-reported motivation and concentration and on the total playing time, two-way analyses of variance were used. All children who played the game at least once and for whom these measures were available (n = 138 for total playing time, n = 113 for motivation and concentration) were included in these analyses. Repeated measures analyses of variance were used to analyze the effects of game features on the development of children’s reading interest and reading performance. All children who completed the assessment tasks both in the beginning and at the end of the training period were included (n = 72 for reading interest and n = 68 for reading performance). These analyses were performed with IBM SPSS Statistics 20.

### 3. Results

#### 3.1. Enjoyment

Children’s enjoyment of GraphoGame during the training period was evaluated by an in-game question presented after each play session. We used the linear growth model to estimate the initial level of children’s enjoyment and the amount of change in enjoyment over the
course of the training period. According to the results, children’s initial level of enjoyment was high (Estimate = 4.05, SE = .09, p < .001), and no significant change in the level of enjoyment occurred during the training period (Estimate = .03, SE = .12, p = .794). However, there was between-individual variance both in the initial level of enjoyment (Estimate = .36, SE = .07, p < .001) and in the change of enjoyment (Estimate = .40, SE = .16, p = .013). We tried to explain this variance by the level of challenge of the game and the presence of reward games, but no significant effects were found. The results are presented in Table 1.

3.2. Playing time

The average total playing time per child was 179.60 min (SD = 141.37), when all the children who played the game at least once were included. According to the analysis of variance, neither rewards, $F(1, 134) = 2.42, \eta^2_p = .02, p = .122$, the level of challenge, $F(1, 134) = 2.43, \eta^2_p = .02, p = .121$, nor their interaction $F(1, 134) = .04, \eta^2_p = .00, p = .838$ had a significant effect on the total playing time.

Next we analyzed the variation of session durations over the course of the training period. Initially, the play sessions lasted on average 13.42 min (SE = .68, p < .001). However, play sessions tended to shorten over time (Estimate = −2.63, SE = .80, p = .001). We found a significant amount of between-individual variance in the session durations (Estimate = 24.08, SE = 5.70, p < .001) and in the change of durations over time (Estimate = 14.83, SE = 7.09, p = .036). According to the analysis (Table 2), the presence of the reward games initially increased the duration of a play session (p = .006), but, again, the play sessions tended to shorten over time (p = .015). This effect is illustrated in Fig. 4.1.

3.3. Parent’s evaluation of motivation and concentration

We also asked parents to evaluate their children’s motivation and concentration during GraphoGame sessions. The mean rating was 3.34 (SD = .91) for motivation and 3.84 (SD = .69) for concentration on a 5-point scale. The analysis of the results did not reveal any significant effects of the game features on children’s motivation, but the presence of the reward system had a significant effect on concentration, $F(1, 109) = 5.24, \eta^2_p = .05, p = .024$. According to the parents, the children playing with the reward system versions were concentrating better ($M = 3.98, SD = .55$) than the children playing without the reward system ($M = 3.70, SD = .78$).

3.4. Interest in reading

Children’s interest in reading was measured by an in-game question at the beginning and end of the training period. The ratings of the 72 children who answered the questions both times were included in the following analysis. Repeated measures ANOVA found no change in children’s reading interest from the beginning of the training period ($M = 3.86, SD = 1.45$) to the end of the training period ($M = 3.76, SD = 1.43$), $F(1, 68) = .04, \eta^2_p = .00, p = .841$. The only significant interaction was between time and the level of challenge, $F(1, 68) = 6.13, \eta^2_p = .06, p = .037$. Of the 72 children, those who played with the high challenge version reported higher interest at the beginning ($M = 4.11, SD = 1.24$) than in the end ($M = 3.65, SD = 1.48$) of the training period, but this difference was only marginally significant, $F(1, 34) = 3.23, p = .081$. We found no difference in the ratings among those who played with the high success version.

3.5. Reading performance

The following analysis contains the results from the 68 children who completed the in-game reading tests both at the beginning and at the end of the training period. Although we found a significant effect of test session on children’s reading performance, $F(1, 64) = 16.34, \eta^2_p = .20, p < .001$, with the children’s scores improving from the beginning of the training period ($M = 19.48, SD = 12.08$) to the end ($M = 24.92, SD = 12.16$), no significant interactions between test session and the game features were found.

4. Discussion

We investigated if certain game features have an impact on children’s engagement during digital game-based learning of reading and if they help to sustain children’s engagement over time. In this study children’s own and parent’s reports suggested that children enjoyed playing GraphoGame and that the enjoyment level was not affected by the game features investigated in the present study. However, when we looked at the times spent playing the game, we found that the novel reward system initially seemed to trigger children’s interest,

---

Table 2
The results of a two-level random coefficient linear growth model estimating the initial level and change of GraphoGame session durations during the training period with the level of challenge and the presence of reward games as predictors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Initial level</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>Challenge</td>
<td>−.03</td>
<td>1.68</td>
</tr>
<tr>
<td>Rewards</td>
<td>4.97</td>
<td>1.81</td>
</tr>
<tr>
<td>Challenge × Rewards</td>
<td>−1.14</td>
<td>2.57</td>
</tr>
<tr>
<td>$R^2 = .09$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

1 There were many children who used GraphoGame only a few times during the training period. Because these children’s data may not give sufficient information of the effects of long-term use of GraphoGame, we also performed the analyses described above including only the players with at least 10 play sessions ($n = 65$). The results did not differ from the ones we report here.
showing as longer playing duration during the first few sessions, but subsequent session durations were reduced back to typical level. Moreover, there were no differences in the total playing time between the reward and nonreward groups.

Overall, the total playing time was quite low regardless of the game version, about 3 h during 8 weeks, which is not enough especially for children with serious difficulties learning to read. Also, only about half of the children completed the training by responding to the post-training interest survey and the reading test. The informal feedback received from the parents after the study suggests that, in some cases, gaming ceased because the children learned to read and the training became unnecessary, but some of the children quit playing because the game lost its appeal. These findings seem to support those of Kerawalla and Crook (2005), who provided families with novel educational CD-ROMs to use at home and found that, despite initial enthusiasm, the time children spent using the software declined quickly after the first week.

Considering the design principles of Malone (1980), Lepper and Malone (1987), and Sweetser and Wyeth (2005), we noticed some shortcomings in the control, goal setting, and feedback features of the GraphoGame versions used in this study. First, there were generally few opportunities for choice and for impacting the game world. In the reward system versions the child was able to choose between seven alternative reward games, but this may not have been enough. The player needed to collect five tokens before a choice could be made, which may have felt more like a restriction of choice than an opportunity, resulting in waning intrinsic motivation (Ryan et al., 2006). Second, the criterion for multiple clear goals (Lepper & Malone, 1987) was not fulfilled. In the nonreward versions, the only goal was to finish the game tasks and there were no long-term goals. In the reward game versions, the player could aim at gaining access to all reward games, which may have taken a few sessions, but, after this was accomplished, there were no more long-term goals. Third, the game versions did not provide feedback about the player’s progress in the game or about skill development, which did not help the children understand how playing the game benefited their development in reading. Fourth, GraphoGame did not include features that would have encouraged collaboration, for example between the player and the supervising parent, but was designed as a single-user game. Lack of parental involvement in the learning process may be one of the key factors why access to educational software does not seem to engage children in learning at home (Kerawalla & Crook, 2005). These shortcomings in the game design may help to explain the declining engagement over the course of the training period.

Children’s own ratings of enjoyment were not affected by game features or the passing of time. Children generally gave very positive ratings, which reflects children’s tendency to favor the positive end of the rating scale when they are asked to evaluate computer games (Read & MacFarlane, 2006; Sim, MacFarlane, & Read, 2006) often making it difficult to find any differences between game versions. The method (a single question) may also have been too simple to completely capture the nature of children’s feelings about playing GraphoGame. However, also parental reports suggested that children’s motivation to play GraphoGame was generally quite high. More interesting, the parents whose children played with the reward versions, reported that their children concentrated better during the play sessions, than did the parents whose children played without the reward system. The fantasy elements and the greater variety of learning tasks may have increased children’s attention during gaming. However, we cannot rule out the possibility that the parents themselves thought the reward versions were more interesting, which affected their perception of their children’s behavior.

The level of challenge did not affect children’s engagement in playing, and we found no interactions between the level of challenge and the reward system. Both investigated levels of challenge (success rates of 60 and 80) may have been appropriate for the children or, at least, the difference between them may not have been significant enough to produce differences in engagement. Moreover, the level of challenge itself may not have been as important as the child’s perception of the level of challenge. As the flow theory suggests, the balance between perceived challenge and perceived skill is the essential condition of flow and both should be above the average level (Nakamura & Csikszentmihalyi, 2002). It seems that to be able to draw any conclusions as to which level of challenge is most engaging to the children, we should gather information about children’s own perceptions of their skills and challenges.

We found a marginal decrease in the reading interest of the children who used the high challenge version of the game. This may imply that the version was too difficult for the children although this did not show in the playing times or the ratings of enjoyment. Many failures and negative feedback during learning of reading at school may have a negative effect on children’s self-concept as a learner (Aunola, Leskinen, Onatsu-Arvilommi, & Nurmi, 2002; Chapman & Tunmer, 2003) and may lead to negative attitudes toward reading (McKenna, Kear, & Ellsworth, 1995; Morgan, Fuchs, Compton, Cordray, & Fuchs, 2008). However, the evaluation of reading interest by a single question within a computer game is problematic and may not give a comprehensive view of how the children really feel about reading in their everyday life. More research is needed to find out whether negative feedback in a learning game can indeed decrease children’s reading interest.

Children’s performance in the in-game reading test clearly improved regardless of the game version used. The high challenge game version exposed children to new learning content more quickly than the high success version, but this did not seem to benefit children’s
learning. Also, it is not surprising that the reward system did not benefit learning because it did not increase the total playing time. The generally low playing times and frequencies may explain why no differences emerged between the game versions.

This study shows that web-based methods can be used to investigate children’s learning and motivation in the home environment. However, it is possible that these results were specific to this learning condition (at home under parental supervision) which is why these findings should not be generalized to other learning conditions without caution. In more formal learning settings, there are often (a) fewer activities competing for children’s time and interest and (b) restrictions on the time allotted to playing learning games, which means that the playing time would not be a good measure of the player’s motivation. Moreover, most of the participants of this study had played some earlier version of GraphoGame before, which may have affected the results. The reward system was a new game feature, and the children probably enjoyed exploring it initially, which resulted in longer playing times during the first few sessions. However, the effect could have been different if GraphoGame had been totally new to all of these children. Also, the participants of the study had parents who were eager to support their children’s learning and were able to provide the children the computer equipment needed for the training. The results may not therefore apply to children from more disadvantaged backgrounds.

Conducting the study via the Internet caused some limitations concerning the assessment methods. The enjoyment and interest surveys were technically limited to the GraphoGame environment, and the answers might have been somewhat different if they had been obtained by more traditional methods, such as by interviewing children face to face. We were not able to monitor how well the children understood the instructions and the questions presented by the game. The surveys were deliberately kept as short and simple as possible, but this limited the type of data that we were able to collect. It is important to continue developing reliable and valid assessment methods for studying young children’s motivation in online environments.

As a conclusion, this study suggests that although fantasy elements and novel task types may increase children’s engagement in playing digital learning games, this effect may not be long-lasting, at least if there are shortcomings in the game design. When developing digital learning games for young children with poor academic skills, it is important that the game fulfills the design principles that support long-term engagement and interest in the subject matter. To find out what kinds of features are most efficient at achieving this goal, more studies with long-term observations of player engagement and interest are needed.

Acknowledgments

We thank Rachel Ferlatte Kuisma for helpful comments and assistance in the language editing of this paper.

References


