We don't need no stinkin' badges: The impact of reward features and feeling rewarded in educational games

Brian McKernan a,⇑, Rosa Mikeal Martey b, Jennifer Stromer-Galley c, Kate Kenski d, Benjamin A. Clegg e, James E. Folkestad f, Matthew G. Rhodes e, Adrienne Shaw g, Emilie T. Saulnier h, Tomek Strzalkowski a

a University at Albany, SUNY, The Institute for Informatics, Logics, and Security Studies, Room 262, Social Science Building, 1400 Washington Avenue, Albany, NY 12222, USA
b Colorado State University, Department of Journalism and Technical Communication, Clark C-218, CD 1785, Fort Collins, CO 80523, USA
c Syracuse University, School of Information Studies, 220 Hinds Hall, Syracuse, NY 13244, USA
d University of Arizona, Department of Communication and School of Government and Public Policy, 1103 E. University Blvd., Communication Building #25, Room 211, Tucson, AZ 85721, USA
e Colorado State University, School of Psychology, Fort Collins, CO 80525, USA
f Colorado State University, School of Education, 244 Education Building, Fort Collins, CO 80525, USA
g Temple University, Department of Media Studies and Production, 2020 N. 13th St., Annenberg Hall, Room 203a, Philadelphia, PA 19122, USA
h 1st Playable Productions, 5 Third Street, Suite 300, Troy, NY 12180, USA

Article info

Article history:

Keywords:
Educational games
Video games
Rewards
Learning
Cognitive biases

ABSTRACT

Drawing from recent research on the ability of video games to satisfy psychological needs, this paper identifies how the presence of rewards influences learning complex concepts and tasks using an educational video game. We designed and developed two 60-min educational games with and without a range of reward features and examined learning outcomes among 242 participants in university laboratories. Although both games improved learning, analyses suggest that the quantity of in-game rewards did not have an impact on biased behavior avoidance or knowledge about biases. To further illuminate these findings, we examined perceptions of feeling rewarded and found that those who felt more rewarded had more favorable views of the gameplay experience, but they did not demonstrate different learning outcomes.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Over the last two decades, prominent figures from a variety of fields have explored how features that account for video games’ massive popularity may also make them powerful educational tools (Gee, 2007; McGonigal, 2011; Squire, 2011). This work has consistently identified in-game rewards as an important source of video games’ overarching appeal (Wang & Sun, 2011; Yee, 2006), leading scholars to suggest that the reward systems in popular commercial games should also be used in educational games (Garris, Ahlers, & Driskell, 2002; Kelle, Klemke, & Specht, 2013). Indeed, rewards such as points, badges, or achievements have become key components of proposals for “gamifying” education (Kapp, 2012; Lee & Hammer, 2011). Only limited research has explored the impact of rewards on learning, however, and few authors distinguish between perceptions of feeling rewarded and in-game reward features themselves.

Drawing from recent research on the ability of video games to satisfy psychological needs (Przybylski, Rigby, & Ryan, 2010; Przybylski, Ryan, & Rigby, 2009; Ryan, Rigby, & Przybylski, 2006), this study examined how the presence of rewards influences learning complex concepts and tasks using an educational video game. We designed and developed two versions of a 60-min educational game, one with and one without a range of reward features. We examined learning outcomes among 242 participants in university laboratories. Analyses suggest that although both versions of the game had significant and strong effects on participants’ behavior and knowledge, the quantity of in-game rewards did not have an impact on these outcomes. To further illuminate these findings, we examined perceptions of feeling rewarded and found that there was no difference in how rewarded participants felt when in-game reward features were and were not present. Rewards affected how favorably participants viewed the gameplay experience, but they had no impact on learning outcomes. We examine the implications of these findings for educational game design in the concluding sections.
2. Literature review

2.1. Theories of rewards and feeling rewarded in game-based learning

The scholarly literature on learning identifies two separate, but interrelated, reasons why educational games should include reward systems. First, scholars suggest that rewards may strengthen learning outcomes by motivating individuals to pursue challenging tasks or goals that they otherwise would be less interested in or attempt less diligently (Cameron, Pierce, Banko, & Gear, 2005; Mesch, Johnson, & Johnson, 1988; Pierce, Cameron, Banko, & So, 2012; Skinner, Williams, & Neddennriep, 2004). According to this logic, in-game features that players find appealing such as unlocking new challenges, earning accessories to customize their avatar, or receiving trophies may motivate players to both continue to play the game and to play the game more carefully than they otherwise would.

Second, many game-based learning scholars treat certain in-game rewards as providing a valuable form of performance feedback. For example, Garris et al. (2002) hold that game score-keeping may motivate players to replay the game in order to improve their performance. Similarly, Kelle et al. (2013) suggested that point systems allow players to gauge their skill level and thus encourage them to ameliorate deficiencies. McGonigal (2011) identified performance feedback systems as a defining feature of all games and suggested that it provides players with a sense of progress or accomplishment that motivates them to pursue the game’s ultimate goal.

Overall, many scholars consider some reward mechanisms to be vital components to game-based learning. Kelle et al. (2013) examined this relationship and found that rewards in the form of a score system strengthen learning outcomes only when combined with a time limit. Much of the other research on this topic emphasizes the importance of rewards in educational games without empirically examining the particular impact rewards have on learning outcomes (Huang, Huang, & Tschopp, 2010; Wang & Sun, 2011). More generally, Garris et al. (2002) warned that without more rigorous empirical and scientific testing, the field runs the risk of “designing instructional games that neither instruct nor engage the learner” (p. 442). Bedwell, Pavlas, Heyne, Lazara, and Salas (2012) and Guillén-Nieto and Aleson-Carbonell (2012) issued similar warnings ten years after Garris et al.’s original critique. In this work, we attempt to remedy this gap in the literature by testing whether or not rewards influence learning outcomes in educational games.

Recent research has drawn from the basic psychological needs literature to help explain video games’ considerable popularity (Przybylski et al., 2009, 2010; Ryan et al., 2006). The basic psychological needs literature posits that individuals are intrinsically motivated to pursue activities that satisfy their basic psychological needs (see Deci & Ryan, 2000). Accordingly, video games may provide players with a certain degree of freedom in setting goals or selecting strategies, thus satisfying players’ psychological need for autonomy (Przybylski et al., 2010; Ryan et al., 2006). Successful games may satisfy players’ need for competency by including challenges that are optimized to players’ abilities without being unfair (Ryan et al., 2006). For many game designers and scholars, providing players with a meaningful sense of control over their actions and challenges that are progressively difficult without being unfair are fundamental principles of quality game design (Fullerton, 2008). Consequently, these perceptions of autonomy, challenge, and control may allow a game to feel rewarding to players even when features game designers identify as in-game rewards are not present or are minimal.

2.2. Games for teaching cognitive biases

Scholars have suggested that games may be particularly effective for training behaviors relevant to complex conceptual tasks such as decision-making due to their tendency to generate Intrinsic Motivation among players (Garris et al., 2002; Gee, 2007; McGonigal, 2011; Ricci, Salas, & Cannon-Bowers, 1996; Salen & Zimmerman, 2004). Individuals are said to be intrinsically motivated to engage in a task or activity when they find the activity itself to be pleasurable or enjoyable (Deci, 1971; Deci & Ryan, 1985). Many scholars consider the majority of players to be intrinsically motivated to play non-educational games given that players do so voluntarily and usually receive no financial rewards in return. Scholars and developers in this area endeavor to make complex and challenging subject matter more enjoyable for students by adopting mechanics and features of popular commercial games into educational games.

The process of learning about cognitive biases and the necessary steps to avoid them can be considered a complex task. Cognitive biases are systematic errors in judgments that are often the result of heuristics used in decision-making, and they are especially common when people examine large amounts of information or when they have to make quick judgments (Larrick, 2004; Tversky & Kahneman, 1974). Cognitive biases are especially pervasive and resistant to change because they are often unconscious (Kahneman, 2011). Various techniques for reducing biased behavior have been explored in experimental settings, although many have had only limited success (Dale, Kehoe, & Spivey, 2007; Sanna, Schwarz, & Small, 2002; Simmons, LeBoeuf, & Nelson, 2010). Psychological research has examined how these biases manifest and ways to overcome them, but little research has addressed the use of games for this purpose, with the exception of projects associated with the present study (Shaw et al., 2013; Martey et al., 2014).

3. Our game

The game designed for this study trains players to change their decision-making behavior and increase their knowledge of three specific cognitive biases. Anchoring bias occurs when an individual is influenced by a number or topic (an anchor) before making an assessment and results in estimates that resemble the anchor too closely (Tversky & Kahneman, 1974; Wilson, Houston, Etling, & Brekke, 1996). Projection bias is defined as assuming others share one’s own values, preferences, skills, or habits (Loewenstein, O’Donoghue, & Rabin, 2002). Representativeness bias is a term that refers to a set of errors that ignore or miscalculate the actual likelihood of events, such as drawing conclusions based on a limited (non-representative) or misleading set of examples (Ajzen, 1977; Kahneman & Tversky, 1973; Micallef, Dragicevic, & Fekete, 2012).

To examine the impact of in-game rewards on players’ ability to recognize and avoid these three cognitive biases, we built a 60-min educational game with a professional game company. The game is based on Transformative Learning Theory (Mezirow, 1990), a cognitive learning theory oriented around behavior change through task-oriented problem solving and understanding others’ values, ideas, and decisions. Further, the game was structured with learning taxonomy by Bloom (1956; see also Anderson & Sosnai, 1994) to train cognitive, affective, and psychomotor skills. Two versions of the game were developed for this experiment: one with frequent in-game rewards and one with very few of these features.

Games scholars generally define a reward as any game item or feature that reinforces particular in-game behaviors (Fullerton, 2008; Hsu, Wen, & Wu, 2009; King, Delfabbro, & Griffiths, 2010;
According to these scholars, this form of behavior reinforcement, or “pat on the head” (Bateman & Boon, 2005, p. 46), may motivate players to continue playing the game. That is, rewards provide players with incentives to keep playing. Based on this understanding of rewards, we conceptualized rewards for this study as incentives in the forms of sound, visuals, items, and text that mark player achievements.

The game, *Cycles Carnivale*, is a point-and-click, two-dimensional, Flash-based puzzle game that uses first- and third-person views set in an alien carnival in outer space where the player's character has crash-landed. In order to purchase the parts they need to fix their ship, players must collect currency by winning a series of classic carnival games such as spinning a wheel or knocking down a tower of bottles. Winning requires players to identify and avoid the relevant cognitive bias that can influence behavior in each challenge.

An alien companion helps the player win by providing additional explanations of the biases and how to avoid them, performance feedback, and guidance. Text, animation, sound, cut scenes, and an interactive “heads-up display” (HUD) provide feedback throughout the game. Short quizzes allow players to test their knowledge of the biases between levels. No voice-over audio was used. Teaching content structure leverages the spacing effect (Bahrick & Hall, 2005; Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006) by requiring players to learn content and then return to it later in the game. Player progression and teaching content are the same in both versions (see Fig. 1).

Both versions of our game use features that the basic psychological needs literature suggests players may find rewarding. Although our game is predominantly a linear experience, players can select between two challenges to complete first in some places, which may contribute to a sense of autonomy. Additionally, the challenges in our game become progressively more difficult in terms of the complexity of the learning material and the tasks players must complete. Both versions of our game provide clear indications of player success. The combination of both these features may provide players with a significant degree of competency.

We removed all of the rewards that we could without varying the teaching content for the low rewards version of our game. Table 1 lists the rewards features that we used in the high rewards version and how they were adjusted in the low rewards version of our game. For example, players earned parts that they could use to repair their ship in order to escape the *Cycles Carnivale* and return home. These rewards were removed in the low rewards gameplay and all ship parts were given to the player at the end of the game.

The tokens and prizes players receive at the end of each carnival booth in the high reward version provide players with a sense of the quality of their performance in that challenge (see Fig. 2). Players receive different levels of each depending upon how well they avoided biased behavior and how many rounds they won. These tokens and prizes are needed in the game’s narrative to “pay” for the parts players need to repair their ship and contribute to a sense of game progression. Accordingly, these features serve to incentivize players to engage in the desired behaviors by connecting game progression to learning objectives (Chen, Liao, Cheng, Yeh, & Chan, 2012; Hsu et al., 2009; King et al., 2010; Wang & Sun, 2011).

The badges players earn in the high reward version for engaging in more challenging or demanding in-game tasks also function as a reward. For example, players need to label correctly one of two creatures in a challenge to complete it successfully. However, players who correctly label both creatures receive the “Label Maker” badge. Wang and Sun (2011) classified such badges as a glory rewards system. Although they possess no narrative significance and are not required to progress in the game, badges provide players with a sense of achievement and thus encourage players to think carefully about their in-game behavior and play the game in challenging or more difficult ways (Wang & Sun, 2011).

The performance score players receive at the end of the high rewards game also functions as a glory rewards system. Players are told at the start of the high rewards game that their efforts will receive a final score when they finish. The end of the game includes a screen summarizing performance, including tokens earned, successes on quiz questions, and a letter grade. Knowing they will receive a final score may encourage players to play the game more carefully than they otherwise would.

The high rewards version of our game also includes audio and visual feedback that functions as a type of reward. The SmartScanner, an in-game instrument the player uses to help avoid biased behaviors, provides feedback on the player’s performance. The combination of these types of rewards may provide players with a significant degree of competency.

Table 1 lists the rewards features that we used in the high rewards version and how they were adjusted in the low rewards version of our game.

---

**Table 1**

<table>
<thead>
<tr>
<th>Feature</th>
<th>High Rewards Version</th>
<th>Low Rewards Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parts to repair ship</td>
<td>Earned</td>
<td>Given at end of game</td>
</tr>
<tr>
<td>Tokens</td>
<td>Earned</td>
<td>Given at end of game</td>
</tr>
<tr>
<td>Prizes</td>
<td>Earned</td>
<td>Given at end of game</td>
</tr>
<tr>
<td>Badges</td>
<td>Earnings</td>
<td>None</td>
</tr>
<tr>
<td>Performance score</td>
<td>Earned</td>
<td>None</td>
</tr>
</tbody>
</table>

---

Fig. 1. Map of game challenges for both versions of *Cycles Carnivale*. 

---

*Salen & Zimmerman, 2004; Wood, Gupta, Derevensky, & Griffiths, 2004.*
behavior, provides a ding sound and flashes green when players make a correct choice. Challenges and quiz questions also include victory animations and sounds. These features provide players with pleasurable audiovisual experiences, which may encourage players to continue their efforts (Fisher & Griffiths, 1995; Wang & Sun, 2011).

The detailed, animated cut scenes that bookend some levels in the high rewards game can also be considered as a type of reward. Such scenes may motivate players because they are visually appealing and indicate progress towards completing the game (Clanton, 1998; King et al., 2010; Wang & Sun, 2011).

The high rewards version of our game also includes encouraging text based on the player’s performance. For example, players receive praise such as “Good job!” or “Excellent work!” after making the correct choice in a challenge. Many reward scholars have found that positive verbal feedback motivates individuals to pursue the task at hand by strengthening their sense of competence and self-determination (Deci, Koestner, & Ryan, 1999; Pittman, Davey, Alafat, Wetherill, & Kramer, 1980; Ryan, 1982; Sheldon & Filak, 2008).

Although much of the literature on game rewards focuses on intrinsic versus extrinsic rewards, this project does not distinguish between these. It focuses on characteristics internal to games that can be assessed independently of extra-game structures such as progress and scoring, that are considered by some to be extrinsic rewards because they can be shared with other players (Huang et al., 2010; Wang & Sun, 2011). The current project, however, focuses on extensive versus sparse in-game rewards, not differences among types.

4. Hypotheses and research question

Many game scholars suggest that a greater quantity of rewards leads to greater learning outcomes (Cameron et al., 2005; Mesch et al., 1988; Pierce et al., 2012; Skinner et al., 2004). A game with varied and extensive in-game reward features, therefore, should demonstrate greater improvements in learning than a game without such features. We therefore hypothesize that:

H1. The high rewards version of the Cycles Carnival game will show greater improvement in avoiding biased behavior than will the low rewards version.

H2. The high rewards version of the Cycles Carnival game will show greater improvements in bias knowledge than will those who play a game with very few rewards.

In addition, given the widely different ways that the literature defines the specific mechanics and forms of rewards in games and the lack of clarity on the relationship between in-game rewards and participants’ perceptions of feeling rewarded during play, the current study explores the following research question:
RQ1. What is the relationship between in-game reward features and participants’ sense of feeling rewarded during play?

5. Methods

To examine whether frequent in-game reward features have an impact on bias behavior and knowledge, we compared two groups, low rewards versus high rewards, to examine learning outcomes. Although the original experimental design used a comparison video, this study focuses on differences between game conditions. Therefore, learning outcomes for those assigned to the video condition are not examined here.

5.1. Study procedures

Participants recruited from college classes and psychology pools at three universities in the United States were randomly assigned to conditions: digital game with high rewards and digital game with low rewards. Participants came to a computer lab where they answered online questionnaires before and after the experimental stimuli were given. The questionnaires measured knowledge and mitigation of three cognitive biases, along with several demographic and personality variables. The post-test also included measures of engagement, usability, a manipulation check, and a debriefing statement that included the purpose of the study, additional resources, and contact information for human subjects review boards. Sessions lasted an average of 2.5 h. Participants were asked to move onto the final survey if they took longer than 75 min to play the game. Thirteen people were excluded from analysis due to incomplete post-session surveys. 72 people were excluded from analysis for failing the attention check question included in our questionnaires. The attention check question states: “Public opinion polling suggest American approval of Congress is at an historic low. Nevertheless, we would like you to choose prefer not to respond to this question to show that you are reading carefully. What is your opinion of Congress?” Participants who selected any choice besides “prefer not to respond” were identified as inattentive and thus removed from analysis.

5.2. Study participants

The sample consisted of 242 participants randomly assigned to a high rewards game (49%) or a low rewards game (51%). Participants were 65% female, 35% male, and were an average age of 20 years old. Most (60%) were in their first or second year of school, and the majority reported their race as white (58%), 5% as black, 11% as Hispanic/Latino, and 26% as mixed and/or other races. Only 7% were psychology majors.

The majority agreed or strongly agreed they like playing video games (53%), 35% said they have a lot of experience playing computer games, and 86% reported their level of computer experience was intermediate or better. In addition, 16% said they considered themselves to be “gamers.” On average, participants had some experience with educational games, with 36% agreeing they had “a lot of experience” playing them.

5.3. Cognitive bias measures

This study used two measures of cognitive bias: bias avoidance behavior and knowledge of biases. The scales used address three specific biases: anchoring bias (AN), projection bias (PR), and representativeness bias (RB). They are based on the work of Tversky and Kahneman (1974) and current scholarship in the field (Loewenstein, O’Donoghue, & Rabin, 2002; Micallef et al., 2012; Wilson et al., 1996) and were developed by a team of researchers across six institutions performing similar studies of educational games that teach about these biases (see Bush, Symborski, Martey, Saulnier, & Veinott, 2014). The scales were developed via extensively testing using Amazon’s Mechanical Turk with over 1500 participants and in university laboratories with approximately 300 participants.

5.3.1. Bias knowledge

The bias knowledge scale consists of 19 multiple choice questions similar in content and form to university-level exam questions testing definitions and applications of the three biases. The items describe a scenario and ask participants to identify which of four possible biases was evident. For example, the following was used to assess knowledge of RB: “Sally is drawing cards out of a deck. She draws a card, puts the card back, and shuffles the deck. The first 4 cards she has drawn have all been black cards. She assumes the next card will be red because red is ‘due.’ What bias is Sally demonstrating?” Answers were scored as correct (1) or incorrect (0) and averaged across items for each bias. Average reliability using Cronbach’s alpha was .54 for the pre-session scale and .81 for the post-session scale.

5.3.2. Bias avoidance behavior

To assess whether or not participants displayed biased behavior, a series of 67 questions were developed based on current laboratory research on bias manifestations in the psychological literature. The behavior scale has three subscales. Anchoring (18 items) was assessed by asking participants to first consider a number or concept, then to make an estimate about a trivial fact. Numerical estimates were open text responses, and conceptual estimates were 9-point Likert scales. Unbiased answers were those that were farthest away from the anchor provided. Projection (23 items) used 7-point Likert scales and open text responses. Items first asked participants to rate their agreement with a statement about their own characteristics, then to rate the extent to which they believe others have that characteristic. Unbiased answers were those that demonstrated the greatest difference between self- and other-assessments. Representativeness bias (26 items) was assessed using multiple choice questions that first presented a scenario, and then asked participants to make assessments about the likelihood of specific events related to that scenario. Unbiased responses were those that reflected the correct probability of the event. Average reliability for these measures using Cronbach’s alpha was .69 for the pre-session scale and .86 for the post-session scale.

Three unique versions of the bias knowledge and the bias avoidance behavior scales were developed and randomly assigned to ensure participants answered different questions on the pre- and post-tests. For each scale, items were scored from 0 to 1 where 0 is completely biased and 1 is completely unbiased. Items were then averaged to generate an overall bias knowledge score and an overall bias avoidance behavior score.

5.4. Engagement and game experience measures

To assess how rewarded participants felt playing the game, three questions in the post-survey were used. These were: “The game gave me a lot of rewards for getting things right,” “I received praise and recognition from the game for doing the right thing,” and “Getting things right in the game felt rewarding.” Other measures included engagement scales for Positive Affect, Negative Affect, Workload, Fatigue, and Intrinsic Motivation. These were assessed by asking participants to rate on 7-point Likert

1 Full items and additional detail on the bias measure scales are not included here for space considerations but are available upon request.
Table 2

Engagement and game experience reliabilities.

<table>
<thead>
<tr>
<th>Scale</th>
<th># of items</th>
<th>Reliability α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental immersion</td>
<td>7</td>
<td>.860</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>5</td>
<td>.899</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>5</td>
<td>.816</td>
</tr>
<tr>
<td>Workload</td>
<td>2</td>
<td>.700 (r = .536)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>3</td>
<td>.858</td>
</tr>
<tr>
<td>Intrinsic Motivation</td>
<td>4</td>
<td>.796</td>
</tr>
<tr>
<td>Technical usability</td>
<td>11</td>
<td>.865</td>
</tr>
<tr>
<td>Likeability</td>
<td>4</td>
<td>.826</td>
</tr>
</tbody>
</table>

Table 3

Bias measure descriptive statistics by condition.

<table>
<thead>
<tr>
<th></th>
<th>Low rewards</th>
<th>High rewards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior PRE</td>
<td>.44 (.07)</td>
<td>.58 (.11)</td>
</tr>
<tr>
<td>Behavior POST</td>
<td>.58 (.11)</td>
<td>.24 (.13)</td>
</tr>
<tr>
<td>Knowledge PRE</td>
<td>.26 (.14)</td>
<td>.54 (.22)</td>
</tr>
<tr>
<td>Knowledge POST</td>
<td>.54 (.23)</td>
<td></td>
</tr>
</tbody>
</table>

6. Results

6.1. Bias avoidance and bias knowledge

To examine whether a game with frequent and extensive reward features trains better than one with few reward features (H1 and H2), we conducted a MANOVA on each bias scale (behavior and knowledge) as dependent variables. Table 3 shows means with standard deviations in parentheses for the pre- and post-session bias behavior measures.

A repeated measures MANOVA showed no difference between the game conditions for improved bias avoidance behavior; H1 was not supported. Analysis of pre-post measures by condition found a main effect of pre-post (Wilks’ λ = .36, F(1,240) = 420.52, p < .0005), no main effect of condition (F(1,240) < 1, MSE = 0.012), and no interaction between pre-post and condition (Wilks’ λ = .99, F(1,240) = 2.07, p > .05). Examining the post-test performance using a planned comparison showed no difference between game conditions (t(240) = 0.28, p > .05). Cohen’s d effect size for the change in bias exhibited comparing pre- to post-test was substantial for both types of training, 1.5 for the low rewards game, and 1.2 for high rewards.

Analyses showed no differences between game conditions for improved bias knowledge, lending no support to H2. There was a main effect of pre-post (Wilks’ λ = .34, F(1,240) = 471.97, p < .0005), no main effect of condition (F(1,240) < 1, MSE = 0.045), and no interaction between pre-post and condition (Wilks’ λ = .99, F(1,240) < 1). Examining the post-test performance using a planned comparison showed no difference between game conditions (t(240) = 0.07, p > .05). Cohen’s d effect size for the change in knowledge seen in pre- to post-test differences were again very large, 1.5 for the low rewards game, and 1.3 for high rewards.

6.2. Feeling rewarded

In order to examine the relationship between feeling rewarded and learning outcomes (RQ1), we also examined the scales for feeling rewarded, engagement, usability, likeability, and learning perceptions for those study participants who played a version of the game. Table 4 shows descriptive statistics for these scales.

First, we examined condition differences in the felt rewarded items to determine if, as assumed, the high rewards game felt more rewarding to players. Responses to these items were examined with correlations and independent sample t-tests. Results show that two of the three rewards items were significantly different across conditions, as shown in Table 5. There was no difference in how rewarding the games felt across conditions.

We then examined Pearson correlations among game participants to examine associations between engagement, usability scales and the rewards items. We also examined correlations with a dummy variable for condition (low rewards = 1), shown in Table 6. To examine relationships with learning changes, we also calculated the post-session minus pre-session levels of bias behavior (mean = .13, SD = .10) and of knowledge (mean = .29, SD = .21). Positive values on these measures indicate less biased behavior and more knowledge, respectively. Because there were no significant differences by game condition, correlations for bias measures are across both games.

Analysis revealed that feeling rewarded is related to all the engagement scales, usability and likeability. Intrinsic Motivation and the rewards variables had the most consistently strong relationships, and the scales overall showed the most consistently strong relationships with the question “felt rewarding.” Feeling rewarded, however, is not related to bias improvements. Game condition was related to two of the rewards conditions, but not to any of the scales or bias measures.

7. Discussion

This study examined whether a digital game with extensive and varied rewards could train people to improve their behavior and knowledge of biases better than a digital game with few reward features. Although both games improved learning, comparing the two game conditions did not reveal any differences in bias avoidance behavior or bias knowledge; H1 and H2 were not supported.
This suggests that learning is not affected by the quantity of reward features in an educational game.

When examining how perceptions of feeling rewarded, receiving rewards, and receiving praise influenced these results, we found that feeling rewarded was not related to the quantity of reward features in the game. T-tests on individual items show that the game conditions differed in “lots of rewards” and “receiving praise and recognition” but not in “feeling rewarded.” In short, players did not perceive the version of our game with almost no reward features to be any less rewarding than the version with many reward features. This may be because players felt rewarded more by core features of the game such as progression and successful challenge completion than the in-game reward features such as the badges or visual fanfare themselves. This is supported by the strong relationships between feeling rewarded and the set of scales assessing perceptions of the gameplay experience (engagement, fatigue, usability). These findings suggest that, although learning is not affected by feeling rewarded nor by reward features, a game’s appeal is higher when players feel rewarded. Although the direction of this influence is not identifiable from these data, analyses suggest that players feel more rewarded when they like a game, and understand the controls. Overall, these findings suggest that fulfilling psychological needs may not always be related to learning outcomes per se, even if fulfilling those needs influences the appeal and motivation for educational games (Przybylski et al., 2009, 2010; Ryan et al., 2006).

8. Conclusion

We found that there were no differences in learning between playing an educational game with and without a large number of reward features. These findings suggest that the quantity of rewards does not have an impact on learning outcomes. The strong relationships between feeling rewarded and game assessments such as engagement and usability emphasize the importance of feeling rewarded as an assessment of a game’s appeal. In general, players have better responses to the game when they feel rewarded while playing regardless of how many reward features are present.

This study has some limitations. Most importantly, the sample is not representative of a national sample. The college students recruited, however, were located in different regions of the country. Similarly, as a lab-based study, players may have been less motivated to learn about biases than they might have been in other settings such as a classroom, where students may be inherently interested in a subject and thus more motivated. Another key limitation is that this study used only a single game, with a specific genre and style. Future research should examine these relationships among a range of game types and populations in additional settings.

Overall, this study shows that rewards are important in educational games, but not necessarily in the way scholars have previously believed. Our results suggest that rather than focusing on adding points, badges, or other in-game rewards in an effort to make a game more appealing to potential players, educational game designers should focus on core game mechanics that players may find satisfying, such as intuitive controls and progressively difficult but fair challenges.

Acknowledgements

We would like to thank Elizabeth McLaren, Sarah Taylor, and the entire CYCLES team for their invaluable feedback throughout this research project. This work was supported by the Intelligence Advanced Research Projects Activity (IARPA) via the Air Force Research Laboratory Contract Number FA8650-11-C-7176. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, AFRL, or the U.S. Government.
References


