The Role of Curiosity-Triggering Events in Game-Based Learning for Mathematics

Pieter Wouters, Herre van Oostendorp, Judith ter Vrugte, Sylke Vandercruysse, Ton de Jong, and Jan Elen

Abstract In this study, we investigate whether cognitive conflicts induced by curiosity-triggering events have a positive impact on learning and motivation. In two experiments, we tested a game about proportional reasoning for secondary prevocational students. Experiment 1 used a curiosity-triggering vs. control condition pretest–posttest design. The control condition received the game without curiosity-triggering events. The results provided evidence that the game improves proportional reasoning skills. Although game performance was positively related to posttest performance, the hypothesized higher increase in learning and motivation after curiosity-triggering events was not found. Based on the results of Experiment 1, the game was adapted. Experiment 2 showed basically the same pattern of results, but we did not find a learning effect after playing the game. In the Discussion, we suggest additional research with think-aloud and/or eye-tracking to map the actual thoughts after the curiosity-triggering events. In addition, we propose some alternative implementations to evoke cognitive conflicts.

Keywords Curiosity • Game-based learning • Cognition • Motivation • Mathematics

The last decade shows an increasing attention for the use of computer games in learning and instruction, often referred to as serious games or game-based learning (GBL). However, recent meta-analytic reviews have shown that GBL is only moderately more effective and not more motivating than traditional instruction (Sitzmann, 2011; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013).

P. Wouters (✉) • H. van Oostendorp
Institute of Information and Computing Sciences, Utrecht University, P.O. Box 80.089, Utrecht 3508 TB, The Netherlands
e-mail: p.j.m.wouters@uu.nl; H.vanoostendorp@uu.nl

J. ter Vrugte • T. de Jong
Department of Instructional Technology, University of Twente, Enschede, The Netherlands
e-mail: j.tervrugte@utwente.nl; A.J.M.deJong@utwente.nl

S. Vandercruysse • J. Elen
Center for Instructional Psychology & Technology, KU Leuven, Leuven, Belgium
e-mail: sylke.vandercruysse@kuleuven-kulak.be; jan.elen@ppw.kuleuven.be

J. Torbeyns et al. (eds.), *Describing and Studying Domain-Specific Serious Games*, Advances in Game-Based Learning, DOI 10.1007/978-3-319-20276-1_12

For example, Wouters et al. reviewed 39 empirical studies for their meta-analysis and found a moderate effect size for learning of $d = .29$ in favor of GBL. Likewise, they found a moderate, but statistically nonsignificant, effect for motivation in favor of GBL.

A plausible explanation for the limited effect of GBL on learning is that players act in computer games and see the outcome of their actions directly in changes in the game world. This may lead to a kind of intuitive learning: players know how to apply knowledge, but they cannot explicate it. In other words: they don't necessarily acquire the underlying rules (Leemkuil & de Jong, 2011). It is possible that studies therefore find no relation between success in the game and success on an explicit knowledge test. Yet, it is important that learners articulate and explain their knowledge because it urges them to *organize* new information and *integrate* it with their prior knowledge (Mayer, 2011, Wouters, Paas, & van Merriënboer, 2008).

Sense of control regarding decisions during game play (e.g., when to leave the game, go back in the game, conduct specific actions in the game) is deemed an important determinant for intrinsic motivation (Deci, Koestner, & Ryan, 1999; Ryan, Rigby, & Przybylski, 2006). It is neglected that GBL often lacks control because the game that is used and the playing time are generally defined by the curriculum and not by the player (Wouters et al., 2013), resulting in low motivation. In addition, it is plausible that the lack of motivational appeal in GBL environments is a reflection of the fact that the world of game design and instructional design are not yet integrated (Wouters, van Oostendorp, Boonekamp, & van der Spek, 2011). Take, for example, the situation in which a designer uses a pop-up screen with a message that prompts the player to reflect. From an instructional design perspective such a focus may yield learning, but it is also likely that such an intervention will disturb the flow of the game and consequently undermine the entertaining nature of the game and reduce motivation and learning as well.

The question raised in this study is how we can stimulate players to engage in relevant cognitive processes that foster learning without jeopardizing the motivational appeal of the game. In this respect, the role of curiosity is often neglected in GBL. It is interesting for two reasons. To start with, curiosity is regarded as a motivator for active (cognitive) explorative behavior (cf. Berlyne, 1960; Litman, 2005; Loewenstein, 1994). Second, active exploration is a key aspect of contemporary computer games (Dickey, 2011) which might be beneficial for learning.

Curiosity

In his review, Loewenstein (1994) proposes an information-gap theory in which curiosity is supposed to arise when attention becomes focused on a gap in one's knowledge. Such an information gap produces the feeling of deprivation labeled curiosity. The curious individual is motivated to obtain the missing information in order to reduce the gap and to eliminate the feeling of deprivation.

An information gap can be interpreted in two ways. The first interpretation is related to conceptual change which can be defined as the process of connecting prior knowledge (ideas, beliefs, knowledge) with new knowledge (Limón, 2001; Merenluoto & Lehtinen, 2004). From an information gap perspective a cognitive conflict can be used as a strategy to promote conceptual change. Such a cognitive conflict can be induced by presenting information that is incongruent with the prior knowledge (e.g., it contradicts prior knowledge). The cognitive conflict is supposed to be a drive for information-seeking questions in order to reconcile the conflict between prior knowledge and new incongruent information (Graesser & McMahen, 1993; Graesser & Olde, 2003).

The second interpretation is related to Berlyne's concept of a cognitive conflict (Berlyne, 1960; Loewenstein, 1994). This construct encompasses "collative" variables such as complexity, novelty, and surprisingness. The presence of these stimulus characteristics would arouse cognitive conflict and stimulate curiosity. In this case, an information gap occurs when stimuli present contradictory or incongruent information. For example, in the game a learner is told that a presented problem can be solved but the game environment appears to offer no opportunities to solve the problem. Although this interpretation is not related to conceptual change, it can also be regarded as a cognitive conflict namely the conflict in the current mental representation of the learner between:

1. The expectations of the learner (based on the assurance that the problem can be solved).
2. The affordances in the learning environment to solve the problem.

The assumption—in line with Jirout and Klahr (2012)—is that this information gap will motivate students to explore the environment and find relevant information for constructing appropriate solution methods. More specifically, we assume that based on Loewenstein (1994) and Berlyne (1960) ideas that externally inducing the information gap will stimulate curiosity, raise arousal, and consequently enhance explorations in the game environment and in this way improve learning.

The advantage of curiosity induced by an information gap is that individuals are cognitively active in an engaging way. Scholars have emphasized the potential of curiosity in GBL (Dickey, 2011; Malone, 1981; Wouters et al., 2011), but empirical research is rather scarce. In this study, we present the results of two experiments in which we investigate the impact of curiosity-triggering events on learning. We expect that these events will improve learning because they will motivate learners to engage in explorative behavior. We used the GBL environment "Zeldenrust" that was specifically developed for learning proportional reasoning in secondary prevocational education (see Vandercruysse et al., this volume). Proportional reasoning was chosen because it is a relevant and well-defined domain and existing methods for proportional reasoning are often ineffective (Rick, Bejan, Roche, & Weinberger, 2012). Furthermore, secondary prevocational education students are often associated with lower levels of motivation for school which makes this population particularly suitable for GBL.

Experiment 1

In this experiment, we examine three hypotheses:

1. Playing the game yields learning in proportional reasoning.
2. Game performance is predictive for (off-line) posttest performance.
3. The game with curiosity-triggering events improves learning and increases motivation more than the game without these events.

Method

Participants and Design

The participants were 67 students (28 male, 39 female) from third-year prevocational education with a mean age of 15.5 (SD = .75) recruited from four classes of four schools.

We adopted a pretest–posttest design with a control (N = 34) and a curiosity (N = 33) group. Participants were randomly assigned to the conditions. Dependent variables were proportional reasoning skill, motivation, and game performance.

Materials

Domain. The domain of proportional reasoning comprises three problem types: comparison problems, missing value problems, and transformation problems (cf., Tourniaire & Pulos, 1985). In comparison problems, learners have to find out whether one proportion is "more than," "lesser-than," or "equal to" another proportion. In missing value problems, one value in one of two proportions is missing. Learners have to find this "missing value" in order to ensure that both proportions are equal. Transformation problems involve two proportions as well and all values are known, but the proportions are not equal. Learners have to find out how much has to be added to one or more of the proportions in order to make both proportions equal (for a more extensive description, see Vandercruysse et al., 2014).

Game environment. In the game Zeldenrust students have a summer job in a hotel. By doing different tasks the students can earn money that they can use to select a holiday destination during the game: the more money they earn, the further they can travel. During the game, the player is accompanied by the manager, a non-playing character, who provides information about the task and gives feedback regarding the performance on the task. The game comprises a base game and several subgames. The base game provides the structure from which the subgames can be started. It allows the player to select an avatar, it presents the context of the game in a sort of animation and features the "Student room" from which the student can control the game (e.g., by choosing a specific subgame). Each task is implemented as a subgame and covers a specific problem type in the domain of proportional reasoning. The tasks are directly related to proportional reasoning (e.g., mixing two drinks to make a cocktail according to a particular ratio directly involves proportional reasoning skills). In addition, mental operations with respect to proportional reasoning are connected with the game mechanics (e.g., in order to get the correct amount of bottles in the refrigerator the player has to drag the correct number of bottles in the refrigerator). Each task/subgame can be played on four levels, ranging from easy to difficult. Players first have to finish the three subgames in one level before they can proceed to the next level. Each task (on each level) consists of four assignments. The structure of these assignments is the same, but the numbers vary. For example, in one assignment the student is asked to refill a refrigerator in such a way that the ratio between cola and fanta is 6–12. In the next assignment, this ratio can be 16 cola and 4 fanta, etc. (for a more extensive description, see Vandercruysse et al., 2014).

In the *control* condition, all assignments were presented in an identical way and all information required to perform the assignment was available. See Vandercruysse et al. (this volume) for a description of the control condition.

In the *curiosity* condition, two types of curiosity-triggering events (respectively curiosity type 1 and type 2) were implemented in the Refrigerator and the Blender subgames. The main reason to introduce several types is to have variation in curiosity-triggering events. The Jugs subgame did not use curiosity-triggering events because each assignment in the subgame comprised only two jugs which made the implementation of these events less meaningful. As mentioned before, we define curiosity-triggering events as stimuli that present incongruent information which induce curiosity. The operationalization of curiosity type 1 is as follows:

1. The manager character appears and tells that something strange has happened. He does not exactly know what has happened, but he is sure that the current problem (the assignment) can be solved. In this stage, an expectation is created consisting of the assurance that the problem can be solved.
2. When the character has disappeared, the students cannot see bottles or crates with a caption indicating their numerical value, but only large crates (Refrigerator subgame) or shopping bags (Blender subgame) with a large question mark (see Fig. 1a). The students already may have a hypothesis or idea how to solve the problem, but the opportunities in the game environment (the large crates, shopping bags) are incongruent with what was told them before. Consequently, the perceptual information is incongruent with the verbal information provided by the manager.

They have to explore the contents in the crates and bags and decide how they can solve the problem the best. For example, the blackboard in Fig. 1b makes clear that 4 bottles have to be moved into the refrigerator. The learner can hover the crates/shopping bags and reveal their content. The left crate in Fig. 1b contains three smaller packages with 2, 4, and 6 bottles. By exploring the different crates/bags, the learner can decide which crate/shopping bag contains the packages that can best be used to solve the problem. With a mouse click the large crates/shopping bags are unpacked and the smaller packages become available (Fig. 1c).

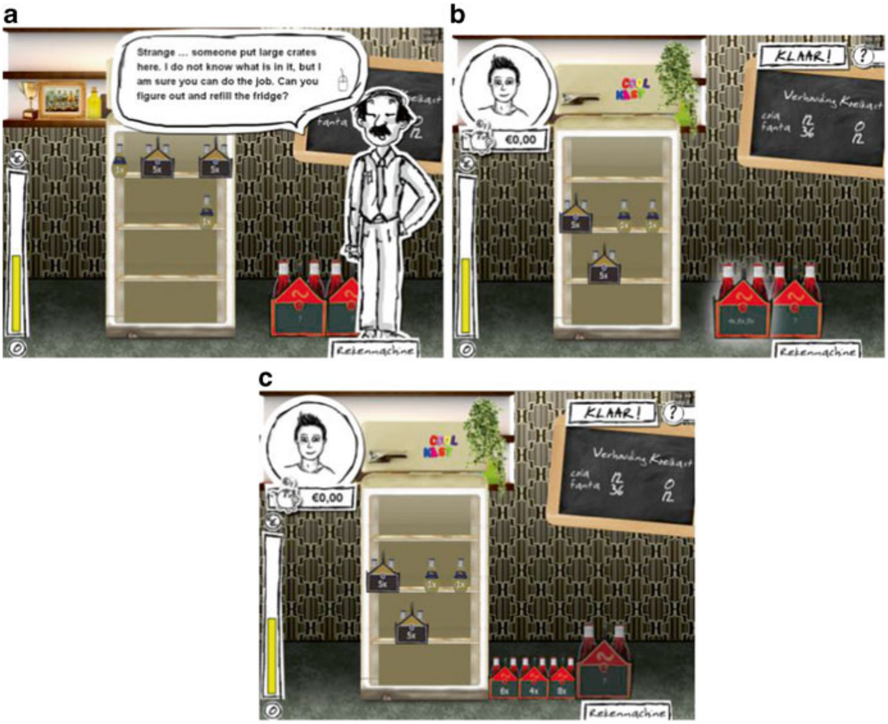

Fig. 1 The implementation of curiosity type 1. (**a**) (*Upper left*) depicts the initial situation. (**b**) (*Upper right*) shows the content when hovering over the crate with the mouse. (**c**) (*Under left*) shows the situation when the crate is unpacked

The operationalization of curiosity type 2 is as follows:

1. The manager character also appears and tells that something has happened but that the problem can still be solved (creating an expectation).
2. The game environment shows a series of crates (Refrigerator subgame) or bottles (Blender game). The first two crates/bottles have a caption with the amount that they represent; the other crates/bottles have a question mark (see Fig. 2a). Again the opportunities in the game environment (crates/bottles with a question mark) are incongruent with what was told them before. So again there is an incongruency between perceptual and verbal information.

The learner can hover the crates/bottles and reveal their content (Fig. 2b). By exploring the content, learners can discover and decide which option best fits the solution of the problem. A crate/bottle can be activated with a mouse click and then be moved to the refrigerator or blender used (Fig. 2c). The crates/bottles always represent an arithmetic relationship (e.g., 12, 18, 24, 30, 36, 42, 48).

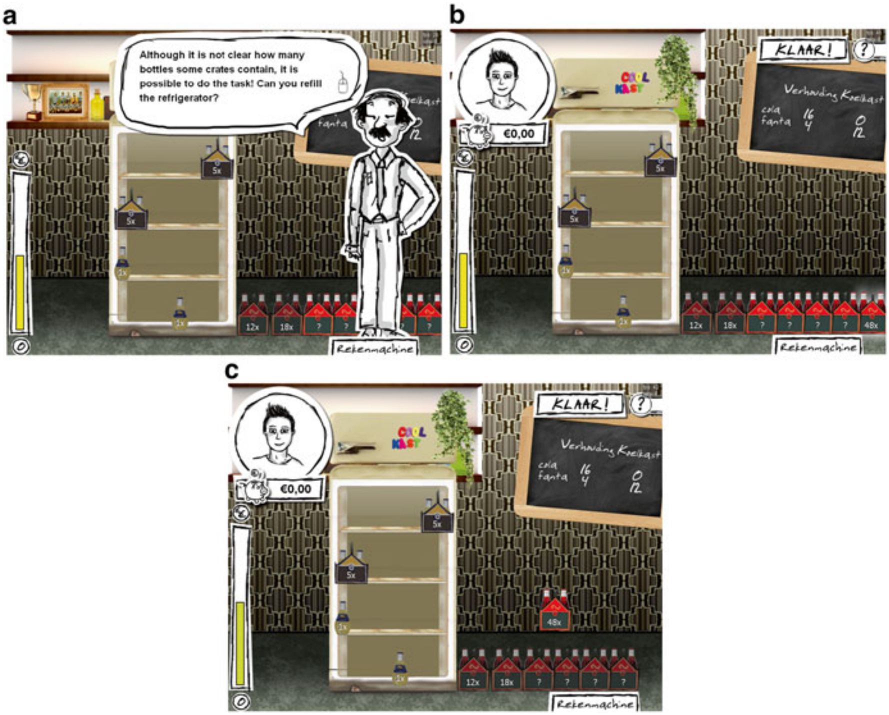

Fig. 2 The implementation of curiosity type 2. (**a**) (*Upper left*) depicts the initial situation. (**b**) (*Upper right*) shows the content when hovering over the crate with the mouse. (**c**) (*Under left*) shows the situation when the crate is selected and dragged to the refrigerator

In each level, two curiosity type 1, one curiosity type 2, and one normal assignment were presented in a random order.

Tests. Proportional reasoning skill was measured with a test consisting of 12 open questions: four questions for each problem type. The questions were comparable with the assignments in the game. An example (missing value) is:

"For a banana milkshake you have to use 28 bananas and 48 units of ice. How many units of ice do you need if you are going to use 56 bananas and you want to remain the same proportion?"

There were two versions of the test. The structure of these versions was the same, but the numbers were different. The comparability of both versions was tested in pilot study.

Motivation was measured with the enjoyment subscale (7 items) from the Intrinsic Motivation Inventory (Ryan, 1982) with a 7-point Likert scale (ranging from "strongly disagree" to "strongly agree"). All items were translated into Dutch and tested in a pilot study (reliability Cronbach's alpha = .74).

Procedure

The experiment was run on the computers of the schools. The experiment took 150 min divided into three sessions of 50 min. In the first session, the experiment was introduced and the pretest was administered (40 min). When participants had finished the pretest, they could do their homework. The second and third sessions, a week later, were two successive lessons with a break of 10 min. In the second session, the participants played the game (40 min) and filled in the motivation questionnaire (10 min). At the beginning of the session, the participants were seated at a designated computer and received a login code. All actions of the players during playing the game were logged. After the break, the posttest was administered in the third session (40 min). One version was used in the pretest, the other version in the posttest.

Scoring

Skill test. Each answer of the pretest and posttest was coded as 0 (wrong answer or no answer) or 1 (correct answer). For the analysis, we focused on the performance on the three problem types (4 questions each) and on the overall performance (12 questions).

Motivation questionnaire. For each participant, a mean score was calculated.

Game performance. Due to technological problems during logging, the data of six participants was removed from the dataset. Two variables were calculated for each participant:

1. The total time they spent in a subgame to perform the assignments.
2. The number of assignments they correctly solved in a subgame.

Results and Conclusion

For all statistical tests, a significance level of .05 was applied. Effect sizes will be expressed in Cohen's *d*. Table 1 shows the results for each condition on proportional reasoning skill and motivation.

A paired-samples *T*-test on the pretest and posttest scores confirms hypothesis 1 arguing that playing the game improves learning (overall $t(66) = 3.31$, $p = .002$, $d = .44$; missing value problems: $t(66) = 2.30$, $p = .025$, $d = .32$; comparison problems: $t(66) = .16$, $p > .05$; transformation: $t(66) = 4.83$, $p = .000$, $d = .28$).

Table 2 provides an overview of both game performance variables for each subgame.

Table 1 Mean scores and standard deviations on the dependent variables for both conditions in Experiment 1

Experiment 1				
	Control		Curiosity	
	Pretest	Posttest	Pretest	Posttest
Proportional reasoning overall [0–12]	3.77 (3.25)	5.00 (3.62)	4.90 (3.07)	6.53 (3.30)
Missing value (Refrigerator) [0–4]	1.42 (1.26)	2.12 (1.56)	2.07 (1.31)	2.59 (1.37)
Comparison (Jugs) [0–4]	1.71 (1.37)	1.61 (1.25)	2.07 (1.26)	2.26 (1.26)
Transformation (Blender) [0–4]	.64 (1.35)	1.27 (1.46)	.77 (1.19)	1.68 (1.51)
Motivation [1–7]	n.a.	3.54 (.38)	n.a.	3.55 (.37)

Note: numbers between [ ] indicate the range of possible scores

Table 2 Mean scores and standard deviations on the game performance variables in Experiment 1

Experiment 1				
	Control		Curiosity	
	M	SD	M	SD
Missing value (Refrigerator)				
Correct assignments	5.67	3.50	6.81	3.52
Time on task (s)	641	355	596	245
Comparison (Jugs)				
Correct assignments	4.42	2.55	5.07	2.94
Time on task (s)	135	102	140	87
Transformation (Blender)				
Correct assignments	5.33	3.60	6.28	3.59
Time on task (s)	651	252	818	417

To test the hypothesis that game performance (correct assignments and time on task) predicts posttest performance, we used a hierarchical regression with two blocks. The first block consisted of the pretest score. In the second block, correct assignments and time on task were entered stepwise. By using two blocks, the effect of the pretest score on the posttest can be isolated. From Table 3 can be concluded that hypothesis 2 is partly confirmed: when the variance caused by the pretest score and both game performance variables is accounted for, only the number of correct assignments and pretest score are predictive for posttest performance. In the control condition, number of correct assignments explains 25 % of the variance extra, in the curiosity condition 20 %.

At face value, it seems that hypothesis 3 can be (partly) confirmed because the curiosity condition benefits more from playing the game than the control condition (see Table 1). However, a series of ANCOVAs (condition as fixed factor and pretest score as covariate) reveals no statistical differences: proportional reasoning all: $F(1, 60) < 1$; missing value: $F(1, 60) = 2.57$, $p = .11$; comparison: $F(1, 60) = 1.5$, $p = .22$ and transformation: $F(1, 60) = 1.70$, $p = .20$. Since motivation was only measured after playing the game, a T-test was conducted ($t(65) = -.18$, $p = .86$), indicating no difference in reported motivation.

Table 3 Hierarchical regression on game performance variables in Experiment 1

Experiment 1			
	B	*SE B*	*β*
Control			
Step 1: Constant	2.64	.92	
Pretest score	.56	.19	.51*
Step 2: Constant	−.25	1.10	
Pretest score	.39	.16	.35**
Correct assignments	.22	.06	.53*
Curiosity			
Step 1: Constant	3.28	.93	
Pretest score	.60	.16	.58*
Step 2: Constant	1.03	1.04	
Pretest score	.34	.16	.33**
Correct assignments	.19	.06	.51*

Note: * $p < .005$, ** $p < .05$
Control: $R^2 = .26$ for step 1, $\Delta R^2 = .25$ for step 2
Curiosity: $R^2 = .33$ for step 1, $\Delta R^2 = .20$ for step 2

Discussion

Although playing the game yielded learning (hypothesis 1) and game performance was positively related to posttest performance (hypothesis 2), we did not find a beneficial effect of curiosity-triggering events in posttest performance nor in motivation (hypothesis 3).

Interviews conducted after the experiment revealed that some students did not immediately understand the intention behind the curiosity-triggering events. This was especially true for the curiosity 1 type assignments. For instance, students raised here their hand and told that they could not solve the problem because there were no bottles with a number (e.g., a bottle with a numerical value) that they could use to perform the assignment. We had expected that this would trigger them to find ways in the game to solve the problem, but they did not make this connection. In the first level of the game, we also observed that some students did not know what to do in the game (also in the control condition). Despite a tutorial, it took these students time to understand what they had to do in an assignment and the actions that were needed to perform the assignment. The log files indicated that most students neglected the tutorial. Based on these findings, we adapted the game.

Experiment 2

In this experiment, we adapted the GBL environment. The most important rationale to conduct Experiment 2 is that the curiosity implementation in Experiment 1 was not clear for all students. Some of them could not make a connection between what was told by the NPC character at the beginning of the curiosity event and what they could do in the game world. In Experiment 2, we have made this link more prominent by introducing a dialogue which refers to the situation in the game world. In addition, the curiosity type 2 events were replaced by curiosity type 1 events (see the "Method" section for more detailed information). Finally, the tutorial was redesigned in order to support students during game play. The same hypotheses were tested.

Method

Participants and Design

Again a pretest–posttest design was used. Students from two small classes were randomly assigned to the control ($N = 11$) and curiosity ($N = 14$) condition.

Materials

Domain. The same domain was used as in Experiment 1.

Game environment. The same game was used as in Experiment 1 with major adaptions. First, the passive tutorial was replaced by an interactive learn-while-you play tutorial which is more in line with contemporary games. Second, the graphics were redesigned in order to limit the size of the game. The game used in Experiment 1 was quite large (40 MB) and sometimes caused long download times. Figure 3 gives an impression of the new visual design (size of the game is 1.2 MB).

The *control* condition was exactly the same as in Experiment 1. The *curiosity* condition was different from the curiosity condition in Experiment 1 on two points:

First, the expectation is now created by a short dialogue between the manager and the aunt character. For example, "Manager: What a strange situation! Yet I know for sure that it is possible to perform the task!" <aunt appears> "Aunt: But how then? …. I only see crates with question marks. Where are the bottles?" It puts more emphasis on the fact that something strange had occurred but that the assignment is still solvable.

Secondly, experts criticized that the arithmetical relationship in the curiosity type 2 events was in fact an additional instructional aid which could confound with the curiosity intervention. Both the curiosity intervention and the (arithmetic) instructional aid can potentially explain a positive effect of the curiosity type 2 events. Therefore, the curiosity type 2 events were replaced by curiosity type 1 events.

Tests. The pretest and posttest were the same as used in Experiment 1. Due to time constraints, the motivation questionnaire was omitted.

Fig. 3 The new design of the GBL

Procedure

The procedure was the same as used in Experiment 1 with the difference that the posttest was not administered directly after playing the game, but in a third session that took place 3 days later.

Results and Conclusion

The same analyses as in Experiment 1 were conducted. Two participants (one from each condition) were identified as outlier and therefore removed from the dataset.

Table 4 shows the results for each condition on proportional reasoning skill.

Hypothesis 1 is rejected because the paired-samples *T*-test reveals no learning effect from playing the game (overall $t(22)=1.01$, $p>.05$). Also tests on subgame level did not reveal differences (missing value problems: $t(22)=.92$, $p>.05$; comparison problems: $t(22)=1.15$, $p>.05$; transformation: $t(22)=1.35$, $p>.05$).

Table 5 gives an overview on game performance.

The hypothesis that game performance (correct assignments and time on task) predicts posttest performance (H2) can be partly confirmed: when the variance caused by the pretest score and both game performance variables is accounted for, only the number of correct assignments and pretest score are predictive for posttest performance. The number of correct assignments explains 46 % and 62 % of the variance extra in the control and curiosity condition, respectively (see Table 6).

Table 4 Mean scores and standard deviations on the dependent variables for both conditions in Experiment 2

Experiment 2				
	Control		Curiosity	
	Pretest	Posttest	Pretest	Posttest
Proportional reasoning overall [0–12]	2.44 (1.01)	2.33 (1.58)	3.71 (1.98)	3.00 (2.38)
Missing value (Refrigerator) [0–4]	.55 (1.01)	.78 (.83)	1.14 (1.34)	1.29 (1.11)
Comparison (Jugs) [0–4]	1.89 (1.17)	1.55 (1.13)	2.28 (.95)	1.42 (1.27)
Transformation (Blender) [0–4]	0 (0)	0 (0)	.25 (.71)	.25 (.70)

Note: numbers between [ ] indicate the range of possible scores

Table 5 Mean scores and standard deviations on the game performance-dependent variables in Experiment 2

Experiment 2				
	Control		Curiosity	
	M	SD	M	SD
Missing value (Refrigerator)				
Correct assignments	4.93	2.66	4.09	2.26
Time on task (s)	632	323	602	238
Comparison (Jugs)				
Correct assignments	4.50	3.23	3.00	1.84
Time on task (s)	168	77	122	100
Transformation (Blender)				
Correct assignments	3.93	2.84	3.09	2.47
Time on task (s)	672	331	444	343

Table 6 Hierarchical regression on game performance variables in Experiment 2

Experiment 2			
	B	*SE B*	*β*
Control			
Step 1: Constant	2.07	.93	
Pretest score	.11	.26	.15
Step 2: Constant	.37	.99	
Pretest score	.12	.21	.16
Correct assignments	.11	.05	.68*
Curiosity			
Step 1: Constant	1.36	1.93	
Pretest score	.30	.40	.29
Step 2: Constant	−1.35	1.48	
Pretest score	−.19	.30	−.19
Correct assignments	.41	.13	.92*

Note: * $p < .05$
Control: $R^2 = .02$ for step 1, $\Delta R^2 = .46$ for step 2
Curiosity: $R^2 = .08$ for step 1, $\Delta R^2 = .62$ for step 2

Hypothesis 3 indicating that the curiosity condition will yield higher learning gains than the control condition must be rejected. Neither on the overall game nor on the different problem types (subgames) the ANCOVAs revealed differences between both conditions (all $F(22) < 1$, $p > .05$).

Despite the modifications in the game environment, we found no learning effect after playing the game (hypothesis 1). The regression analysis on the other hand provides some evidence that game play improves proportional reasoning skills: 46 % and 62 % of the variance in the posttest score (respectively, the control and curiosity condition) can be attributed to the number of correct assignments. This implies that better performance on the posttest can be explained by a higher number of correct assignments which indicates that effective game play improves learning (hypothesis 2). We did not find evidence that curiosity-triggering events improved learning more than a game without these events (hypothesis 3).

The absence of a learning effect can be the result of a lack of motivation during the administration of the posttest. The comparison of the pretest and posttest scores provides some support for this assertion. We found that three participants with rather high scores on the pretest (8, 8, and 7 where the maximum score could be 12) scored very low on the posttest (respectively 2, 0, and 3). It is possible that the game has such a strong negative learning effect, but this is not in line with the results of the control condition in Experiment 1 which was comparable with the control condition in Experiment 2. In Experiment 1, the participants scored significantly higher on the posttest then on the pretest. In our view, it is more likely that these differences arise from low motivation during the administration of the posttest in Experiment 2.

Another explanation may be that the posttest was administered 3 days after the game (in Experiment 1, this was immediately after the game), but there is some evidence that learning effects of GBL environments do not decrease with delayed testing, at least not after 3 days (Wouters et al., 2013).

General Discussion

The goal of the experiments was to investigate whether the use of a GBL environment for proportional reasoning enhances learning. In addition, we examined whether our implementation of curiosity in the GBL environment would further increase the learning effect. Our operationalization of curiosity was based on Loewenstein's information-gap theory (1994). It views curiosity as a reference-point phenomenon, with the reference point being the information that the player wants to know. We concur with Jirout and Klahr (2012) that the information-gap theory combines elements from Gestalt psychology, Social psychology, and behavioral decision theory.

The operationalization involved two phases. First, the student was told that a strange situation had occurred but that the current problem could still be solved. In this way, we created an expectation in the student. Second, the student was confronted with a game environment in which it was not immediately clear how this problem could be solved. Taken together, we regard this as a cognitive conflict namely the conflict between the expectations of the learner and the affordances in the learning environment. Our assumption was that students had to explore the game environment and find the objects (crates/bottles) that would enable them to implement the solution that they had conceived.

In Experiment 1, we found that playing the game had a learning effect. In Experiment 2, game play did not yield learning, though in both studies performance on the game assignments contributed strongly to off-line posttest performance (see the results of the regression analyses). In both experiments, we failed to find a beneficial effect of the curiosity-triggering events. Based on Loewenstein's (1994) and Berlyne's (1960) ideas, we hoped that these situational determinants would induce curiosity. The game environment however had a strong repetitive character which made it perhaps difficult to maintain a curiosity effect. Our implementation of curiosity depended on an incongruity between what players were told and what they saw. Some remarks can be made regarding this implementation. Firstly, can an incongruity that is materialized in two different modalities (verbal and visual) evoke the intended cognitive conflict? It was difficult for some students to make a connection between what was told in the verbal channel and what was shown in the visual channel. This may also explain the confusion that some students experienced when they were confronted with the curiosity-triggering events. It is worth to investigate the impact of the curiosity-triggering events when they occur in only one modality. For example, by focusing on an incongruity within the verbal mode. An additional character can be introduced in the game who challenges the actions of the student or who provides information that is incongruent with the information that is already mentioned. Malone discerned cognitive and perceptual curiosity (Malone, 1981). An interesting addition to cognitive curiosity as we included so far is to implement also some form of pure perceptual curiosity, for example, by hiding bottles in the room in such a way that students have to "collect" the bottles that they want to use.

Secondly, we also don't know if players experienced a cognitive conflict or that they were just confused. For this reason, an obvious next step might be to understand what players think or experience when the curiosity-triggering events occur. Interesting methods in this respect are the use of think aloud protocols and/or eye-tracking.

Thirdly, our interpretation of information gap neglected cognitive conflict as a strategy to promote conceptual change. According to the theory, this can be achieved by presenting new information that challenges preexisting knowledge of the learner.

Whether a cognitive conflict can induce conceptual change depends on various (meta)cognitive factors. The learner must have sufficient prior knowledge in order to identify the gap with the incongruent information. This means that the gap should not be too big and that the learner must feel confident that the knowledge conflict can be reconciled (see Limón, 2001; Merenluoto & Lehtinen, 2004). In GBL environments, such as a math game, this requires a student model with reliable information about a student's prior knowledge and self-confidence. In the current generation of GBL environments, this is still difficult to implement. An application of cognitive conflict to promote conceptual change that we currently investigate starts with a set of proportional reasoning problems that are designed in such a way that one strategy is the most obvious to solve the problem. We assume that this strategy will become prior knowledge because students have to apply it repeatedly. Then, an event in the game changes the characteristics of the problem that the student is working on. The strategy that was used in the preceding problems is now less appropriate so that the student will have to devise a different strategy, which all together could enhance learning.

Some clear limitations adhere to the studies reported here. First, the number of participants in Experiment 2 is very low so the results should be interpreted with care. In addition, we were not able to administer a motivation questionnaire. Second, the comparison of both studies is complicated because in Experiment 1 not only another curiosity-triggering event is used but also because of the changes in the game environment.

Besides the proposals for future research that we have already mentioned, there is another interesting point for further research. Our curiosity manipulation is an example of an added-value comparison in which the effect of a specific game characteristic is investigated (Mayer, 2011; Wouters & Van Oostendorp, 2013). It is possible that the control condition is already effective and that additional curiosity manipulation has no added-value. For future research, we propose research which combines a curiosity and a control condition with a non-game condition. The comparison of the latter two is a media comparison which may give an understanding of the efficacy of the game control condition (see also Girard, Ecalle, & Magnan, 2013).

Acknowledgment This research is funded by the Netherlands Organization for Scientific Research (project Number No. 411-00-003).

References

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York, NY: WH Freeman.

Berlyne, D. E. (1960). *Conflict, arousal and curiosity*. New York, NY: McGraw-Hill.

Cameron, B., & Dwyer, F. (2005). The effect of online gaming, cognition and feedback type in facilitating delayed achievement of different learning objectives. *Journal of Interactive Learning Research, 16*, 243–258.

de Castell, S., & Jenson, J. (2003). Serious play. *Journal of Curriculum Studies, 35*, 649–665.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668.

Dickey, M. D. (2011). Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment. *British Journal of Educational Technology, 42*, 456–469.

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning, 29*(3), 207–219.

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology, 85*, 136–151.

Graesser, A. C., & Olde, B. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology, 95*, 524–536.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*(2), 125–160.

Leemkuil, H., & de Jong, T. (2011). Instructional support in games. In S. Tobias & D. Fletcher (Eds.), *Computer games and instruction* (pp. 353–369). Charlotte, NC: Information Age.

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*(4), 357–380.

Litman, J. A. (2005). Curiosity and the pleasures of learning: Wanting and linking new information. *Cognition and Emotion, 19*, 793–814.

Litman, J. A., Hutchins, T. L., & Russon, R. K. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behavior. *Cognition and Emotion, 19*, 559–582.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*, 75–98.

Malone, T. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 4*, 333–369.

Mayer, R. E. (2011). Multimedia learning and games. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 281–305). Greenwich, CT: Information Age.

Merenluoto, K., & Lehtinen, E. (2004). Number concept and conceptual change: Towards a systemic model of the processes of change. *Learning and Instruction, 14*(5), 519–534.

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology, 97*, 117–128.

Rick, J., Bejan, A., Roche, C., & Weinberger, A. (2012). Proportion: Learning proportional reasoning together. In A. Ravenscroft, S. Lindstaedt, C. D. Kloos, D. Hernández-Leo (Eds.), *Lecture notes in computer science: Vol. 7563. 21st century learning for 21st century skills* (pp. 513–518). Berlin, Germany: Springer.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450–461.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*, 347–363.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64*, 489–528.

Tobias, S., Fletcher, J. D., Dai, D. Y., & Wind, A. P. (2011). Review of research on computer games. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 127–222). Charlotte, NC: Information Age.

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. *Educational Studies in Mathematics, 16*, 181–204.

Vandercruysse, S., terVrugte, J., de Jong, T., Wouters, P., van Oostendorp, H., & Elen, J. (2015). "Zeldenrust": A mathematical game-based learning environment for vocational students.

Wouters, P., Paas, F., & van Merriënboer, J. J. M. (2008). How to optimize learning from animated models: A review of guidelines base on cognitive load. *Review of Educational Research, 78*, 645–675.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*, 249–265.

Wouters, P., & Van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education, 60*(1), 412–425.

Wouters, P., van Oostendorp, H., Boonekamp, R., & van der Spek, E. D. (2011). The role of Game Discourse Analysis and curiosity in creating engaging and effective serious games by implementing a back story and foreshadowing. *Interacting with Computers, 23*, 329–336.