Computers in Human Behavior 29 (2013) 2471–2479

Contents lists available at SciVerse ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Does individual or collaborative self-debriefing better enhance learning from games?

Hans van der Meij *, Henny Leemkuil, Juo-Lan Li

University of Twente, Faculty of Behavioral Science, Department of Instructional Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Available online 26 June 2013

Keywords:
Collaborative learning
Self-debriefing
Game-based learning

ABSTRACT

The primary aim of this study is to find out whether use of different self-debriefing modes affects learning from a game. In self-debriefing participants are led to reflect upon their game experiences by a set of debriefing questions. Two conditions were compared: Individual and Collaborative self-debriefing. The 45 participants first played the game of Lemonade Tycoon Deluxe, were tested for knowledge and self-debriefed in pairs or alone. Then they played the game once more and were tested again. Game scores increased significantly from the first to the second round of gameplay to an equal degree in both conditions. Knowledge scores of participants in individual self-debriefing increased significantly more than those of participants in the Collaborative condition. The study shows that game-based learning can be effectively scaffolded with self-debriefing. Future studies might investigate whether the type of self-debriefing differentially affects game motivation. In addition, attention to the role of feedback is called for.

1. Introduction

Games are big business. In the United States alone, people spend about $11.7 billion on computer and video game software per year (Tobias & Fletcher, 2011). The popularity of games is understandable. Games have the capacity to capitalize on people's intrinsic motivation and to keep them engaged in gameplay for a (very) long time. Both factors are also important for learning, as Tobias, Fletcher, Dai, and Wind (2011) acknowledge in their recent review on game-based learning when they indicate that the motivating and engaging properties of games make them promising for use in education.

A noteworthy issue in the educational usage of games is the provision of support. Games capitalize on experiential learning. There is hardly any direct instruction and scaffolding also tends to be lacking. What this means for learning can be grasped from research on simulations that take a similar approach. Simulations also do not offer direct instruction, and students must get to know the domain from experience or examples. Several large-scale evaluations have shown that simulation-based learning is effective only when supported (e.g., Eysink et al., 2009; Gijlers & De Jong, in press; Gum, Greenhill, & Dix, 2011). A diverse set of cognitive and metacognitive scaffolds for simulations have already been designed and examined. Scaffolding for games, however, is still largely unexplored territory. As Leemkuil and De Jong (2011) state, "This road is still ahead for game-based learning" (pp. 354). Considerable research is needed on finding suitable and effective means to support game-based learning (Garris, Ahlers, & Driskell, 2002; Leemkuil & De Jong, 2012).

This paper examines debriefing as an instructional scaffold for learning from games. Debriefing can be defined as "facilitated or guided reflection in the cycle of experiential learning" (Fanning & Gaba, 2007, p. 116). Debriefing prompts participants to reflect on their game experience. It engages them in an analytic process that revolves around review and analysis of the events that occurred during gameplay (Lederman, 1992; Warrick, Hunsaker, Cook, & Altman, 1979). The main purpose of debriefing lies in changing participants' viewpoints from subjective to objective, and clearing up misunderstandings or mistakes (e.g., Linn, Lee, Tinker, Husic, & Chiu, 2006; van Ments, 1983).

Debriefing sessions, like those that are customary in the military and in healthcare, are usually led by experts. But debriefing can also be self-led (e.g., Boet et al., 2011; Fanning & Gaba, 2007; Moreno-Ger, Burgos, Martínez-Ortiz, Sierra, & Fernández-Manjón, 2008). The players then do the debriefing on their own, using a set of debriefing questions for guidance. Self-debriefing is an attractive mode in educational settings. This study examines the effect of self-debriefing mode (i.e., individual or collaborative) on game-based learning.

2. A conceptual model for debriefing

In games that offer rich, dynamic, ever-present environments, learners often stick to active experimentation based on trial and

* Corresponding author. Tel.: +31 053 4893656; fax: +31 053 4892895.
E-mail addresses: h.vandermeij@utwente.nl (H. van der Meij), h.h.leemkuil@utwente.nl (H. Leemkuil).

0747-5632/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.chb.2013.06.001

error (Kirriemuir & McFarlane, 2004). Their information processing is data-driven and reactive, which often does not lead to a deeper understanding of the content of the game (e.g., Hickey, Kindfield, Horwitz, & Christie, 2003; Koops & Hoevenaar, 2012). Reflection is needed to build such an understanding[1]. In other words, participants need to engage in the process of thinking over earlier experiences with the aim of enhancing comprehension.

Two main methods of reflection are generally considered: reflection-in-action and reflection-on-action (Lederman & Kato, 2005). In the first approach, reflective opportunities are (made) available during gameplay. The main way to do this involves introducing cognitive conflicts (see Blumberg, Rosenthal, & Randall, 2008; Garris et al., 2002). Such an approach was taken in a recent study by Koops and Hoevenaar (2012), who designed a game that suddenly and substantially increased in difficulty each time a new concept was introduced. These events were meant as a trigger mechanism, which should get participants temporarily out of their game mode and into their learning mode. The other approach is to stimulate reflection after gameplay. This is the avenue taken in the present study, and for which we believe debriefing holds considerable promise.

Debriefing is widely considered to be a critical component in experiential learning (e.g., Fanning & Gaba, 2007; Kolb, 1984; Koops & Hoevenaar, 2012; Kriz, 2010; Lederman, 1984, 1992). Lederman (personal communication) puts it very succinctly: "The game generates the behavioral data and the debriefing is where the learning is guided". Tannenbaum and Cerasoli (2013) recently conducted a meta-analysis of research in various disciplines to examine the empirical support for the claimed effectiveness of debriefing. Their search of the literature yielded a total of 46 controlled studies, most of which involved within-group comparisons. Their analysis led to the conclusion that debriefing yields a significant advantage over non-debriefing conditions. An average improvement of approximately 25% ($d = 0.67$) on learning outcomes was found.

The debriefing process is often characterized as a set of consecutive stages or phases with descriptions or example questions (e.g., El-Shamy, 2001; Lederman, 1992; Lennon, 2006; McGaghie, Issenberg, Petrusa, & Scalese, 2010; Petranek, Corey, & Black, 1992; Sims, 2002; Steinwachs, 1992; Thiagarajan, 1992; van Ments, 1983; Vollmeyer & Rheinberg, 1999). Authors differ in the number and type of stages that are distinguished. For instance, Van Ments (1983) suggests three main stages in which participants begin by establishing the facts, follow up with analyzing the causes of behavior and end in planning actions. Sims (2002) adopted Kolb's (1984) experiential learning cycle in which the following four phases are distinguished: concrete experience, reflective observation, abstract conceptualization, and active experimentation. In Sims's model the first phase of debriefing accordingly revolves around questions that ask participants to describe their perceptions, thoughts and feelings during the past game. Next, the participants are asked to consider these experiences from different points of view (e.g., concerns, virtues, values, and characteristics). Thereafter, participants are stimulated to reflect on (new) concepts, theories or models. In the last phase of debriefing, the participants are prompted to come up with some rules or guidelines and consider future actions.

Many authors acknowledge that an important foundation for their stage descriptions comes from Kolb's (1984) experiential learning cycle (Fanning & Gaba, 2007). Just like Sims (2002), we also took Kolb's framework as the basis for our conceptual model for debriefing in our study on game-based learning (see Fig. 1). We expanded Kolb's framework with the seven E's of Petranek et al. (1992), who suggest that in debriefing one should attend to events, emotions, empathy, explanations, every day, employment, and evaluation. For each topic, we then formulated one or more leading questions (compare Lederman, 1984; Lennon, 2010; McGaghie et al., 2010) to guide participants into reconstructing their experiences step by step, appreciating their own and others' perspectives, discovering or realizing learning objectives, and planning possible future application, respectively. This conceptual model served for generating the specific debriefing questions for the participants in our empirical study (see Section 5).

[1] Not all game-based learning depends on reflection. For instance, video games such as Grand Theft Auto and Super Mario Cart can modify perceptual skills without debriefing (Green & Bavelier, 2003). The present study revolves around a simulation game from which the participants must learn the underlying business model with concepts, principles and heuristics.

3. Important modes of debriefing

Debriefings can vary in three important ways or modes: (a) expert-led or self-debriefing; (b) oral or written debriefing; and (c) individual or collaborative debriefing. Each dimension is discussed below, as is the timing factor.

3.1. Expert-led versus self-debriefing

Debriefing originated with the military. In a debriefing session, participants in a military campaign or war game were brought together to describe their experiences, to reason about the actions that occurred and to advance new strategies. Within the same domain, debriefing also became a means for defusing combatants. Debriefing became therapeutic, in the sense that it was employed to reduce the psychological damage of traumatic war events. Somewhat later, debriefing came to serve a similar purpose in clinical practice where it was employed to mitigate stress or helped desensitize patients with phobias. Invariably, these debriefing sessions were led by a facilitator or expert, and this has since remained the standard practice (Fanning & Gaba, 2007; Lederman, 1992). In educational settings it is often impossible, or simply too costly, to have an expert lead the debriefing. Thus, educators have been looking for an alternative, which they found in self-debriefing.

To date only a few empirical studies on self-debriefing have been conducted. A survey among 10.166 pilots who had received flight simulation training found no difference in appraised effectiveness for expert-led or self-debriefing (Neill & Wotton, 2011). In addition, the survey found no difference for the usage of video in either method. The incorporation of videotaped recordings of participant behavior was not considered to enhance the usefulness of the debriefing (see also Sawyer et al., 2012).

Only one experiment contrasting instructor-led debriefing with a self-debriefing condition was found in the literature. In that experiment (Moreno-Ger et al., 2008), fifty anesthesiology residents began with a briefing about crisis management, after which they played a high-fidelity simulation of an anesthesia crisis scenario. Thereafter, half of the participants engaged in an instructor-led debriefing, while the other half engaged in video-assisted self-debriefing. In both conditions, the debriefing revolved around the same tool for assessment that the participants had learned about prior to the training. Immediately after the debriefing, the participants played another simulation of a crisis scenario. The behavior of the participants during both simulations was observed and scored to yield a pre-test and post-test score. Data analyses indicated that there was a significant improvement in performance, but there was no difference between conditions. Both debriefing methods were found to be equally effective.

In short, the findings from the few empirical studies that exist suggest that there is no difference in effectiveness between

Phases	Topics	Leading question(s)
Concrete experience	Events	What happened?
	Emotions	How did you feel?
Reflective observation	Empathy	How do you value this experience?
Abstract conceptualization	Explanations	What did you learn? What would have happened if . . . ?
Active experimentation	Every day	How are the game events and reality connected?
	Employment	How do you go on from here?
	Evaluation	What would you do differently?

Fig. 1. A conceptual model for debriefing, with phases, topics, and leading questions.

expert-led and self-debriefing. Fanning and Gaba (2007) draw the same conclusion in their literature review. However, they also note that it is important to structure the self-debriefing with aids to self-assessment, and/or collaboration scripts.

3.2. Oral versus written debriefing

Debriefings originally used to be conducted orally, and to this very day this has remained the most common format (e.g., Fanning & Gaba, 2007; Linn et al., 2006). The reasons for this choice are often not explicitly stated. An oral debriefing is presumably preferred over a written one for the following reasons: (a) historically the debriefing has always been conducted orally; (b) an oral debriefing accommodates all participants at the same time; and (c) an oral debriefing enables rapid exchanges between facilitator and participants.

Written debriefing is only occasionally employed. Even so, Petranek (2000) strongly argues in its favor when compared to its oral counterpart. "After working in the field for the past 25 years, one of my most significant discoveries is that participants learn much more by writing about the simulation than just by playing or orally debriefing them" (p. 108). An argument supporting his claim is that a written debriefing lengthens the time for reflection. Where oral debriefings can sometimes be pressed for time, a written debriefing gives participants more opportunities to deliberate on their past experience. In addition, the written format offers private time for reflection. The participant can address his or her own personal feelings and insights, and even the voices of participants who normally remain quiet during group discussions can now be heard (compare De Vries, Van der Meij, Boersma, & Pieters, 2005; Oertig, 2010). Yet another argument in favor of written debriefing is that having to write down one's thoughts and feelings is more demanding than just talking about them. Writing invites explicit articulation and can be effective in revealing misconceptions. A final argument is that a written debriefing is easily evaluated. It can be used by both facilitators and participants, and in a formative or summative way, to analyze and assess what the participant has learned from the experience and the debriefing.

With self-debriefing, the use of written debriefing falls naturally into place. The debriefing can be structured with support such as concepts, suggestions or leading questions on paper. The written responses to these questions are then open for self-reflection, discussion or feedback.

3.3. Individual versus collaborative debriefing

In discussing this contrast the perspective should be extended a bit to include the organization of the experience also, since it is closely linked with the form of debriefing. An individual experience is usually followed by an individual, written debriefing, whose benefits have already been discussed.

When the experience has been undergone as a team, the debriefing also tends to be conducted with the team as a whole. Such a set-up is now regularly employed within the healthcare industry, among others. Nurses and doctors who work together in their daily lives may jointly engage in a simulation plus debriefing (Kriz, 2010). The experience provides all participants with an opportunity to jointly practice task and team-related skills in a safe environment. The debriefing provides individuals an opportunity to reflect on their own contributions and perspectives, as well as those of other team members (Kriz, 2010). In addition, it may bring to light new ideas, reveal misconceptions, and contribute to the construction of common understanding (Schoepfle & Werner, 1999). According to Gum et al. (2011), a debriefing with peers can stimulate participants to 'listen for confusion at the edges of understanding' (pp. 24), evoking a discussion that addresses knowledge that otherwise remains intuitive or tacit. Besides the individual benefits that accrue to simulated practice followed by debriefing, the joint actions of the participants undertaken together also contribute to team building and collaboration.

3.4. Timing

Two factors are often mentioned in discussions about timing, namely the interval between the simulation or game and the debriefing, and the duration of the debriefing (e.g., Fanning & Gaba, 2007; Norman, 1993). There is considerable consensus on the best choice for both factors.

For the time interval, the advice is mainly to keep the delay between the participants' experience and the debriefing minimal. A short delay can be useful when emotions have run too high during the experience and participants must cool down (Schön, 1983). However, the preferred option is for the debriefing to take place immediately after completing the simulation or game (Cantrell, 2008; Oertig, 2010; Wotton, Davis, Button, & Kelton, 2010). This is also the most common practice (Fanning & Gaba, 2007). One argument in favor of immediate debriefing is that the participants' memory is still fresh. Participants can still recall how they felt during the experience, and they can also more easily remember the events and actions they engaged in than when debriefing is delayed. When debriefing is immediately followed by another round of practice, the insights gleaned can be directly applied in another round of simulation or gameplay. Such a set-up makes the most of the intertwining of the experience and the debriefing. That is, as the debriefing advances the participants' knowledge of game events and concepts, and of the learning objectives, it also has an attention-directing function for new rounds of gameplay (compare Qudrat-Ullah, 2007; Reed, 2012).

The debriefing should be long enough to be effective (Sawyer et al., 2012). That is, in setting the time limit one should consider the fact that the debriefing usually progresses through various stages and that different topics such as events, emotions, and explanations are to be addressed. Although practices vary considerably, the debriefing is usually fairly time-consuming. It is not uncommon for it to take about as much time as the experience itself. Several studies report a duration that varies between 20 and 40 min (e.g., Koops & Hoevenaar, 2012; McGaghie et al., 2010; Moreno-Ger et al., 2008; Norman, 1993; Oertig, 2010; Qudrat-Ullah, 2007; Reed, 2012; Sawyer et al., 2012; Wotton et al., 2010).

4. Design of the study and research questions

The primary aim of this study is to find out whether use of different self-debriefing modes affects learning from a game. Participants in the study first respond to questionnaires to assess their prior game experience and game motivation. These measures are intended to check on the equivalence of the groups produced by random assignment. Thereafter, participants play the game individually and are tested for knowledge. Next, the participants engage in self-debriefing, followed by another round of individual gameplay and final testing.

The self-debriefing is structured. Participants are supported by a set of Debriefing Questions (see Section 5). There are two debriefing conditions: Individual and Collaborative. In Individual debriefing each participant completes the debriefing alone. In the Collaborative self-debriefing condition two participants jointly work through the Debriefing Questions.

Question 1: Is there an effect of time and condition on game score?

The prediction for time is that game scores in the second round of gameplay will be higher than in the first. Debriefing, in combination with testing, is presumed to be the main cause for the increase, as it is likely to make the participants aware of important gaps regarding what they do not yet know about game concepts, principles and heuristics, as well as about learning objectives. These factors presumably have an attention-directing effect in game replay (compare Finley & Benjamin, 2012; Reed, 2012; Roediger, Agarwal, McDaniel, & McDermott, 2011).

The prediction for condition is that Collaborative debriefing supports reflection more strongly than Individual debriefing, and thus will lead to higher game scores for the participants in that condition in the second round of gameplay.

Question 2: Is there an effect of time and condition on learning?

The prediction for time is that learning will increase in both conditions. More specifically, it is expected that scores on Test 2 will be higher than scores on Test 1. This expected effect is attributed to the combination of debriefing, testing and replay.

The prediction for condition is that participants will gain more from Collaborative debriefing than from Individual debriefing. Research on collaborative learning reveals that scaffolding of the interaction is an important condition for learning (Gijlers & De Jong, 2009; Gijlers & De Jong, in press; Van der Meij, Albers, & Leemkuil, 2011). In this study the set of debriefing questions provide such a scaffold. The questions offer structural support for the participants' discussion. For individuals, the questions should stimulate self-reflections. For participants who collaborate, the debriefing questions provide a specific focus for their dialogue, inviting each to provide arguments and explanations that the other participant can understand and evaluate, and to contribute to the discovery of misconceptions.

Question 3: Is there a relation between game score, test score and time?

We explore the relation between game scores and test scores at different time points. The prediction is that time affects the correlation between game score and test score. Specifically, we expect a positive relation only at the end of the study (i.e., game score for round 2 and test score on Test 2). During the first round of gameplay participants must necessarily spend time on getting to know game features such as game objects and permissible moves, whereas in the second round they already know these and can pay more attention to game features that pertain to the underlying business model of the game.

5. Method

5.1. Participants

The participants (mean age 17 years) were 45 male students from a single-sex Senior High School in Taiwan. All participants came from the same classroom and had already known each other for at least one school year. Fifteen participants engaged in Individual self-debriefing, and 30 participants in paired Collaborative self-debriefing. Participants were randomly assigned to condition. Because the native language of the participants was Mandarin, all materials were in that language, with the exception of the game. Two pilot tests revealed that this student population could work with the English interface from the game. Even so, a translation in Mandarin for all key terms in the game was presented on the interactive whiteboard. Participants could also ask the experimenter, or look up an unknown word in the dictionary.

5.2. Materials

5.2.1. Game

Participants played "Lemonade Tycoon Deluxe". The main goal of this commercial off-the-shelf (COTS) strategy game is to set-up a successful lemonade business. The game revolves around the law of supply and demand. It requires the skillful use of strategies in manipulating variables such as hiring staff, recipe(s) for making lemonade, stock, location, stand, price of lemonade and marketing expenditure. Success further depends on non-manipulable factors such as weather situations, news, and popularity of and satisfaction with the lemonade. In the ''Time Challenge'' variant, the player must try to optimize profit within a simulated 15-day time frame. Lemonade Tycoon keeps a record of the profit (or loss) for each day, and also presents an average for the whole game period. In this study, the game score is the mean daily profit earned during the fifteen simulated days of gameplay.

An attractive feature of the game is that players can repeatedly improve their knowledge and skills under slightly changing conditions. In the ''Time Challenge'' variant that was used, players can explore and test their insights fifteen times. The game score from the previous day(s) limits the possibilities, but even so the game frequently gives players the chance to develop and adjust their understandings of fundamental game features, along with the possibility of adapting to what happens during the day's play.

Players can also obtain a daily update on the consequences of their actions. They can request detailed information about the day's outcomes by looking at the constitutive elements on the statements from their bank account (i.e., permits, staff, marketing, and stock). In addition, the player can consult a daily report with specific information about, in particular, cost for types of stock (e.g., sold stock, spoiled stock, and wasted overstock), gross profit or loss, and gross margin.

5.2.2. Game experience questionnaire

This questionnaire consists of five closed questions that inquire about the participants' prior computer game experience. One question specifically asks whether participants have played Lemonade Tycoon earlier. Another question asks about the mode of playing (e.g., alone, with a remote partner, with multiple players). Three questions cover the amount of experience in playing computer games (i.e., ''How many hours per month on average did you play games in recent months?''; ''How much experience do you have with computer games?''; ''How many hours on average did you play a strategy game like the Sims, SimCity or Civilization last week?''). Reliability for the last three experience questions was moderate (α = 0.53).

5.2.3. Game motivation questionnaire

This questionnaire is an adapted version of the FAM-questionnaire developed by Rheinberg, Vollmeyer, and Burns (2001), which measures the students' appraisals of Probability of success, Anxiety, Interest, and Challenge for a future undertaking (also see Vollmeyer & Rheinberg, 1999). First, the participant receives a description of the task they will be facing. Thus, they are told that they are going to play the game Lemonade Tycoon Deluxe. Next, the questionnaire asks the participants for their self-appraisals for this undertaking. Probability of success refers to the student's belief that he or she can succeed (e.g., ''I think I can successfully complete this task''). Anxiety represents the negative incentive of failure (e.g., ''When I think of this task I am worried''). Interest refers to positive affect and evaluations (e.g., ''I like it that you learn new things with this task''). Challenge assesses whether the assignment is perceived as an achievement situation in which the student wants to have success (e.g., ''I'm really going to try as hard as I can on this task''). All constructs are measured with five questions. Answers are given on a 7-point Likert-scale with one extreme described as ''Does not fit me'' and the other as ''Fits me''. Reliabilities for the four constructs were satisfactory: Probability of success (α = 0.73), Anxiety (α = 0.75), Interest (α = 0.86), and Challenge (α = 0.85).

5.2.4. Knowledge test

The paper-and-pencil test measures knowledge about game concepts, principles and heuristics. True/false and multiple-choice items are awarded 1 point for the correct answer. The open-ended answers are scored with the aid of a codebook, which specifies the units of meaning that should be present in the answer. Each unit is awarded 1 point. The maximum possible score for the whole knowledge test is 25 points. There are two parallel tests (Test 1 and Test 2) with the same distribution of questions and an identical scoring procedure. Reliability for Test 1 was low (α = 0.30), for Test 2 it was moderate (α = 0.57). The low to moderate reliability scores for the tests are ascribed to (a) developing game knowledge, and (b) variations in the nature and the content of the test items.

Questions about *concepts* ask for definitions or descriptions of phenomena such as customer satisfaction and popularity. There are five questions about concepts, of which the first two are true/false items (e.g., ''Indicate whether the following statement is true or false: Assets in the results include cash, equipment and stock''). The other three questions are open-ended (e.g., ''Mention three reasons that make customers dissatisfied''). The maximum score for the five concept questions is 9 points.

Questions about *principles* inquire about the ways in which events and actions influence each other, as well as influencing outcomes. There are two multiple-choice questions (e.g., ''Which statement is true about locations: (a) You do not have to pay for the park, (b) The rent for the beach is higher than for the mall, (c) The rent downtown is as high as in the railway station, and (d) The rent is highest in Magic Gardens''). Four questions are open-ended (e.g., ''Which do you think is better: To buy in small amounts so that there will be no leftover stock at the end of the day, or to buy all ingredients (except ice cubes) in large quantities? Please explain your choice.''). The maximum score for the six principles questions is 10 points.

Questions about *heuristics* refer to the coherence between the various principles of the game. There are three heuristics questions. Each question first describes a situation or an event (e.g., ''You have just moved to the park. You have been open on location for one day and your popularity is now 10%. The weather forecast predicts sunny weather with a temperature of 15°. Your current recipe consists of 6 lemons, 3 units of sugar and 3 ice cubes. You own the upgrade of a teller. You sold 43 out of 60 cups yesterday.''), and then asks what specific actions the player should be taking, and what effects of these actions are expected. The maximum score for these questions is 6 points.

5.2.5. Debriefing questions

The debriefing was scaffolded with a set of open-ended questions, preceded by an introduction that informed participants about how these were to be used: ''These questions are to help you to recall what you have just experienced, to re-examine your actions and those that you may have missed, and to make you think about strategies and future applications.'' Participants who debriefed individually were asked to think about these questions and write down their answers. In addition, they were informed that their answers would not be assessed, serving only for personal reflection and understanding. Participants who debriefed collaboratively were also asked to discuss answers before writing them down. Participants in the two conditions received the same debriefing questions.

The debriefing questions were organized around two main points in the game, namely 'before starting a day' and 'after a day'. Following the scheme presented in Fig. 1, there were questions revolving around: (a) concrete experiences (e.g., ''What kinds of customer complaints have you experienced?''), (2) reflective observations (e.g., ''What were your reasons for selecting a location for your lemonade stand?''), (3) abstract conceptualizations (e.g., ''One player stated that money spent on advertising is not profitable. Do you agree? Why (not)?'', and (4) active experimentation (e.g., ''Let's suppose there is a festival on the beach but it is rainy and cold. What would you do in such a situation?'').

5.3. Procedure

The experiment was conducted in a large computer room with rows of computers and an interactive whiteboard in front for presentations. Following a brief, plenary description of the game 'Lemonade Tycoon Deluxe', participants completed the questionnaires on game experience and game motivation. Then, participants could play the game individually until they completed the 15-day Time Challenge (a maximum of 25 min sufficed for all participants). Thereafter, participants were individually tested for knowledge (Test 1). After completing the test, participants in the Individual debriefing condition worked through the Debriefing Questions on their own, writing down their answers on paper. Participants in the Collaborative debriefing condition worked through the Debriefing Questions jointly in pairs. After discussing their views, one team member wrote down their answers on paper. Debriefing took a maximum of 30 min. Then participants played the game individually again, for another 15-day Time Challenge. The experiment was rounded off with another individual test (Test 2). Neither test answers nor answers to debriefing questions were disclosed at any time during the study.

5.4. Analyses

The data from the Game Experience Questionnaire and the Game Motivation Questionnaire were analyzed to check on the random distribution of participants across conditions. These analyses revealed that only one of the forty-five participants had limited prior experience with Lemonade Tycoon; all others had not played the game before. In addition, there was no statistically significant difference between conditions for any of the other game experience questions. Likewise, there were no statistically significant differences between conditions for any of the four motivational constructs.

Explorations of the relations between amount of initial game experience and test scores yielded non-significant correlations, all of which were close to zero (ranging between $r = -0.03$ and $r = 0.07$). Likewise, explorations of the relations between all initial game motivation measures and test scores yielded non-significant correlations, with scores ranging between $r = -0.23$ (for Anxiety and Test 1) and $r = 0.22$ (for Challenge and Test 2).

Repeated measures ANOVAs were computed to examine changes due to time for game scores and test scores. Correlations were used to examine the relations between test scores and game scores. All tests are two-tailed with alpha set at 0.05. For significant results, or trends in the predicted direction, we use Cohen's (1988) d-statistic to report effect size. These tend to be qualified as small for $d = 0.2$, medium for $d = 0.5$ and large for $d = 0.8$.

6. Results

Question 1: Is there an effect of time and condition on game score?

A repeated measures ANOVA revealed that there was a main effect of time on game scores, $F(1,43) = 4.53$, $p = 0.039$, $d = 0.38$. Table 1 shows that the game scores in round two were higher than in round one. The analysis also revealed that there was a main effect of condition on game scores, $F(1,43) = 5.47$, $p = 0.024$, $d = 0.74$. The game scores in the Collaborative debriefing condition were higher than in the Individual debriefing condition (see Table 1). This stems from their higher starting condition. That is, participants in the Collaborative debriefing condition tended to do better than those in the Individual debriefing condition on the game scores for round one, $F(1,43) = 3.72$, $p = 0.060$, $d = 0.64$, and this difference was maintained in round two, $F(1,43) = 3.70$, $p = 0.061$, $d = 0.65$. There was no interaction between condition and time. In other words, participants in both conditions showed similar gains in game scores.

Table 1
Game Scores for the first and second round of gameplay.

	Game score round 1 Mean (SD)	Game score round 2 Mean (SD)
Individual debriefing ($n = 15$)	10.99 (4.95)	13.29 (4.66)
Collaborative debriefing ($n = 2 * 15$)	14.74 (6.66)	17.41 (7.59)
Total	13.49 (6.34)	16.03 (6.98)

Question 2: Is there an effect of time and condition on test score?

A repeated measures ANOVA revealed that there was a main effect of time for test scores. $F(1,43) = 35.1$, $p = 0.000$, $d = 0.91$. Table 2 shows that the test scores in round two were higher than in round one. There was no main effect for condition. However, there was an interaction between time and condition, $F(1,43) = 4.18$, $p = 0.047$. Table 2 shows that this interaction effect is due to higher knowledge gains for Individual debriefing. That is, while both conditions had almost the same test score for round one, $F < 1$, n.s., in round two the Individual debriefing condition scored significantly higher, $F(1,43) = 5.83$, $p = 0.020$, $d = 0.75$.

Question 3: Is there a relation between game score and test score, and time?

Taking all participants together, significant correlations were found between the two game scores, $r = 0.41$, $p = 0.01$, and between the two test scores $r = 0.38$, $p = 0.010$. No relations were found between the game and test score at any point in time. The correlation between the first game and test score was close to zero ($r = 0.08$). There was likewise no relationship between these scores on round two ($r = -0.10$).

Table 3 shows the results from the correlational analyses split for the two conditions. In the Collaborative debriefing condition a significant correlation was found only between Test 1 and Test 2. In the Individual debriefing condition there was a significant correlation only for the two game scores.

7. Discussion

The study revealed a significant improvement in *game scores* over time. Participants made more effective decisions when they played the game the second time compared to their first round of gameplay. Contrary to what was expected, there was no effect of condition on the increase in game scores. Participants who had debriefed together did not increase their game score more than those who did so alone. The finding is a signal that collaborative debriefing did not result in reflection that prepared the participants better for another round of gameplay.

The study further showed that participants achieved considerable *knowledge gains*. The reported effect size of 0.9 indicated the presence of a large effect. This outcome indicates that game-based

Table 2
Test Scores for the Two Administrations.

	Test 1 Mean (SD)	Test 2 Mean (SD)
Individual debriefing ($n = 15$)	12.00 (3.98)	17.20 (4.36)
Collaborative debriefing ($n = 2 * 15$)	11.50 (2.52)	14.03 (4.04)
Total	11.67 (3.05)	15.09 (4.37)

Table 3
Correlations between Test Scores and Game Scores.

	Test 1	Test 2	Game score round 1	Game score round 2
Test 1		*0.20*	*0.09*	*−0.30*
Test 2	**0.54**[**]		*−0.03*	*−0.38*
Game score round 1	**0.12**	**0.02**		*0.62*[*]
Game score round 2	**0.10**	**0.11**	**0.30**	

The italic cells above the diagonal display the scores for participants from the Individual debriefing condition ($n = 15$).
The bold cells below the diagonal display the scores for participants from the Collaborative debriefing condition ($n = 2 * 15$).
[*] $p < .05$.
[**] $p < .01$.

learning was effectively stimulated by the combination of testing, debriefing and another round of gameplay. Because this gain was achieved within a short time frame, this bodes well for the possibility of creating scaffolds that contribute to learning from games. The difference between the scores for Test 2 and Test 1 was both statistically significant and substantial.

A significant interaction for time and condition on test scores was found. In contrast to what was expected, participants who had engaged in individual debriefing outperformed those who had debriefed collaboratively. An explanation could be that the collaboration was not sufficiently scaffolded (compare Gegenfurtner, Veermans, & Vauras, 2013). The collaboration may have engaged the partners in a discussion on superficial rather than fundamental game features. That is, it is possible that the debriefing between the collaborating partners revolved more around a discussion of each partner's personal game experiences, choices and events, while individuals pushed more for deeper understanding. In a related study, Van der Meij et al. (2011) also found no benefit of collaborative over individual gameplay. This was ascribed to the superficial nature of the communication. Partners were found to discuss mainly what game move to make rather than giving arguments for or against such a choice (see also Blumberg & Randall, 2013). Perhaps the players' discussion of the set of debriefing questions was of a similar nature, revolving more around differences in game events, choices and personal experiences rather than stimulating partners to articulate conceptions and misconceptions of game fundamentals. Individual players are not 'distracted' by another player, and therefore may concentrate more on reflecting about game features that build understanding of the underlying business model. It is possible to verify whether this speculation is correct by recording and analyzing the partners' communication about the set of debriefing questions.

When considered from a pragmatic perspective, the favorable finding for individual debriefing is good news. In schools it is much easier to let students play the game on their own and also do the debriefing on their own than to arrange for the latter to be done collaboratively. Both the study by Van der Meij et al. (2011) and the current one show that communication with others is not needed to boost gameplay and learning.

For all participants together, the results showed that there were significant correlations between the two *game scores*, and between the two *test scores*. These findings indicate that there is consistency for these measures. Participants who do relatively well in the first round of gameplay and testing perform similarly in the second round. Condition had some influence on these outcomes. For the participants in the Collaborative debriefing condition, the only significant correlation was between the Test scores for the two rounds, whereas for the participants who debriefed individually only the correlation of Game scores was significant. It is not clear why condition influenced these correlations.

The absence of a relationship between the Game scores for round 1 and Test 1 was expected. However, the failure to find a relationship between Game scores for round 2 and Test 2 was unexpected, particularly because Van der Meij et al. (2011) did find a significant and positive correlation. The explanation may lie in the length of the gameplay. The present study employed the 15-day variant of Lemonade Tycoon, whereas in the other study the 30-day variant was used. It is fair to assume that when participants have more, or longer opportunities for gameplay, their game score is more likely to correlate to a test score that assesses fundamental game knowledge (compare Delacruz et al., 2010; Wolfe & Roberts, 1983).

It is important to examine game scores and their relations with test scores for several reasons. One, the game scores in Lemonade Tycoon reflect a critical set of variables that play a role in a small-business enterprise. To achieve a high score, players must jointly consider factors such as producing lemonade for the lowest possible price, considering weather conditions in deciding about the production levels, selling lemonade for the highest possible price, and keeping customers happy. In other words, the game score has construct validity and as such should correlate positively with test scores. Two, game scores reflect knowledge-in-action. They are based on authentic, on-the-spot choices and as such show the results from decision-making within the boundaries of gameplay. Three, a game score is an unobtrusive measure of the player's competence. Shute (2011) refers to this phenomenon as "stealth assessment", by which she means an evaluation that does not disrupt flow and blurs the distinction between assessment and learning (see also Hickey & Zuiker, 2012).

8. Conclusion

Education has been interested in game-based learning for a long time. Piaget, Vygotsky and Papert are well-known advocates for the integration of play in education (Games & Squire, 2011). One of the reasons for the educational interest lies in the motivational qualities of games. Games offer their players an opportunity to engage in an experience that they perceive as intrinsically motivating. Games can keep a player motivated for a lengthy period of time. They can induce a state of flow, with players being completely absorbed in the gaming activity (Games & Squire, 2011). It is precisely this quality of games that can also be a drawback for learning. With reflection-in-action being hard to design into the game, we examined the possibility of stimulating reflection after gameplay.

We proposed self-debriefing as the means for doing so, and our research questions focused on the participants' developing knowledge of game strategies and fundamentals (i.e., game scores and test scores). Less attention was given to their developing motivation. That is, we attended to student motivation only at the start of the experiment to check on the equivalence of conditions on this factor. Exploratory analyses indicated that participants' initial motivation did not relate to their test scores.

However, this leaves open the intriguing possibility that the type of debriefing may influence the participants' motivation. For instance, it is conceivable that only participants who engage in individual debriefing will attribute successes in round 2 of gameplay to feelings of control and competence and, as a result, might end with higher self-efficacy beliefs and learning outcomes. To examine this possibility it would be necessary to measure motivation repeatedly during the experiment. Measuring student motivation at different occasions during gameplay is desirable anyway (see Orvis, Orvis, Belanich, & Mullin, 2005; Pavlas, 2010). Because the introduction of games in education originated in part from their motivating character, it seems logical that experimental research should also keep a close watch on the development of student motivation during gameplay.

The recent meta-analysis by Tannenbaum and Cerasoli (2013) contributes at least two valuable insights about research on debriefing. One, the authors identified the presence of a considerable number of experimental studies on debriefing in various disciplines. In doing so they have successfully brought together for the first time a large set of controlled studies in which the effectiveness of debriefing was quantified. Two, the meta-study revealed that the effects of debriefing were quite robust. The authors note that a limitation of their study is that much of the data came from quasi-experimental designs and that causal inferences should be drawn with care.

The meta-analysis by Tannenbaum and Cerasoli (2013) also reveals that there can be a subtle but important difference in how one defines debriefing. Our study was based upon the description by Fanning and Gaba (2007), who characterize debriefing as facilitated or guided reflection in the cycle of experiential learning. Tannenbaum and Cerasoli define debriefing as consisting of four essential elements: (a) active self-learning, (b) developmental intent, (c) specific events, and (d) multiple information sources. The set-up of our study accorded with the first three elements. That is, the participants were not passive recipients but were stimulated to reflect on their experiences. In addition, there was a clear, primary intent to achieve learning. Furthermore, the debriefing was structured around a set of questions that invited participants to reflect on specific events or game episodes rather than about their general performance or competence. The definitional difference from our study occurs in the fourth element. Tannenbaum and Cerasoli define this element as "Includes input from multiple team members or from a focal participant, *and at least one external source, such as an observer or objective data source*" (p. 233, emphasis added). It is in the latter part of this element that their definition varies from ours. Tannenbaum and Cerasoli consider the provision of (external) feedback to be an essential element in debriefing. This is a debatable issue. How one defines feedback, and whether it is considered to be an essential component in debriefing varies considerably across studies (e.g., Boet et al., 2011; Fanning & Gaba, 2007; Kriz, 2010; Petranek, 2000; Rudolf, Simon, Raemer, & Eppich, 2008). When considering the merits of self-debriefing in game-based learning it is therefore important to attend to this issue.

References

Blumberg, F. C., & Randall, J. D. (2013). What do children and adolescents say they do during video game play? *Journal of Applied Developmental Psychology, 34*, 82–88. http://dx.doi.org/10.1016/j.appdev.2012.11.004.

Blumberg, F. C., Rosenthal, S. F., & Randall, J. D. (2008). Impasse-driven learning in the context of video games. *Computers in Human Behavior, 24*, 1530–1541. http://dx.doi.org/10.1016/j.chb.2007.05.010.

Boet, S., Bould, M. D., Bruppacher, H. R., Desjardins, F., Chandra, D. B., & Naik, V. N. (2011). Looking into the mirror: Self-debriefing versus instructor debriefing for simulated crises. *Critical Care Medicine, 39*(6), 1377–1381. http://dx.doi.org/10.1097/CCM.0b013e31820eb8be.

Cantrell, M. A. (2008). The importance of debriefing in clinical simulations. *Clinical Simulation in Nursing, 4*(2), e19–e23. http://dx.doi.org/10.1016/j.ecns.2008.06.006.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Delacruz, G. C., Chung, G. K. W. K., & Baker, E. L. (2010). *Validity evidence for games as assessment environments (CRESST Report 773).National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*. Los Angeles, CA: University of California.

De Vries, B., Van der Meij, H., Boersma, K. T., & Pieters, J. M. (2005). Embedding email in primary schools: Developing a tool for reflection. *Journal of Educational Computing Research, 32*(2), 167–183. http://dx.doi.org/10.2190/AM6M-JACJ-PG51-C8UF.

El-Shamy, S. (2001). *Training games: Everything you need to know about using games to reinforce learning*. Stylus Publishing, LLC.

Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46*(4), 1107–1149. http://dx.doi.org/10.3102/0002831209340235.

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare, 2*(2), 115–125. http://dx.doi.org/10.1097/SIH.0b013e3180315539.

Finley, J. R., & Benjamin, A. S. (2012). Adaptive and qualitative changes in encoding strategy with experience: Evidence from the test-expectancy paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 632–652.

Games, A., & Squire, K. D. (2011). Searching for the fun in learning: A historical perspective on the evolution of educational video games. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 17–46). Charlotte, NC: Information Age Publishing Inc.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation Gaming, 33*(4), 441–467. http://dx.doi.org/10.1177/1046878102238607.

Gegenfurtner, A., Veermans, K., & Vauras, M. (2013). Effects of computer support, collaboration, and time lag on performance self-efficacy and transfer of training: A longitudinal meta-analysis. *Educational Research Review, 8*, 75–89. http://dx.doi.org/10.1016/j.edurev.2012.04.001.

Gijlers, H., & De Jong, T. (in press). Using concept maps to facilitate collaboratuve simulation-based inquiry learning. *The Journal of the Learning Sciences*. doi: 10.1080/10508406.2012.748664.

Gijlers, H., & De Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction, 27*, 239–268. doi: 10.1080/07370000903014352.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534–537. doi:10.1038.

Gum, L., Greenhill, J., & Dix, K. (2011). SimTRACT: A reflective conceptual framework for simulation debriefing. *Journal of Transformation Education, 9*(1), 21–41. http://dx.doi.org/10.1177/1541344611428470.

Hickey, D. T., Kindfield, A. C. H., Horwitz, P., & Christie, M. A. (2003). Integrating curriculum, instruction, assessment, and evaluation in a technology-supported genetics environment. *American Educational Research Journal, 40*, 495–538. http://dx.doi.org/10.3102/00028312040002495.

Hickey, D. T., & Zuiker, S. J. (2012). Multilevel assessment for discourse, understanding and achivement. *The Journal of the Learning Sciences, 21*, 522–582. http://dx.doi.org/10.1080/10508406.2011.652320.

Kirriemuir, J., & McFarlane, A. (2004). Literature review in games and learning. *Nesta Futurelab Series. Report 8*.

Kolb, D. A. (1984). Experiential learning: Experience as the source of learning and development.

Koops, M., & Hoevenaar, M. (2012). Conceptual change during a serious game: Using a lemniscate model to compare strategies in a physics game. *Simulation and Gaming, 1–18*. http://dx.doi.org/10.1177/1046878112459261.

Kriz, W. C. (2010). A systemic-constructivist approach to the facilitation and debriefing of simulations and games. *Simulation and Gaming, 41*(5), 663–680. http://dx.doi.org/10.1177/1046878108319867.

Lederman, L. C. (1984). No-debriefing: A critical reexamination of the postexperience analytic process with implications for Its effective use. *Simulation and Games, 15*(4), 415.

Lederman, L. C. (1992). Debriefing: Toward a systematic assessment of theory and practice. *Simulation Gaming, 23*(2), 145–160. http://dx.doi.org/10.1177/1046878192232003.

Lederman, L. C., & Kato, F. (2005). Debriefing the debriefing process: Toward an expanded look. In D. Crookall & A. Kiyoshi (Eds.), *Simulation and gaming across disciplines and cultures*. Thousand Oaks, CA: Sage.

Leemkuil, H. H., & De Jong, T. (2012). Adaptive advice in learning with a computer-based knowledge management simulation game. *Academy of Management Learning and Education, 11*, 653–665. http://dx.doi.org/10.5465/amle.2010.0141.

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

Lennon, J. L. (2006). Reports & communications: Debriefings of web-based malaria games. *Simulation and Gaming, 37*, 350–356. http://dx.doi.org/10.1177/1046878106291661.

Lennon, J. L. (2010). Debriefing a health-related educational game: A case study. *Simulation and Gaming, 41*(3), 390–399. http://dx.doi.org/10.1177/1046878109332810.

Linn, M. A., Lee, H. S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science, 313*, 1049–1050. http://dx.doi.org/10.1126/science.1131408.

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation-based medical educations research: 2003–2009. *Medical education, 44*(1), 50–63. http://dx.doi.org/10.1111/j.1365-2923.2009.03547.x.

Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., & Fernández-Manjón, B. (2008). Educational game design for online education. *Computers in Human Behavior, 24*(6), 2530–2540. http://dx.doi.org/10.1016/j.chb.2008.03.012.

Neill, M. A., & Wotton, K. (2011). High-fidelity simulation debriefing in nursing education: A literature review. *Clinical Simulation in Nursing, 7*, e161–e168. http://dx.doi.org/10.1016/j.ecns.2011.02.001.

Norman, D. A. (1993). *Things that make us smart. Defending human attributes in the age of the machine*. Reading, MA: Addison-Wesley Publishing Company.

Oertig, M. (2010). Debriefing in Moodle: Written feedback on trust and knowledge sharing in a social dilemma game. *Simulation and Gaming, 41*(3), 374–389. http://dx.doi.org/10.1177/1046878108325982.

Orvis, K. A., Orvis, K. L., Belanich, J., & Mullin, L. N. (2005). *The influence of trainee gaming experience and computer self efficacy on learner outcomes of video game based learning environments (Vol. TechnicalReport, No.1164)*. Arlington, VA: U.S. Army Research Institute for the Behavioural and Social Sciences.

Pavlas, D. (2010). *A model of flow and play in game-based learning: The impact of game characteristics, player traits, and player states*. (PhD), University of Central Florida, Orlando, Florida.

Petranek, C. F. (2000). Written debriefing: The next vital step in learning with simulations. *Simulation Gaming, 31*(1), 108–118. http://dx.doi.org/10.1177/104687810003100111.

Petranek, C. F., Corey, S., & Black, R. (1992). Three levels of learning in simulations: Participating, debriefing, and journal writing. *Simulation Gaming, 23*(2), 174–185. http://dx.doi.org/10.1177/1046878192232005.

Qudrat-Ullah, H. (2007). Debriefing can reduce misperceptions of feedback: The case of renewable resource management. *Simulation Gaming, 38*(3), 382–397. http://dx.doi.org/10.1177/1046878107300669.

Reed, S. J. (2012). Debriefing experience scale: Development of a tool to evaluate the student learning experience in debriefing. *Clinical Simulation in Nursing, 8*, e211–e217. http://dx.doi.org/10.1016/j.ecns.2011.11.002.

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica, 47*(2), 57–66. http://dx.doi.org/10.1026//0012-1924.47.2.57.

Roediger, H. L., Agarwal, P. K., McDaniel, M., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied, 17*(4), 382–395. http://dx.doi.org/10.1037/a0026252.

Rudolf, J. W., Simon, R., Raemer, D. B., & Eppich, W. J. (2008). Debriefing as formative assessment: Closing performance gaps in medical education. *Academic Emergency Medicine, 15*, 1010–1016. http://dx.doi.org/10.1111/j.1553-2712.2008.00248.x.

Sawyer, T., Sierocka-Castaneda, A., Chan, D., Berg, B., Lustik, M., & Thompson, M. (2012). The effectiveness of video-assisted debriefing versus oral debriefing alone at improving neonatal resuscitation performance. *Simulation in Healthcare, 7*(4), 213–221. http://dx.doi.org/10.1097/SIH.0b013e3182578eae.

Schoepfle, G. M., & Werner, O. (1999). Ethnographic debriefing. *Field Methods, 11*(2), 158–165. http://dx.doi.org/10.1177/1525822x9901100206.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Aldershot, England: Arena, Ashgate Publishing.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Charlotte, NY: Information Age Publishing Inc.

Sims, R. R. (2002). Debriefing experiential learning exercises in ethics education. *Teaching Business Ethics, 6*(2), 179–197. http://dx.doi.org/10.1023/A:1015237928724.

Steinwachs, B. (1992). How to facilitate a debriefing. *Simulation Gaming, 23*(2), 186–195. http://dx.doi.org/10.1177/1046878192232006.

Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors, 55*(1), 231–245. http://dx.doi.org/10.1177/0018720812448394.

Thiagarajan, S. (1992). Using games for debriefing. *Simulation Gaming, 23*(2), 161–173. http://dx.doi.org/10.1177/1046878192232004.

Tobias, S., & Fletcher, J. D. (2011, October 19–21). Learning from computer games: A research review. *Paper presented at the Serious Games: The challenge: ITEC/CIP/T 2011: Joint Conference of the interdisciplinary research group of technology, education, communication, and the scientific network on critical and flexible thinking, Ghent, Belgium*.

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. 127–221). Charlotte, NY: Information Age Publishing Inc.

Van der Meij, H., Albers, E., & Leemkuil, H. H. (2011). Learning from games: Does collaboration help? *British Journal of Educational Technology, 42*(2), 655–664. http://dx.doi.org/10.1111/j.1467-8535.20101.01067.x.

van Ments, M. (1983). *The effective use of role-play: A handbook for teachers and trainers*. London, UK: Kogan Page.

Vollmeyer, R., & Rheinberg, F. (1999). Motivation and metacognition when learning a complex system. *European Journal of Psychology of Education, 14*(4), 541–554. http://dx.doi.org/10.1007/bf03172978.

Warrick, D. D., Hunsaker, P. L., Cook, C. W., & Altman, S. (1979). Debriefing experiential learning exercises. *Journal of Experiential Learning and Simulatlon, 1*, 91–100.

Wolfe, J., & Roberts, C. R. (1983). A longitudinal study of the external validity of a business management game. *Developments in Business Simulation & Experiential Learning, 10*, 9–12.

Wotton, K., Davis, J., Button, D., & Kelton, M. (2010). Third-year undergraduate nursing students' perceptions of high-fidelity simulation. *Journal of Nursing Education, 49*(11), 632–639.