Original article

doi: 10.1111/jcal.12104

Journal of Computer Assisted Learning

When a game supports prevocational math education but integrated reflection does not

J. ter Vrugte,* T. de Jong,* P. Wouters,† S. Vandercruysse,‡ J. Elen‡ & H. van Oostendorp†

*Department of Instructional Technology, University of Twente, Enschede, The Netherlands
†Institute of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands
‡Center for Instructional Psychology and Technology, KU Leuven, Kortrijk, Belgium

Abstract The present study addressed the effectiveness of an educational math game for improving proportional reasoning in prevocational education, and examined the added value of support in the form of reflection. The study compared four conditions: the game with reflection prompts, the game with reflection prompts plus procedural information, the game with procedural information only and the game without additional support. It was found that students' proportional reasoning skill improved after playing the game. The game managed to target prevocational students with low prior knowledge, a group that has the potential to understand proportional reasoning but has not yet encountered the right learning situation to live up to their potential. However, it was also found that students need to be computational fluent to profit from the game. Furthermore, no added value of the support was found. The way the support was structured may have been too demanding for most of the students. The fact that the prevocational students (and specifically those with low prior knowledge) improved by playing the game is noteworthy, because the topic of proportional reasoning is demanding for this group of students who often have lower abilities as well as in some cases a high resistance to learning.

Keywords game-based learning, math, prevocational education, reflection.

Introduction

Games seem to offer an ideal circumstance for high-quality learning (Girard, Ecalle, & Magnan, 2013) because they provide students with an interactive decision-making context in which game players are stimulated to analyse a situation and evaluate the effects of their decisions (Kebritchi & Hirumi, 2008). By providing learners with control (Vogel *et al*., 2006), feelings of competency (Ryan, Rigby, & Przybylski, 2006) and situatedness (Habgood & Ainsworth, 2011), games also create engaging environments that stimulate personal motivation (Kebritchi, Hirumi, & Bai, 2010; Wrzesien & Alcañiz Raya, 2010), which is then thought to facilitate learning (Squire, 2005).

The motivational and engaging nature of computer games makes them particularly attractive for educating students who have lower levels of intrinsic motivation (e.g., prevocational students). Prevocational education (in Dutch: VMBO) is a specific secondary school track in the Dutch educational system where students are prepared for intermediate vocational education. It is the least advanced of three tracks that are offered in secondary education in the Netherlands, and it brings together students who vary highly in their cognitive capability and potential. Quite a few prevocational students are dealing with motivational and/or cognitive

Accepted: 07 December2014

Correspondence: Judith ter Vrugte, Instructional Technology, University of Twente, Cubicus 41, Postbus 217, Enschede 7500 AE, The Netherlands. Email: j.tervrugte@utwente.nl

issues. Many prevocational students have struggled with subjects such as mathematics for years, and their teachers often face educational resistance. These students could especially benefit from an alternative instructional method to keep them interested, motivated and engaged.

However, recent overviews of the effects of game-based learning show that although educational games have potential, the use of computer games for education is not always effective in terms of knowledge acquisition (e.g., Li & Tsai, 2013; Kebritchi *et al.*, 2010; O'Neil, Wainess, & Baker, 2005; Vandercruysse, Vandewaetere, & Clarebout, 2012). One overall conclusion is that support is necessary in order to facilitate learning in game-based education (Garris, Ahlers, & Driskell, 2002; Leemkuil & de Jong, 2011; ter Vrugte & de Jong, 2012). The current study discusses games as an educational tool for prevocational students, and specifically focuses on the effects of incorporating support in the form of reflection prompts in an educational math game.

Motivation and learning from games

Motivation is one of the core aspects that make games appealing for education (Papastergiou, 2009). Motivation can be described as the willingness and desire to engage in a task (Garris *et al.*, 2002). It refers to the individual's choice to engage in an activity and the individual's intensity of effort or persistence during the activity (Wolters, 1998). Garris *et al.* (2002) describe the motivated learner as enthusiastic, focused and engaged. Games and motivation seem to coincide, which means that games should offer a viable means of generating motivated learners. Motivational aspects of educational games that have been identified in prior research include enjoyment, task persistence and engagement (Garris *et al.*, 2002; Lepper & Cordova, 1992). Paras and Bizzocchi (2005, p. 1) explain that 'games foster play, which produces a state of flow, which increases motivation, which supports the learning process'. A recent study by Liu, Horton, Olmanson, and Toprac (2011) demonstrated the positive relationship between motivation and learning in a digital learning environment. However, although motivation and engagement support the learning process, computer games that are engaging and motivational are not guaranteed to result in learning gains (Garris *et al.*, 2002).

Support in games

From prior research in the field of open, inquiry-flavoured media environments, it can be concluded that these environments generally need support structures in order to create an effective learning situation (Alfieri, Brooks, Aldrich, & Tenenbaum, 2010; ter Vrugte & de Jong, 2012). Recent research from Erhel and Jamet (2013) shows that educational games can promote learning, provided that they include features that prompt learners to process the educational content actively. These findings are confirmed by the observation that players of educational games often have difficulty with representing, reproducing and generalizing the knowledge they have learned in the game. This demonstrates that the knowledge that students gain from gameplay is often more intuitive and implicit rather than explicit. This lack of explication can be partially attributed to game flow (Johnson & Mayer, 2010; Leemkuil & de Jong, 2011; Paras & Bizzocchi, 2005; Sweetser & Wyeth, 2005). For actual in-depth learning to take place, we need the players to be conscious of the educational material and how to work with it. Game flow inhibits players from thinking about this explicitly during game play, which implies that support that encourages reflection (as a way to stimulate the development of explicit rather than implicit knowledge) needs to be built into the game to facilitate learning (Ke, 2008; Wouters, Paas, & van Merriënboer, 2008).

Reflection fosters meaningful understanding and therefore is often thought of as an essential element of the learning process (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005; Ke, 2008). Reflection creates awareness of processes that are normally experienced as self-evident. It enables learners to become mindful of the way they handle problems and helps them to critically evaluate the effects of decisions they have made. In addition, it encourages them to integrate newly learned information with prior knowledge, which makes for stronger knowledge structures with increased accessibility (Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

Ke (2008, p. 7) identified reflection as 'a major knowledge construction format for game based learning', but also concluded that, although it is essential for learning, reflection is not a standard element in educational computer games. Johnson and Mayer (2010) found that adding reflection components to an educational computer game furthers knowledge acquisition.

Incorporation of prompts that make players aware of the educational material can turn the players into learners, but is also likely to disturb the game flow and thus interfere with players' engagement and motivation (Johnson & Mayer, 2010; Sweetser & Wyeth, 2005). Loss of engagement and motivation can, in turn, disrupt learning effects. The way reflection is initiated is therefore important.

There are many different ways to initiate reflection. One of the most salient differences between the various forms of reflection prompts is that between open and directed reflections. Open reflection means that a student is simply prompted to reflect. An open prompt can take the form of a direct question or an action that requires reflection. Directed reflection is when students are not only prompted to reflect, but are additionally guided or assisted in completing the reflection (Davis, 2003). Directive prompts can consist of a series of questions to scaffold the reflective process or to direct attention to specific areas of learning. Both open and directed reflections have advantages and disadvantages. For example, Berthold, Eysink, and Renkl (2009) concluded that learners are not always capable of responding appropriately to open reflection questions. On the other hand, it is also likely that directed reflection may restrict learners' options for reflection and can thus limit their opportunities to learn (Chi, 2000).

Johnson and Mayer (2010) compared the effects of providing open, directed and no reflection prompts within a computer game environment. They found that the directed condition yielded significantly better results (students showed more progress on a domain knowledge test) than the open and no reflection conditions, and that there were no differences between the open and no reflection conditions. They offered several explanations for these results: the open reflection could have been too difficult or too disruptive to the game flow, or it might not have been the reflection in the directed condition that caused the effects, but the information that was provided through the multiple-choice answers in the directive prompts. Expanding on these findings of Johnson and Mayer (2010), and taking into account the findings of Berthold *et al.* (2009) about the fact that open reflection is often too difficult for learners, the current study implemented directed reflection prompts to support prevocational students when learning in the context of an educational computer game.

Current study

The current study evaluates prevocational students' learning with an educational math game. Math was chosen because it is a fundamental skill for future school achievement, and prevocational students' math skills are often inadequate (CvE, Dutch Board of Examinations, 2014). More specifically, the math subdomain of proportional reasoning was selected. Besides the fact that recent reports of Cito, Dutch Central Institute for Test Development (2011) show a severe deficiency of prevocational students in proportional reasoning skill, the selection of proportional reasoning was driven by the following reasons: First, proportional reasoning is a fundamental skill for future math achievement and mathematical understanding (Rick, Bejan, Roche, & Weinberger, 2012). Second, proportional reasoning is a well-defined domain. Third, traditional instructional methods for proportional reasoning are often ineffective (Rick *et al.*, 2012), and therefore students regularly lack proportional reasoning skills (Lawton, 1993; Tourniaire & Pulos, 1985). Difficulties with proportional reasoning seem to emerge from students' possession of fragile domain-specific concepts. Proportional reasoning problems vary in structure and this can create difficulty in applying these already fragile concepts (Lawton, 1993; Tourniaire & Pulos, 1985). Instruction of proportional reasoning is likely to benefit from game-based learning because, in addition to the traditional word problems, a game can provide students with a variety of motivational and vivid contexts and opportunities to interact with the material. The active, multimodal nature of the environment can help students develop a more solid and concrete understanding of the normally abstract proportions and ratios that make up the core of proportional problems.

The current study targets a specific group of secondary students: prevocational students. This group includes a significant number of at-risk students with a history of poor learning. These students often encounter numerous unsuccessful instructional interventions and have grown resistant to the traditional educational material. Educational games can create an alternative approach that might motivate such learners to reengage with the educational material. In addition, the interactive multimodal features may provide them with new insights they would have missed with more traditional methods of instruction.

The current study investigated whether prevocational students could benefit from an educational game and whether in-game reflection could foster their learning. It was expected that students in this study would not possess the metacognitive skill and content knowledge for successful open reflection (Berthold *et al*., 2009; Johnson & Mayer, 2010). For this reason, we used directed reflection prompts that focused students' reflection towards specific aspects of domain knowledge. In addition, we took into account the possibility that these students could still experience difficulty in coming to an explicit realization of the knowledge that resulted from their reflection. Therefore, we added procedural information that was designed to help students structure what they learned. In summary, our study explored the effects of providing directed reflection prompts (and procedural information) in an educational math game by comparing groups that received either reflection prompts, procedural information, a combination or no support. The following hypotheses were developed based on the research literature:

1. Playing the game will help students learn about proportional reasoning.
2. Students who receive reflection prompts during the game will outperform students who do not receive these prompts.
3. Students who receive reflection prompts combined with procedural information will outperform other students.
4. The effects of the game and its support are influenced by students' computational skills and prior knowledge.

We also investigated whether the addition of support influenced students' perceptions of the game, in particular whether it affected the way students perceived the usefulness and playfulness.

Method

Participants and design

The study was conducted in two schools for prevocational education in the Netherlands. The sample involved 145 students, 78 boys and 67 girls, aged 13.3–17.5 years old [with mean (M) = 14.88 and standard deviation (SD) = 0.79]. The students participated in the second (59 students) or third year (86 students) of the program of study. All students possessed basic computer skills, which are part of the national Dutch curriculum. Students were familiar with educational software, but new to the game that was used in the current study.

The study utilized a 2 × 2 factorial design. The four conditions involved were identical in terms of embedded learning objectives (proportional reasoning) and learning material (game environment), and differed on two variables: the presence or absence of reflection prompts and the presence or absence of procedural information.

Materials

Domain

The domain involved in this study is proportional reasoning. Three types of proportional problems were identified: comparison problems, missing value problems and transformation problems (e.g., Harel & Behr, 1989; Kaput & West, 1994; Tourniaire & Pulos, 1985). *Comparison problems* always involve two ratios. Students must determine the relationship between two ratios. Possible answers to these problems are that the first ratio is 'more than', 'less than' or 'equal to' the second ratio. Comparison problems can be divided into three levels of difficulty. The *first* level (the easiest level) includes problems that can be solved directly by qualitative reasoning. The answer to these problems can be achieved by reasoning because either the values of the antecedents or consequents in both ratios are equal (e.g., 1:4 vs. 3:4), or the comparison involves ratios that are obviously quite small and quite large (e.g., 100:31 vs. 42:100). The *second* level includes problems that can be solved by estimation. In this case, the answer can be estimated because the internal or external terms of the proportion show an easy multiplication (e.g., 2:4 vs. 4:6) or the internal or external terms of the proportion match a simple reference point (e.g., 1/2, 1/3, 1/4, or 1/10). The *third*, and hardest, level must be solved using full calculation. The answer cannot be determined directly by reasoning or estimation, but must be computed (e.g., 14:63 vs. 18:81). The other two problem types are missing value and transformation problems.

Missing value problems concern a proportion in which one value is missing. Students must calculate the missing value, assuming that both ratios are in proportion (e.g., 3:6 = ?:12). *Transformation problems* concern two ratios that are not (yet) in proportion (e.g., 3:6 ≠ 4:12). Students must calculate how much has to be added to one ratio to make both ratios equal. Both missing value and transformation problems can be divided over four levels of difficulty depending on whether the multiplicative relations between the internal and/or external terms of the proportion are integer or not (e.g., Kaput & West, 1994; Tourniaire & Pulos, 1985; van Dooren, de Bock, Evers, & Verschaffel, 2009). The educational game intervention focused on practice and knowledge gains on all three problem types. An example of each type of problem as implemented in the game can be found in Table 1.

Game

The intervention consisted of a newly developed computer game application, 'Zeldenrust', in which students take on the role of a hotel employee. The goal of the game is to gather as much money as possible (to spend on a holiday) and this can be achieved by completing challenges around the hotel. All challenges require efficient and effective use of proportional reasoning. The amount of money earned for completing a challenge increases in relation to the accuracy of the response, and the accuracy of the actions taken, while it decreases with the use of the calculator, and the number of attempts used to solve the problem. The more money students earn, the farther they can travel on their virtual holiday.

The game consists of the following:

- *The game centre*: a central point in the game where students keep track of their score and receive directions.
- *Four levels*: the levels are of progressively increasing difficulty, each level targets a specific level of proportional reasoning, students get to practise all the proportional problem types in every level, the problem types are presented in specific subgames
- *Three subgames*: games that are designed to practise specific types of proportional reasoning problems; in these games, specific parts of the hotel are represented, the parts have dedicated features for performing specific assignments, like a bar area with pitchers and a serving tray.
- *48 challenges*: the challenges represent problems that require proportional reasoning; the student must complete four challenges at every level for every subgame.

Table 1 provides an overview of the subgames, levels and challenges (including the number of attempts that students are allowed to use, to solve a challenge in a subgame). Figure 1 provides an illustration of the game centre and the three subgames.

When the game starts, students see a short animation that introduces them to the storyline and the goal of the game. After this, they can choose an avatar (out of four options) and enter the game centre, where they meet the hotel owners (non-playable characters) and are taken to their virtual room. This room (Figure 1, upper left illustration) is the game centre. From here, subgames can be entered. Students automatically return to this game centre when a subgame is finished. In the game centre, the hotel owners give the students tasks: that is, fill the fridges, mix cocktails and serve drinks. Each task needs to be completed in a subgame. When a student enters a subgame (by clicking one of the paintings on the wall), the owners introduce the challenge that has to be accomplished. In addition, the first level of each subgame starts with a tutorial. After this, the first challenge is introduced. Students can solve the challenges by dragging and dropping the correct number of objects to the correct place. Once they have given their solution, feedback is provided. Feedback depends on the number of attempts students have made at solving the challenge and whether their solution is correct. After one attempt, the feedback states whether the solution is right or wrong. After a second attempt, the feedback states either that the answer is correct or that the answer is less or more than the expected answer (e.g., 'This number is not correct. You have used too many berries'). After a third attempt, the feedback states whether the answer is right or wrong and the game proceeds to the next challenge. After four challenges, students receive the cash they earned and return to their room. Here they can keep track of their holiday destination on a geographical map, or start a new subgame. Every subgame can be opened only once per level. After completion of all three subgames at one level, students get access to the next level. This structure fosters maximum variation (in context and problem type) in combination with progressive difficulty, which promotes the experience of challenge and reduces feelings of frustration.

Table 1. Overview of Level Structure per Subgame

Subgame	Problem type	Challenges per task	Attempts per challenge	Example of problem	Game level 1	Game level 2	Game level 3	Game level 4
Jugs	Comparison	Four	One	‘There are two pitchers of juice on the counter. A customer asks for the sweetest juice mix. Which juice mix will you give to the customer?’ The ratio of water/fruit was presented on the pitchers. The student had to click on the correct pitcher to answer.	Qualitative reasoning	Estimation	Calculation	Mix of levels 1, 2 and 3
Refrigerators	Missing value	Four	Three	‘This is the reception desk refrigerator. This refrigerator always contains 3 bottles of water for every bottle of juice. It already contains 9 bottles of water. Fill the refrigerator so it will contain the right amount of juice’. The given ratio of 3/1 was presented next to the ratio with the missing value 9/?. The student had to answer the question by dragging and dropping the juice bottles into the refrigerator.	Internal terms, and external terms integer	Internal terms integer, and external terms not integer or vice versa	Internal terms, and external terms not integer	Mix of levels 1, 2 and 3
Blender	Transformation	Four	Three	‘A fruit cocktail contains 10 berries for every 100 mL of yoghurt. How many berries should you add to 500 mL of yoghurt if you want to maintain the flavor?’ The given ratio of 10/100 was presented and the student had to answer the question by dragging and dropping the berries into a blender that contained the 500 mL of yoghurt.	Internal terms, and external terms integer	Internal terms integer, and external terms not integer or vice versa	Internal terms, and external terms not integer	Mix of levels 1, 2 and 3

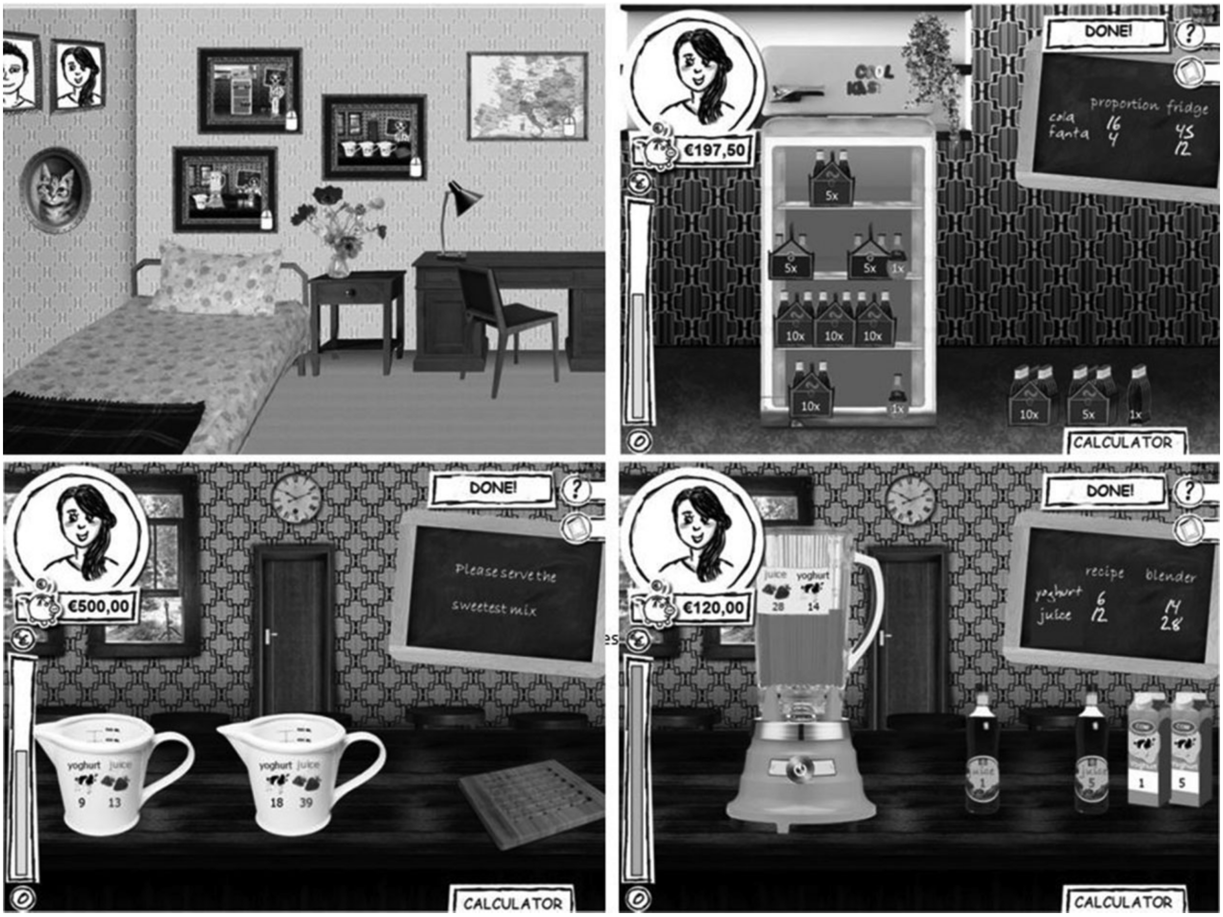
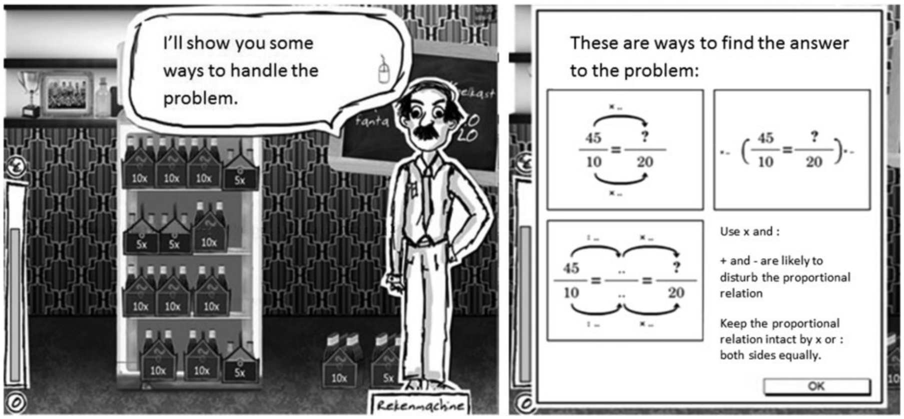


Figure 1 Screencaptures Game Centre Screen (Upper Left) and Subgames

The goal of ‘Zeldenrust’ is to encourage active learning within a game environment. Students have the opportunity to search for and discover information in an interactive environment, to engage in problem solving, to think about concepts presented and to test their understanding of those concepts. Papastergiou (2009) identified a series of elements that can promote student involvement within an instructional gaming environment. In the current game, the following elements were adopted: clear but challenging goals, fantasy linked to the student activity, progressive difficulty elements, and immediate and constructive feedback. In Zeldenrust, goals are introduced in the narrative and intertwined with the gameplay and storyline of the game to assure *clear goals*. Clear goals stimulate engagement and engage players’ self-esteem (Malone, 1981). Because games where the learning content and game content are fully – or intrinsically – integrated are expected to be superior with respect to learning outcomes (Habgood & Ainsworth, 2011), the storyline and the gameplay of Zeldenrust were designed to *integrate* the educational content seamlessly. As advised by Malone (1981), the goal and the theme of the storyline (earning money for a holiday) were tailored so that teenage students could identify with it, and could link the virtual (*fantasy*) world to their daily activity. And finally, to assure *progressive difficulty* and minimize frustration, a level-based structure was incorporated and *feedback* was provided. To promote greater retention and a greater correction of inaccurate strategies, feedback was provided immediately upon response, and was both corrective and constructive (Dihoff, Brosvic, & Epstein, 2003).

To overcome societal issues, the game depicted a gender-neutral and violence-free setting and storyline, and all references to alcohol or other drugs were avoided. Moreover, the following practical conditions



I'll show you some ways to handle the problem.

These are ways to find the answer to the problem:

$$\frac{45}{10}=\frac{?}{20}$$

$$\cdot ..\left(\frac{45}{10}=\frac{?}{20}\right)\cdot ..$$

$$\frac{45}{10}=\frac{..}{..}=\frac{?}{20}$$

Use x and :

+ and - are likely to disturb the proportional relation

Keep the proportional relation intact by x or : both sides equally.

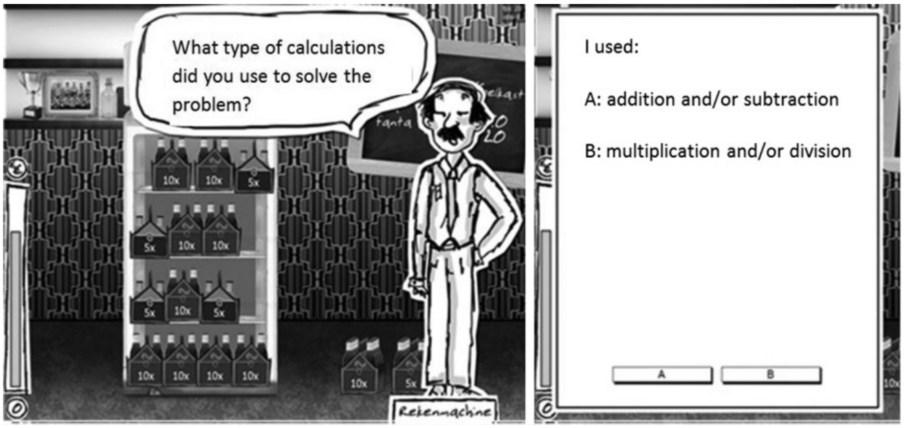
Figure 2 Information Introduction Screen (Left) and Information Screen (Right)

were considered: the available computer hardware at schools, the total time needed to complete the game, and the intuitiveness of the game controls. These practical implications led to some design restrictions: two-dimensional instead of three-dimensional graphics were used, audio fragments were limited, the storyline was kept relatively simple, and all game controls were mouse operated.

Variants of the game for the different conditions

To create the research conditions, two additions to the game were designed: procedural information and reflection prompts. *Procedural information* provided students with possible procedures and the corresponding rules necessary to tackle the problems that were presented in the game (see Figure 2 for an example of procedural information). *Reflection prompts* were presented in the form of multiple-choice questions. These questions directed students' attention to the steps that are the most important when solving a proportional problem: that is, 'what type of calculations did I use?', 'how did I apply these calculations?', 'what would happen if one quantity changed?' (see Figure 3 for an example of a reflection prompt). After answering a multiple-choice reflection question, students received feedback on whether their answer was correct or incorrect.

The reflection prompts and procedural information appeared at predetermined points; students received the support eight times during the game. They received it twice per level: after their first attempt for the second challenge during both the refrigerator and blender subgames. In this way, students could first practise the first challenge for each level, then receive reflection prompts and/or procedural information, and could apply the knowledge they had gained in the following problem. To keep the disruption of the game flow to a



What type of calculations did you use to solve the problem?

I used:

A: addition and/or subtraction

B: multiplication and/or division

Figure 3 Reflection Question Screen (Left) and Answer Screen (Right) minimum, all the information and questions were embedded in the storyline of the game and were presented in an interactive conversation with a non-playable character (i.e., one of the hotel owners).

Test materials

To evaluate computational skills, students completed an arithmetic tempo test, the *TTR* (*Tempo Test Rekenen*). This is a validated test developed in the Netherlands and Flanders which aims to measure students' computational skills in fundamental arithmetic computations, that is, addition, subtraction, multiplication and division (de Vos, 1992). The test consists of the four types of arithmetic computations distributed over five sheets, one sheet for each type of computation and one with all types in mixed order. There were 40 arithmetic problems per sheet, presented in an order of increasing difficulty. The students had 1 min per sheet to solve as many arithmetic problems as possible. The more arithmetic problems the students solved correctly, the better their computational skills.

TTR scores in the current study represent the sum of all the correct answers, with a possible range from 0 to 200. These scores were used to identify whether the students were computational fluent; a score was calculated based on the principle of automation which states that when a student is able to process the calculation and provide the answer within 3 s, the student has an adequate mastery of learned facts and strategies (van de Bosch, Jager, Langstraat, Versteeg, & de Vries, 2009). Applying this principle to the TTR scores meant that students were computational fluent when they were able to provide 100 or more correct answers within the 5 min (300 s) of the test.

Domain knowledge of proportional reasoning was assessed with a *domain knowledge test*. This test was used to measure domain knowledge prior to and after the intervention. Therefore, two parallel versions of this test were developed. Both had the same structure and text, but the numbers in the problems were changed. Both tests were equally difficult and they were administered in a counterbalanced order. Reliability analyses of the test scores show a Cronbach's alpha of 0.78.

The domain knowledge tests consisted of 15 open-ended questions: four questions for each type of proportional reasoning problem (i.e., missing value, comparison and transformation) and three transfer questions. The proportional problems of each type were presented in order of increasing difficulty. Every question presented a proportional problem that was similar in context and structure to the problems posed in the game. The three transfer questions were math problems from adjacent domains, that is, fractions, measurements and geometry.

For each question, the students had to complete a calculation (procedure) and an answer. Both answers and calculations were coded for the missing value and transformation problems. Because of the nature of the comparison (subject to guessing) and transfer problems (no identifiable procedures), only the answers were coded for these. Answers could be coded as correct, incorrect or missing. The score for domain knowledge represented the sum of all the correct answers, with a range from 0 to 12. Scores on transfer represented the sum of all the correct answers on the transfer questions, with a range from 0 to 3. In addition, the number of adequate procedures was identified based on the calculations the students had provided at the missing value and transformation problems. Calculations were coded as adequate procedures when they could be identified as known proportional procedures (i.e., proportional procedures that are taught, or that are known to be effective). Table 2 provides an overview of the different procedures combined with an example per proportional problem type. If the calculation did not fit one of these procedures, it was coded as inadequate. The score on the adequate procedures could range from 0 to 8.

Furthermore, the overall perception of the game was assessed with a *perception questionnaire*. This questionnaire measured students' playfulness and perceived usefulness of the game. The questionnaire consisted of nine 6-point Likert scale items: five items to measure playfulness and four items to measure usefulness. The items that measured playfulness were all based on the 'direct play assessment' subscale from the play experience scale (Pavlas, Jentsch, Salas, Fiore, & Sims, 2012). The items that measured usefulness were all based on the 'usefulness' subscale of the Intrinsic Motivation Inventory (McAuley, Duncan, & Tammen, 1989).

All items represented perceptions of the students in regard to the game and its potential for learning, for example: 'the use of this game is beneficial for me (when I am studying proportional problem solving)' and 'the game felt more like "playing" instead of

Table 2. Scheme for Coding Calculations in the Domain Knowledge Test

Adequate procedure	Example of procedure missing value problem	Example of procedure transformation problem
	Question: *If a recipe for a smoothie prescribes that per 4 strawberries you should add 8 raspberries.* *How many strawberries do you need when you want to use 16 raspberries?*	Question: *The recipe for a smoothie prescribes that per 4 strawberries you should add 8 raspberries. Someone already mixed 6 strawberries with 16 raspberries.* *How many strawberries do you need to add to complete the smoothie?*
Internal rationalization Based on the internal terms of the proportion	$4/8 = ?/16$ 8 became twice as big (*or* 16 is 2 times 8). Therefore, 4 should also be multiplied by 2. 2 times 4 is 8. The answer is 8.	$4/8 \neq 6/16$, *thus* $4/8 = 6 + ?/16$ 8 became twice as big (*or* 16 is 2 times 8). Therefore, 4 should also be multiplied by 2. 2 times 4 is 8. So it should be 8/16 not 6/16. So we need to add two.
External rationalization Based on the external terms of the proportion	$4/8 = ?/16$ 4 is half of 8 (*or* 8 divided by 2 is 4). Therefore, ? should also be half of 16. 16 divided by 2 is 8. The answer is 8.	$4/8 \neq 6/16$, *thus* $4/8 = 6 + ?/16$ 4 is half of 8 (*or* 8 divided by 2 is 4). Therefore, ? should also be half of 16. 16 divided by 2 is 8. So it should be 8/16 not 6/16. So we need to add two.
Simplifying First simplifying the first ratio of the proportion before expanding it	$4/8 = ?/16$ $4/8 = 2/4 = 8/16$ The answer is 8.	$4/8 \neq 6/16$, *thus* $4/8 = 6 + ?/16$ $4/8 = 2/4 = 8/16$ So it should be 8/16 not 6/16. So we need to add two.
Simplifying to one First simplifying the first ratio of the proportion to a 'something-to-one' ratio	$4/8 = ?/16$ How many strawberries do you need for every single raspberry? $4/8 = 0.5/1$ I need half a strawberry for every raspberry. There are 16 raspberries. 16 times 0.5 is 8. The answer is 8.	$4/8 \neq 6/16$, *thus* $4/8 = 6 + ?/16$ How many strawberries do you need for every single raspberry? $4/8 = 0.5/1$ I need half a strawberry for every raspberry. There are 16 raspberries. 16 times 0.5 is 8. So it should be 8/16 not 6/16. So we need to add two.
Correct additive reasoning Calculating one or more equivalent ratios and adding it to the first ratio	$4/8 = ?/16$ 8 plus 8 is 16. 4 plus 4 is 8. The answer is 8.	$4/8 \neq 6/16$, *thus* $4/8 = 6 + ?/16$ 8 plus 8 is 16. 4 plus 4 is 8. So it should be 8/16 not 6/16. So we need to add two.

"studying" '. Reliability analyses of the test scores showed a Cronbach's alpha of 0.68 for the items that measured perceived playfulness, and a Cronbach's alpha of 0.90 for the items that measured perceived usefulness.

The students received this questionnaire after they finished the game and had to indicate the extent to which they agreed with the given perceptions (where higher scores equal greater agreement). Scores in the current study are the sum of the scores of all the items per construct, ranging from 0 to 30 for playfulness, and from 0 to 24 for usefulness.

Procedure

The total time spent on this study was 200 min, spread evenly over four sessions. The first session started with a short introduction. In this introduction, the students were informed about the organization of the upcoming lessons and what was expected from them. After the introduction, the students completed the TTR and the domain knowledge test. Before starting the game in the second session, the students were assigned to the four conditions. Because of large differences in performance between the students, the distribution of the

Table 3. Coding Scheme of In-Game Performance

Variable	Logging	Coding	Range per level	Range per game
Time on task	Time spent on a single challenge	Summation of all the loggings of time	NA	NA
Correct answers	Correct answer on a challenge	Summation of all the correct answers	0–12	0–44
Number of attempts	Attempts needed to complete a challenge	Summation of all the attempts	12–28	48–144

students over the conditions was based on their level of performance on the domain knowledge pretest. The students and the teachers received no information on the different conditions and their content and were not aware of any different groups that were made.

The second session started with a short introduction (approximately 10 min) on how to play the game and on the math problems addressed in the game. The goal was to inform the students and to activate their prior knowledge on proportional reasoning so that they would be able to work independently on the game. Again, expectations were made clear: work individually, no help during the game, keep calm and quiet, and only pay attention to your own screen. Next, the students received codes so they could log in on a version of the game that matched the group they were assigned to. In the third session, the students could resume the game where they left off the previous time. When the students finished the game, the perception questionnaire was administered. In the fourth session, the students completed a parallel version of the domain knowledge test.

Data analyses

A total of 145 students participated in the study. However, because of the duration of the study, and the fact that it was spread across four sessions, a dropout occurred: 2 students failed to attend the pretests session, 9 failed to attend the posttest session, 8 failed to attend both the pre- and posttest sessions, 10 failed to complete the game, and 6 failed to attend both the game and either the pre- or posttest sessions. In total, 35 students did not complete the study (24%). Dropout was evenly spread across conditions. Results of the performance measures are based on the analyses of data from the 110 students who completed all sessions (pretests, game and posttest). For the results of the perception data, eight additional students were left out of analyses because these students failed to complete the perception questionnaire. For the results of the process data, three additional students were left out of the analyses because these three students' loggings during the game were not saved properly.

Several variables were required to answer the research questions. Game-generated loggings were consulted to derive in-game performance measures: number of attempts needed to solve challenges, correctly solved challenges, time on task. Time on task is taken as the total amount of time the students actually spent on all the challenges; time spent navigating around the environment was not recorded. A high score for time on task could have been caused by the student requiring more attempts (and thus more time), performing more calculations while solving a challenge, being slower with his or her calculations, or being distracted during the game. A higher score for time on task could therefore represent a weaker math student and/or a less engaged student. Table 3 provides an overview of all the variables that were derived from the game-generated loggings.

Results

Table 4 summarizes the descriptive statistics for the participants' test scores and game performance per experimental condition. The overall mean of the domain knowledge pretest score was 6.37 (range = 13, *SD* = 3.00) out of a maximum of 15, which was considered to be sufficiently low to assume that the test could be used to register a development in knowledge. Univariate analysis of variance (ANOVA) revealed no differences in computational skills as measured with the TTR, $F(3,\ 106) = 1.33$, $p = 0.269$ (with effect size $\eta_p^2 = 0.036$), and no significant differences in prior domain knowledge, $F(3,\ 106) = 0.12$, $p = 0.949$ (with

Table 4. Summary of Students' Scores by Condition

	Experimental condition							
	Reflection plus procedural information		Reflection		Procedural information		Control	
	(*n* = 28)		(*n* = 29)		(*n* = 28)		(*n* = 25)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Test scores								
Computational skills (TTR)	115.75	18.19	122.55	23.79	127.68	24.85	118.52	27.46
Domain knowledge pretest	5.21	2.51	5.41	2.13	5.21	2.83	5.36	2.71
Domain knowledge posttest	5.96	3.11	6.66	2.61	7.18	2.83	6.52	2.62
Adequate procedures pretest	3.71	2.55	4.41	2.24	3.86	2.46	4.04	2.46
Adequate procedures posttest	4.75	2.61	5.52	2.35	4.46	2.87	5.00	2.52
Transfer pretest	1.28	1.02	0.97	0.91	1.14	.097	1.28	1.02
Transfer posttest	1.24	1.01	1.17	0.93	1.54	.961	1.24	1.01
Perceived usefulness	18.44	5.99	19.46	5.32	17.85	7.39	19.44	8.03
Perceived playfulness	10.96	4.50	14.08	4.35	12.74	4.82	13.32	4.43
In-game performance scores								
Time on task (s)	2350	606.12	2056	580.68	2221	688.21	2335	702.88
Number of correct solutions	31.35	7.55	31.24	8.89	33.82	7.69	35.92	8.50
Number of attempts	62.85	13.14	60.90	13.16	64.48	12.34	65.56	15.68

effect size $\eta_p^2 = 0.003$) between conditions. This proves that even after dropout, conditions are comparable with respect to students' prior knowledge and skills.

The *first hypothesis* was that playing the game would help students to learn about proportional reasoning. A paired sample *t*-test across all participants indicated a significant difference between total pretest and posttest scores, $t(109) = –5.23$, $p < 0.001$, with effect size $d = 0.44$. Further analysis indicated a significant difference between the number of adequate procedures on the pretest and the number of adequate procedures on the posttest, $t(109) = –4.47$, $p < 0.001$, with effect size $d = 0.34$. These results demonstrate that students not only learned to work with proportions implicitly (solve the problem correctly), but also gained explicit knowledge (provide the related procedure).

To evaluate whether playing the game affected students' posttest performance, a stepwise regression analysis was conducted, with computational skills (TTR), pretest domain knowledge score, and game measures (time on task and number of correct solutions) as predictors. All predictors were entered simultaneously. Tests to see if the data met the assumption of collinearity indicated that multicollinearity was not a concern (TTR tolerance = 0.772, $VIF = 1.30$; pretest tolerance = 0.737, $VIF = 1.36$; time-on-task tolerance = 0.833, $VIF = 1.20$; number of correct solutions tolerance = 0.736, $VIF = 1.36$). Statistics on the correlations between the variables that were entered in the regression analyses can be found in Table 5.

The results of the regression indicated that three of the predictors (pretest score, computational skills and number of correct solutions during the game) explained

Table 5. Correlations, Means and Standard Deviations of Regression Variables

Measure	*M*	*SD*	1	2	3	4
1. Domain knowledge posttest	5	2.5				
2. Domain knowledge pretest	7	2.8	0.577**			
3. Computational skills (TTR)	121	23.8	0.462**	0.374**		
4. Number of correct solutions in game	33	8.3	0.486**	0.366**	0.335**	
5. Time on task in game (s)	2234	647.0	–0.103	–0.230*	0.208*	0.173

$^*p < 0.05$, $^{**}p < 0.01$.

52% of the variance in posttest performance, $R^2 = 0.52$, $F(3, 106) = 37.21$, $p < 0.001$. It was found that pretest scores significantly predicted posttest scores, $\beta = 0.426$, $p < 0.001$, as did computational skills, $\beta = 0.247$, $p = 0.002$, and the number of correct solutions in the game, $\beta = 0.237$, $p = 0.003$. Time spent on challenges within the game did not contribute to the prediction of posttest performance, $p = 0.779$. These outcomes concur with the expectation that playing the game fosters knowledge acquisition; not only did the students perform significantly better on the posttest, how well they performed on the posttest is partially predicted by how well they performed during the game (when controlling for the effects of prior knowledge and computational skills).

The *second and third hypotheses* were that reflection during the game would foster learning and that the combination of reflection and procedural information would produce the best results. A mixed-design ANOVA with time (domain knowledge pretest to posttest) as a within-subject factor and reflection and procedural information as between-subject factors revealed a main effect of time, $F(1, 106) = 27.40$, $p < 0.001$, $\eta_p^2 = 0.205$. This effect was not qualified by an interaction between time and reflection, $F(1, 106) = 1.30$, $p = 0.257$, $\eta_p^2 = 0.205$, nor was there an interaction between time and procedural information, $F(1, 106) = 0.28$, $p = 0.600$, $\eta_p^2 = 0.003$. The hypothesized interaction of time, reflection and procedural information was not significant, $F(1, 106) = 3.01$, $p = 0.085$, $\eta_p^2 = 0.028$.

In addition, a mixed-design ANOVA with time (adequate procedures pretest to posttest) as a within-subject factor and reflection and procedural information as between-subject factors revealed a main effect of time, $F(1, 106) = 19.74$, $p = 0.000$, $\eta_p^2 = 0.157$, and no effect of reflection, $F(1, 106) = 0.09$, $p = 0.765$, $\eta_p^2 = 0.001$, or procedural information, $F(1, 106) = 0.01$, $p = 0.920$, $\eta_p^2 = 0.000$, on the number of adequate procedures from pretest to posttest.

Furthermore, whether reflection supported transfer was investigated. Hence, scores on transfer problems were analysed. A mixed-design ANOVA with time (transfer pretest to posttest) as a within-subject factor and reflection and procedural information as between-subject factors revealed no main effect of time, $F(1, 106) = 2.76$, $p = 0.100$, $\eta_p^2 = 0.025$, no interaction between time and reflection, $F(1, 106) = 0.10$, $p = 0.759$, $\eta_p^2 = 0.001$, no interaction between time and procedural information, $F(1, 106) = 0.53$, $p = 0.467$, $\eta_p^2 = 0.005$, and no three-way interaction, $F(1, 106) = 2.84$, $p = 0.095$, $\eta_p^2 = 0.026$. These outcomes, although they support the previous finding that students learned from the game, do not support the hypotheses that reflection prompts, procedural information, or a combination of both, support students' knowledge acquisition during gameplay.

The *fourth hypothesis* stated that students' computational fluency and their prior knowledge would influence the effects of support. To evaluate whether *computational fluency* affected students' ability to learn from the game and their ability to benefit from the support, a mixed-design ANOVA with time (domain knowledge pretest to posttest) as a within-subject factor and reflection, procedural information, and computational fluency as between-subject factors revealed a main effect of time, $F(1, 102) = 7.64$, $p = 0.007$, $\eta_p^2 = 0.070$, and an interaction between computational fluency and time, $F(1, 102) = 4.96$, $p = 0.028$, $\eta_p^2 = 0.046$, with the students who were computational fluent outperforming the students who were not. There was no interaction between time, reflection and computational fluency, $F(1, 102) = 0.39$, $p = 0.535$, $\eta_p^2 = 0.004$, and no interaction between time, procedural information and computational fluency, $F(1, 102) = 0.03$, $p = 0.868$, $\eta_p^2 = 0.000$. All other effects were non-significant.

To evaluate whether *prior knowledge* affected students' ability to learn from the game and their ability to benefit from the support, *s*tudents were grouped as performing either below average or above average based on their domain knowledge pretest scores. Then a mixed-design ANOVA with time (domain knowledge pretest to posttest) as a within-subject factor and reflection, procedural information, and prior knowledge as between-subject factors, revealed a main effect of time, $F(1, 102) = 24.14$, $p = 0.000$, $\eta_p^2 = 0.191$, and an interaction effect from prior knowledge and time, $F(1, 102) = 27.27$, $p = 0.009$, $\eta_p^2 = 0.065$, with the below-average students outperforming the above-average students. There was no interaction between time, reflection and prior knowledge, $F(1, 102) = 3.35$, $p = 0.07$, $\eta_p^2 = 0.032$, and no interaction between time, procedural information and prior knowledge, $F(1, 102) = 0.12$, $p = 0.731$, $\eta_p^2 = 0.001$. All other effects were non-significant.

These outcomes support the assumption that students with different levels of computational fluency or prior knowledge are affected differently by playing the game, but do not support the hypothesis that these students are affected differently by the added support (i.e., reflection prompts and procedural information).

In accordance with our last research question, differences in perception between conditions were investigated. To evaluate whether support in the game affected students' perception of the game, multivariate analysis of variance (MANOVA) with support (reflection and procedural information) as independent factors and perception (playfulness and usefulness) as dependent factors was conducted. The results of Pillai's Trace multivariate test showed no main effect from reflection, $F(1,\ 98) = 0.18$, $p = 0.839$, $\eta_p^2 = 0.004$, no main effect from procedural information, $F(1,\ 98) = 2.76$, $p = 0.068$, $\eta_p^2 = 0.054$, and no interaction between reflection and procedural information, $F(1,\ 98) = 0.99$, $p = 0.375$, $\eta_p^2 = 0.020$. These outcomes indicate that there was no effect of the support (i.e., reflection prompts and procedural information) on students' perception of the playfulness and usefulness of the game.

Further exploration. Results from the analyses of the fourth hypothesis led us to believe that students' basic arithmetic abilities and prior domain knowledge influence whether they are able to learn from the game. Therefore, we evaluated which groups were able to benefit from the game, and explored how they differed on in-game performance. For this evaluation/exploration, we differentiated between students who were computational fluent and students who were not (based on the automation principle of van de Bosch *et al*., 2009, as described earlier). In addition, because the results from the fourth analyses showed that students' prior knowledge affects the effect of the game, we also differentiated in students with above and below average prior knowledge (on condition that the students were computational fluent).

First, we evaluated whether playing the game would help the different groups of students to learn about proportional reasoning. A paired sample *t*-test for the three groups showed a significant difference between total pretest and posttest scores for the computational fluent students with below average prior knowledge, $t(42) = -7.25$, $p < 0.001$, $d = -1.157$, and the computational fluent students with above average prior knowledge, $t(43) = -2.20$, $p = 0.034$, $d = 0.395$. However, the paired sample *t*-test results revealed no significant difference between total pretest and posttest scores for students who were not computational fluent, $t(22) = -0.57$, $p = 0.573$, $d = -0.138$. This indicates that computational fluency is a prerequisite for learning from the game, and that when students are computational fluent, all students, regardless of their prior knowledge, can benefit from the game.

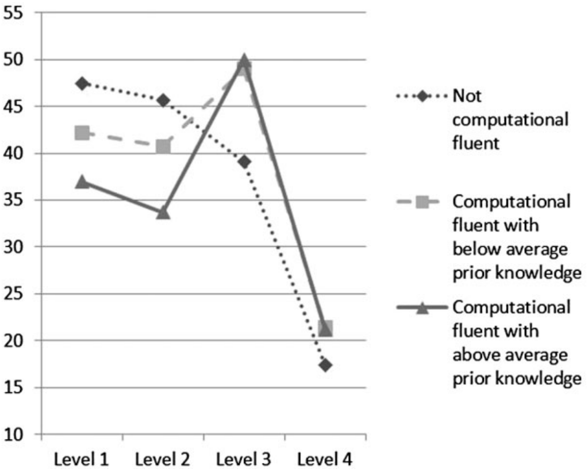


Figure 4 Time per Attempt per Level

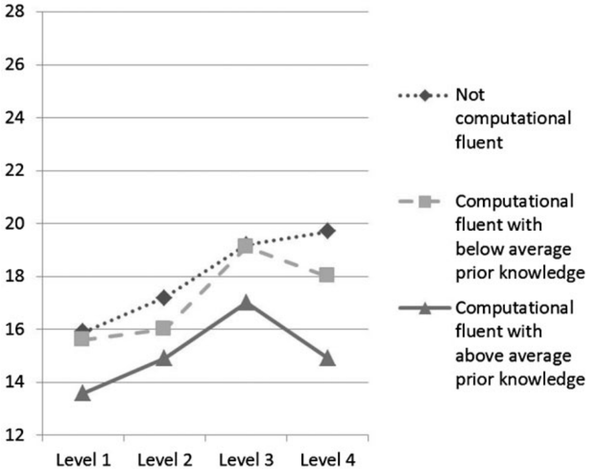


Figure 5 Number of Attempts per Level

Second, we explored the in-game performance from the different groups. Figures 4–6 present the performance of the different groups on the different in-game performance measures per level. In these figures, we can clearly see the increase in difficulty from level 1 (easy) to level 3 (hard): there is an increase in time needed to complete the challenges (time on task) and number of

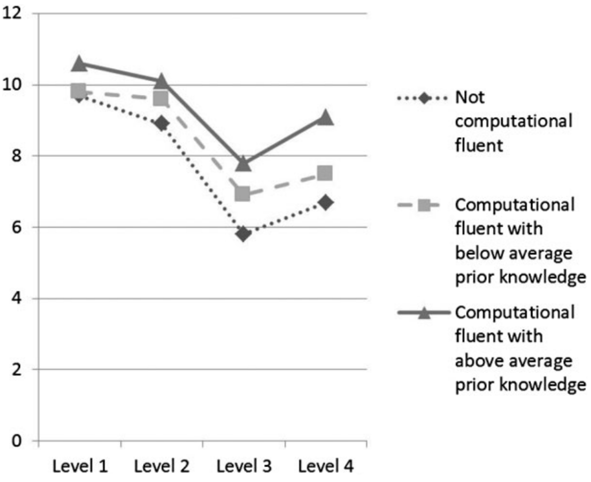


Figure 6 Number of Correct Answers per Level

attempts needed to solve the problems, combined with a decrease in accuracy (number of correct answers).

These graphs also show that the students who are not computational fluent fall behind during the game: in the first level, their number of attempts and number of correct answers are comparable to the computational fluent students with below average prior knowledge, and the only difference is that they invest a little more time per challenge. However, in the second level, their number of attempts increases while they fall behind in the number of correct answers. In the third level, the students who are not computational fluent show a drop in time on task, which could indicate a loss of motivation (instead of mindful play, they just click to finish the game). In addition, we can see a rise in the number of attempts and a decrease in the number of correct answers. The pattern of the latter two measures, however, is similar to the pattern that the other two groups of students show.

These explorative analyses indicate that although students who were not computational fluent were able to start with the game and complete the first level without any severe backlog in score compared with the other students, they fell behind during the subsequent levels. Despite spending sufficient time on the challenges in the second level, they were not able to keep up with the pace of the game, resulting in lower scores – due to an increase in number of attempts and a decrease in accuracy – at the end of the second level. This, combined with the increase in difficulty, could have led to students getting frustrated and giving up, as is demonstrated by the significant drop in time on task, combined with the increase in attempts and decrease in correct answers in the third level.

Discussion

Overall learning and perception

One of the principal outcomes of this study is that students' ability to solve proportional problems increased significantly after playing the game, and that despite the high density of educational content in the game, students generally perceive the games' playfulness as average and usefulness as slightly above average. These were not obvious outcomes because the topic of proportional reasoning is quite demanding for prevocational students, and these students have built up some resistance to learning in general.

Further analyses revealed that students' posttest scores could be predicted by students' computational skills and domain knowledge, and that in-game performance showed an additional (unique) predictive value. This finding suggests that gameplay does indeed matter for acquiring proportional reasoning skills. Although students showed progress on proportional reasoning, analysis of transfer problems showed that there was no transfer. This is in line with previous findings on game-based learning showing that games often foster experiential learning, and that students gather their knowledge in an intuitive and implicit way. This is generally perceived in an inability of the student to explain what was learned, and transfer of learning fails to happen (Chi *et al.*, 1989; Wouters *et al.*, 2008). However, in the current study, students did show some attainment of explicit knowledge representations; students did gain competence in providing explicit representations of the employed procedures. That transfer failed to happen might be because these students were not able to disconnect what they had learned from the context that it was learned in, and therefore were not able to identify how they could use this newly obtained knowledge to help them solve the transfer problems.

Effects of support

The current study evaluated the usefulness of support (reflection and procedural information) for stimulating the process of developing explicit knowledge. Results showed that support in the form of reflection did not affect the performance on proportional reasoning and transfer. The procedural information – which was intended to aid the reflection process and make it more effective for prevocational students – had no additional value as well. This indicates that the provision of procedural information did not serve as an aid for reflection, nor did it help students to extract knowledge from the game. These findings contrast with the work by Johnson and Mayer (2010), who found an effect of a similar implementation of reflection, which they refer to as self-explanation. In the study by Johnson and Mayer, however, participants were university college students, whereas in our work prevocational students were involved. These groups of students have substantially different cognitive capacities and study skills, which may explain the difference in findings. We cautiously conjecture that – based on the cognitive demands reflection requires – our students lacked the capacity to engage in a productive reflection process. In addition, it should be taken into account that the support provided was restricted and was fixed regardless of students' progression. The number of prompts, the timing and the interval could have influenced the effectiveness. Although the game rewarded students who actively participated in the support measures, it might still be that students were reluctant to respond to the prompts and therefore effects of support could have been diminished.

Analysis of students' perception of the game showed that the addition of the support did not affect their opinion on the usefulness or playfulness of the game. Thus, the current addition of explicit representations of educational content to the serious game did not affect how students perceived the game. However, as the results show, the current addition of instructional content did not affect students' learning. This could have been caused by the volume of the explicit representations in the game. A higher density of instructional content in serious games may facilitate students' learning, but can negatively influence the perception of students, and thus disrupt their motivation. Research is needed to explore the balance between instructional content, perception/motivation and effects of the game.

Prior knowledge and computational fluency

Our data suggest that students' capabilities are a decisive factor in the effects of the game. The results of the current study indicate that both students with below average prior knowledge and above average prior knowledge were able to learn from the game, indicating that students who had not previously been able to meet their potential were able to learn successfully about proportional reasoning when using the game. However, computational fluency seems to be a prerequisite for this learning. Students who were computational fluent outperformed students who were not. Only the first group showed significant growth in domain knowledge. Several factors might explain the fact that the students who were not computational fluent failed to learn from the game.

First, computational skills are necessary when solving proportional problems. Research has shown a clear link between students' understanding of the concept of multiplication and their proportional reasoning skills (Tourniaire, 1986). The game is not designed to teach the students computational skills. Although the students are challenged to work on them, and may improve them, their deficit can still create a threshold for actually improving their proportional reasoning.

Second, a lack of computational skills makes the students vulnerable for frustration when the game does not adapt to their level. The current game was designed to progress in difficulty at a fixed pace. Although this progression was designed to fit the pace of prevocational students learning proportional reasoning, students who could not keep up with this pace could fall behind, which could induce frustration and loss of engagement. Explorative analysis of in-game performance of students who were not computational fluent suggests that these students were unable to keep up with the progressive difficulty of the game and lost motivation halfway through the game (around the third level).

Third, students who have not reached computational fluency by the time they are in secondary school are more likely to suffer from severe learning deficits or even didactic resistance. Computational skills are taught in primary school and make up the basis for all future math problem solving (Calhoon, Emerson, Flores, & Houchins, 2007). This means that the group of prevocational students who, despite age, education and remediation, have not mastered these skills is likely to contain a large population of at-risk/low-level students. Furthermore, severe deficits in the development of computational skills have been linked to mathematical disabilities (Calhoon *et al.*, 2007;

Warner, Schumaker, Alley, & Deshler, 1980). It could be that specifically this group is less susceptible for game-based learning because the game creates a setting where the relevant information is clouded by irrelevant additions (e.g., storyline, decorative representations). Mayer (2005) states that decorative representations induce unnecessary processing demands and can distract from learning. Recent findings from Magner *et al*. (2014) imply that this extra processing load can be specifically disturbing for low-level students.

Conclusion

Overall, the results of the current study point towards the fact that math games can be beneficial for the education of prevocational students, and that computational skills can influence their effectiveness. Our results showed that, although most students from our specific target group are able to learn from the game environment, there is also the risk that a number of students will fail to learn. It should be noted that when working with prevocational students, there will be students who will fall behind but who do have the potential to grow and have just missed out on the ideal opportunity, and there will also be students who have already reached their potential, or who have deeper problems – such as learning deficits – that are preventing them from living up to their potential. In the current study, the learning process was stimulated for the students who were computational fluent, but the students who were not computational fluent were not able to learn from the game. This translates to the conclusion that the game was able to stimulate the students who failed to grasp proportional reasoning when working with other educational material but did have the potential to comprehend the subject matter, but that the game could not help the students who seem to have a severe deficit when it comes to math. This might be due to the high processing load of game-based learning environments. Further research should focus on preventing these negative effects and on how to design effective game-based learning environments for those who most need it: the group of learning-resistant students.

The current study set out to research whether support in the form of (structured) reflection could help students' game-based learning. Results indicated no positive effects of reflection. We must, however, interpret these results carefully due to the number of participants. Nonetheless, in line with the focus of this study, it might be interesting to see how adaptive rather than fixed support can affect these students' game-based learning. Providing students with the correct information just when most needed is likely to be more effective than a fixed type of support. Adaptive feedback, adaptive instructional support or even (heterogeneous) collaborative implementation of game-based learning could help accomplish this. In addition, it might be beneficial to compare the motivational and learning effects of game-based learning to those of more traditional education in order to consider the practical implementations of game-based learning within an educational perspective.

Acknowledgements

The authors wish to thank Lieven Verschaffel, Wim van Dooren, Geraldine Clarebout (KU Leuven), Yvonne Mulder (University of Twente), Emily Fox (University of Maryland) and Ludomotion for their contribution to the article. This study was sponsored by NWO under grant number 411-10-900 and FWO under grant number G.0.516.11.N.10.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2010). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*, 1–18. doi:10.1037/a0021017

Barab, S. A., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development*, *53*, 86–107. doi:10.1007/bf02504859

Berthold, K., Eysink, T. H. S., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science*, *37*, 345–363. doi:10.1007/s11251-008-9051-z

Calhoon, M. B., Emerson, R. W., Flores, M., & Houchins, D. E. (2007). Computational fluency performance profile of high school students with mathematics disabilities. *Remedial and Special Education*, *28*, 292–303. doi:10.1177/07419325070280050401

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in* *instructional psychology* (pp. 161–238). Mahwah, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182. doi:10.1016/0364-0213(89)90002-5

Cito, Dutch Central Institute for Test Development (2011). *Cito meting taal en rekenen 2011 (Cito measurement language and math 2011).*

CvE, Dutch Board of Examinations (2014). *Tussenrapportage centraal ontwikkelde examens mbo en rekentoets VO (Report on central examination of math in secondary and vocational education).*

de Vos, T. (1992). *Tempo test rekenen handleiding en verantwoording [tempo test arithmetic manual and justification].* Pearson, Amsterdam.

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences*, *12*, 91–142. doi:10.1207/s15327809jls1201_4

Dihoff, R. E., Brosvic, G. M., & Epstein, M. L. (2003). The role of feedback during academic testing: The delay retention effect revisited. *The Psychological Record*, *53*, 533–548.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers and Education*, *67*, 156–167. doi:10.1016/j.compedu.2013.02.019

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation and Gaming*, *33*, 441–467. doi:10.1177/1046878102238607

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*, 207–219. doi:10.1111/j.1365-2729.2012.00489.x

Habgood, M. P. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, *20*, 169–206. doi:10.1080/10508406.2010

Harel, G., & Behr, M. (1989). Structure and hierarchy of missing value proportion problems and their representations. *Journal of Mathematical Behavior*, *8*, 77–119.

Johnson, C. I., & Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior*, *26*, 1246–1252. doi:10.1016/j.chb.2010.03.025

Kaput, J., & West, M. M. (1994). Missing value proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 237–292). New York, NY: Suny Press.

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers and Education*, *51*, 1609–1620. doi:10.1016/j.compedu.2008.03.003

Kebritchi, M., & Hirumi, A. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers and Education*, *51*, 1729–1743. doi:10.1016/j.compedu.2008.05.004

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers and Education*, *55*, 427–443. doi:10.1016/j.compedu.2010.02.007

Lawton, C. A. (1993). Contextual Factors Affecting Errors in Proportional Reasoning. *Journal for Research in Mathematics Education*, *24*, 460–466.

Leemkuil, H., & de Jong, T. (2011). Instructional support in games. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 353–369). Scottsdale: IAP Inc.

Lepper, M. R., & Cordova, D. I. (1992). A desire to be taught: Instructional consequences of intrinsic motivation. *Motivation and Emotion*, *16*, 187–208. doi:10.1007/bf00991651

Li, M. C., & Tsai, C. C. (2013). Game-based learning in science education: A review of relevant research. *Journal of Science Education and Technology*, *22*, 877–898. doi:10.1007/s10956-013-9436-x

Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research and Development*, *59*, 249–265. doi:10.1007/s11423-011-9192-7

Magner, U. I. E., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, *29*, 141–152. doi:10.1016/j.learninstruc.2012.07.002

Malone, T. (1981). What makes computer games fun? *Byte*, *12*, 258–277. doi:10.1145/1015579.810990

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge University Press.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A Confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*, 48–58. doi:10.1080/02701367.1989.10607413

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *Curriculum Journal*, *16*, 455–474. doi:10.1080/09585170500384529

Papastergiou, M. (2009). Digital game-based Learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers and Education*, *52*, 1–12. doi:10.1016/j.compedu.2008.06.004

Paras, B., & Bizzocchi, J. (2005). *Game, motivation, and effective learning: An integrated model for educational game design*. Paper presented at the Digital Games Research Association (DiGRA), Vancouver, BC, Canada.

Pavlas, D., Jentsch, F., Salas, E., Fiore, S. M., & Sims, V. (2012). The play experience scale: Development and validation of a measure of play. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*, 214–225. doi:10.1177/0018720811434513

Rick, J., Bejan, A., Roche, C., & Weinberger, A. (Eds.). (2012). *Proportion: Learning proportional reasoning together* (Vol. 7563). Berlin, Germany: Springer.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, *30*, 344–360. doi:10.1007/s11031-006-9051-8

Squire, K. (2005). Changing the game: What happens when video games enter the classroom. *Innovate: Journal of Online Education*, *1*(6), 20. http://www.innovateonline.info/pdf/vol1_issue6/Changing_the_Game-__What_Happens_When_Video_Games_Enter_the_Classroom_.pdf

Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computers and Entertainment*, *3*, 3–24. doi:10.1145/1077246.1077253

ter Vrugte, J., & de Jong, T. (2012). How to adapt games for learning: The potential role of instructional support. In S. Wannemacker, S. Vandercruysse, & G. Clarebout (Eds.), *Serious games: The challenge* (Vol. 280, pp. 1–5). Berlin, Germany: Springer Berlin Heidelberg.

Tourniaire, F. (1986). Proportions in elementary school. *Educational Studies in Mathematics*, *17*, 401–412. doi:10.1007/BF00311327

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. *Educational Studies in Mathematics*, *16*, 181–204. doi:10.1007/bf02400937

van de Bosch, E., Jager, J., Langstraat, H., Versteeg, B., & de Vries, M. (2009). *Remedierend Rekenprogramma Automatiseren De Zuidvallei [Remediating Mathprogramme Automation De Zuidvallei]*. Giralis, Ede.

van Dooren, W., de Bock, D., Evers, M., & Verschaffel, L. (2009). Students' overuse of proportionality on missing-value problems: How numbers may change solutions. *Journal for Research in Mathematics Education*, *40*, 187–211.

Vandercruysse, S., Vandewaetere, M., & Clarebout, G. (Eds.). (2012). *Game-based learning: A review on the effectiveness of educational games* (Vol. 1). Hershey, PA: IGI Global.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research*, *34*, 229–243.

Warner, M. M., Schumaker, J. B., Alley, G. R., & Deshler, D. (1980). Learning disabled adolescents in the public schools: Are they different from other low achievers? *Exceptional Education Quarterly: Special Issue on Special Education for Adolescents and Young Adults*, *1*, 27–36.

Wolters, C. A. (1998). Self-regulated learning and college students' regulation of motivation. *Journal of Educational Psychology*, *90*, 224–235. doi:10.1037//0022-0663.90.2.224

Wouters, P., Paas, F., & van Merriënboer, J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, *78*, 645–675. doi:10.3102/0034654308320320

Wrzesien, M., & Alcañiz Raya, M. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers and Education*, *55* (1), 178–187. doi: 10.1016/j.compedu.2010.01.003