The Impact of Choice and Feedback on Learning, Motivation, and Performance in an Educational Video Game

Charles K. Kinzer, Daniel L. Hoffman, Selen Türkay, Nilgun Günbaş,
Pantiphar Chantes, Tatyana Dvorkin, and Apichai Chaiwinij
Teachers College Columbia University, 525 West 120 St., New York, NY 10027
Email: kinzer@tc.columbia.edu

Abstract: This study modified an existing educational video game by varying a learning mechanic and an assessment mechanic. The result was multiple versions of the same game with identical game mechanics but different learning and assessment variables. The impact of these variables was examined to determine their impact on three dependent variables: learning, motivation, and in-game performance. One hundred thirty-eight (N=138) sixth grade students were randomly assigned to play one of the four versions of the game. After thirty minutes of play, results suggest that providing players with a choice of non-player character from whom to receive feedback results in significantly higher learning outcomes and desire to continue playing compared to a non-choice condition. Comparisons between informative and elaborative feedback did not influence student any of the dependent variables. The theoretical and practical implications of these findings are discussed within the context of educational game design and research.

Introduction

Educational video games by definition, regardless of genre or quality, must contain a number of learning-related variables. Unfortunately, many games intended to educate, currently lack coherent connections to theories of learning or underlying bodies of research (Shaffer, Squire, Halverson, & Gee, 2005). This gap between theory and practice has resulted in video games that may be enjoyable, but do not support academic learning (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005). To help reverse this phenomenon, the Games for Learning Institute (G4LI) has urged educational game designers to distinguish between and consider in their designs three categories of mechanics: game, learning, and assessment (Plass et al., in press; Plass, Homer, Kinzer, Frye, & Perlin, 2011a).

Literature Review

Game, Learning, & Assessment Mechanics

When carefully designed and implemented, game, learning, and assessment mechanics can work in concert to create a game experience that is effective both as a play experience and as a learning/instructional experience.

Perhaps the most familiar concept to game designers is that of the game mechanic, since much has been written on the topic (see Bjork & Holopainen, 2005; Fullerton, Swain, & Hoffman, 2008; Salen & Zimmerman, 2004). For the purposes of this study, game mechanics describe the essential game play activity and are distinct from learning mechanics and assessment mechanics. Well known game mechanics include leveling, resource management, and turn taking. In contrast, learning mechanics according to Plass et al. (in press) are grounded in learning theory and describe specialized activities that have learning as their primary objective. Learning mechanics are theoretical in nature and must be instantiated in the game space through the use of game mechanics. For example, the well-documented instructional practice of peer-tutoring (see Topping, 1988), might be incorporated into a game by requiring players to generate authentic problems to be solved by other players. Similarly, assessment mechanics are grounded in test theory and are specialized activities that have assessment as their primary objective (Plass et al., in press). An example, drawing on adaptive testing theory, is a game that progressively challenges players by adaptively adjusting and setting the difficulty level based on player performance.

Choice as a Learning Mechanic

The learning mechanic targeted in this study was choice. Research has shown that providing students with choices can increase self-efficacy, motivation and learning. The motivational aspects of choice have been part of many motivational frameworks, such as Eccles & Wigfield's (1995) expectancy-value theory, Bandura's (1997) social cognitive theory, and cognitive dissonance theory (Collins &

Hoyt, 1972). For example, the concept of self-efficacy is the belief in one's capabilities "to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). Recently, Leotti, Iyengar, & Ochsner (2010) argued that opportunities to exercise control may be necessary to foster self-efficacy beliefs. They further assert that "each choice—no matter how small—reinforces the perception of control and self-efficacy, and removing choice likely undermines this adaptive belief" (p. 4).

Several researchers have examined the relationship between choice and learning. Zuckerman, Porac, Lathin, Smith, & Deci (1978) assigned 40 students each to a choice or no-choice puzzle-solving condition. Individuals in the choice condition reported a greater feeling of control, greater willingness to participate in another solving session, and spent significantly more time on similar puzzles in a free-choice period at the end of the experiment. Cordova & Lepper (1996) investigated the effects of choice on elementary children's learning within a computer game. Subjects made choices on features that are not relevant to the pedagogical aspect of the game. The results showed that even minimal choices produced a significant increase, not only in the participants' motivation, but also in the depth of their engagement in learning, as evidenced by a preference for more challenging versions of the game, the greater use of complex operations, and an emphasis on strategic play. Moreover, the amount they learned increased, as did their perceived competence and level of aspiration.

This brief overview demonstrates that choice in a game environment might be leveraged to influence learning, motivation, and in-game performance. For this study, the variable of choice was operationalized by providing players with a choice as to the non-player character (NPC) that would act as their "guide" during the game. These NPC guides provided feedback to players in the case of incorrect answers. Players in the Choice condition (C) could manually select a guide from six potential NPCs. Players in the Non-choice condition (N) were assigned guides automatically in the same proportions as those players who selected their own.

Feedback as an Assessment Mechanic

The assessment mechanic varied in this study was type of feedback. Feedback is arguably one of the most studied areas of learning and instruction and has a rich history in instructional theory. Research on feedback generally confirms that learners are more effective when they attend to externally provided feedback (Butler & Winne, 1995, p. 246). Furthermore, feedback has "the capacity to turn each item of assessed work into an instrument for the further development of each student's learning (Hyland, 2000, p. 234). In addition, external feedback has been shown to influence how students feel about themselves both positively and negatively, and what and how they learn (Dweck, 1999).

Video games, both educational and recreational, are filled with feedback. Many games use visual and audio feedback to let players know if certain actions have succeeded or failed. Such feedback communicates, to the player, details about the game's inner sates and its core mechanics (Adams, 2009, p. 225). In video games, feedback is half of the "circular model of gameplay," where the "gamer's input and the game's output reciprocally influence each other" (Heaton, 2006).

To operationalize feedback in the studied game, the researchers provided players with one of two types of feedback: informative or elaborative. The informative feedback was similar to what Kluger and DeNisi (1996) called "knowledge of results", which from an assessment perspective, is of little value as it does not elaborate on why the answer was wrong, just that it was wrong and not the desired outcome. The second type of feedback was termed elaborative. The goal of the elaborative feedback was to provide players with more applicable information on what to do to correct an error when an incorrect answer was submitted.

The Current Study

This study aimed to examine the impact of these learning and assessment mechanics on three dependent variables: learning, motivation, and in-game performance. The overarching question was if and how the inclusion of each of these theoretically based, non-game mechanics would alter these important aspects of games for learning.

The first research question focused on the role of player choice in selecting a NPC to serve as an instructional guide throughout the game. Will giving players control over which character provides feedback influence learning, engagement, and in-game performance? It was hypothesized that providing players with choice would positively impact all three variables.

The second research question centered on which type of feedback, elaborative or informative, would have a more positive impact. It was hypothesized that elaborative feedback, which is meant to guide the player toward the correct solution, would result in higher player motivation, ultimately leading to better understanding of the instructional material and more efficient in-game performance. Conversely, informative feedback, which simply confirmed that an answer was incorrect, was hypothesized to offer little additional value to players, resulting in lower motivation and a reduced understanding of the instructional material.

A third research question focused on the potential interaction between choice of NPC and feedback type. If a player is allowed to choose a NPC, but that NPC only gives informative feedback, will the affordance of choice alone be enough to off-set receiving the presumably less valuable informative feedback? If so, to what degree will the results be measurable?

Methodology

Design and Participants

To explore these questions, a two-factor study with an experimental design was conducted. One hundred and thirty-eight (N=138) sixth graders were randomly assigned to one of four conditions. These four conditions were based on two experimental factors: choice of NPC (C) versus no-choice of NPC (N) and style of feedback, informative (I) versus elaborative (E). This resulted in four randomized experimental groups summarized in Table 1.

	Choice (C)	No-Choice (N)
Elaborative Feedback (E)	*N = 35*	*N = 37*
Informative Feedback (I)	*N = 34*	*N = 32*

***Table 1*: The four experimental groups and the number of participants per condition.**

Procedure

The experiment lasted two days, consisting of approximately two instructional periods. The Day 1 session consisted of introducing participants to the project, answering questions, and conducting a 15-minute paper-based pre-test with 21 questions about the game's educational content (see standards 4.G, 4.MD, 4.OA, 5.G, 7.G, and 8.G in National Governors Association Center for Best Practices, 2010). On Day 2, participants were given one 30-minute play session followed by a paper-based post-test. Students were told that the game consisted of six chapters and the goal was to advance as far as possible in thirty minutes. Students played individually at computer consoles, with a pair of headphones and "scratch" paper for note taking. After thirty minutes, students were asked to exit the game and the paper-based post-test was administered. Students were given approximately 15 minutes to finish the test, which marked the end of the study.

Instruments and Measures

Educational Video Game

The educational video game used for this study was *Noobs vs. Leets: the Battle of Angles and Lines*. This game was developed by researchers at the G4LI and was previously shown to be an effective educational intervention (see Plass et al., 2011b). The game teaches angle rules and has a simple story in which players help save their friends by unlocking paths represented by lines and angles. The paths are unlocked by solving for unknown angles. Each of the game's six chapters introduces a new concept about identifying and calculating angles. For example, the first chapter starts with types of angles (e.g., acute, obtuse, right, etc.) and their numerical values. As players progress through the game, they are required to apply more complex concepts, such as the complementary, and supplementary angles rule. At the beginning of each chapter, players are provided a brief video tutorial about the new concepts covered. In total there are six chapters and the game increases in difficulty with each chapter.

For the experiment, modifications to the original game produced by the G4LI were made. The first change appeared before participants started playing the game. Depending on experimental group, players were asked to select a NPC (Choice condition) or auto-assigned a NPC (No-Choice condition). In both conditions, players were told the NPC guide would "give you hints and tell you how you're doing." Students in the choice condition were given an array of six characters to pick a guide

from. These character images were selected from previous G4LI research (Turkay, Hoffman, Gunbas, & Chantes, 2011) and included three of the most liked and three of the least liked avatars. In the No-Choice condition, the game automatically assigned players one of the characters from the six made available in the choice condition.

The other variable manipulated in the game was the type of feedback presented to the player when an incorrect answer was submitted. This feedback was displayed in a panel that would drop down from the top of the screen. In this panel, the player would see their "guide", the character that they had selected or were assigned, and a speech balloon with text. This text was also spoken by a voice-over recording that matched the gender of the NPC. In the Informative condition (I), the feedback told the players what they did wrong. For example, if the player clicked on an angle that was too big, the NPC would say, "The angle you selected is bigger than 90 degrees." In the Elaborative condition (E) the feedback provided information about what the players needed to do to find the correct answer. For example, if a player clicked on an angle that was too big, the non-player character would say, "For 90 degrees, look for two rays which are perpendicular to each other." Each piece of feedback was preceded by a randomly selected preamble, such as "I'm sorry that is incorrect" or "This is not quite right." The preambles were the same for both conditions. The panel in which the NPC appeared could be dismissed at any time by clicking a close button. This feature allowed players to interrupt and remove the feedback at any time. Players also had the option of repeating the audio feedback by pressing a button labeled "Repeat."

Learning Measures

This study used several methods of data collection in assessing the potential gains in learning and engagement motivation. To test both prior knowledge and post-intervention knowledge, a paper-based test was designed by the researchers, which covered the topics introduced in the game. The pre- and post-tests both assessed the participants' knowledge of angle types (9 questions), angles within triangles (4 questions), angles within quadrilaterals (2 questions), and the rules concerning complementary, supplementary and corresponding angles (6 questions).

Motivation Measures

Motivation was measured using in-game questions presented at the end of each chapter. After each of the six chapters, students were asked to answer five questions about their experience. Using a five point Likert scale (1 = "Not at All"; 5 = "Very Much") students were asked about their engagement in the game. The five questions included: 1) How much fun was this part of the game?, 2) How difficult was this part of the game?, 3) How much do you want to continue playing this game?, 4) How interesting was this part of the game?, and 5) How helpful have your character's hints been in this part of the game? Answers to these questions were required in order to proceed to the next chapter. All answers were recorded in log files.

Performance Measures

In-game performance was recorded using detailed log files kept during game play. These log files recorded all actions taken within the game environment. This included speed of game play, correct and incorrect answers, answer attempts, length of time on feedback screens, and more. These files were subsequently parsed and analyzed to extract data about each participant's specific in-game actions.

Results

A preliminary analysis was conducted to ensure the four experimental groups were equivalent in prior knowledge at the beginning of the experiment. This was done by comparing the pre-test scores of the four groups. A one-way between-subjects ANOVA found no significant difference between the four groups.

The researchers' first question asked whether providing players with a choice of NPC would influence learning, motivation, and in-game performance. Since the two groups, choice (C) and no-choice (I), did not differ significantly on their pre-test score, an independent samples *t*-test was conducted. The test found no significant difference in the post-test score of the two groups. The next logical step was to look at whether there was a significant change from the pre-test to the post-test for the two groups. A paired-samples *t*-test found a significant change from pre-test score to post-test score for subjects in the Choice condition ($t = 4.043$, $p < .001$). The mean pre-test score for subjects in the choice

condition (*M* = 11.406, *SD* = 4.860) increased at the post-test (*M* = 12.906, *SD* = 5.283). A paired samples *t*-test found no significant change from pre-test to post-test for the No-choice group.

Next, the researchers turned to whether or not having a choice of NPC would influence subjects' self-reports of engagement. Each question was answered on a five-point Likert-scale (1 = "Not at All"; 5 = "Very Much"). The analysis examined responses provided at the end of chapters one through three. Responses to chapters four through six were not analyzed because not enough players completed those chapters in the allotted time.

The first question asked players to report the amount of fun they had in the chapter they just completed. Independent samples *t*-tests found no significant differences between the C and N groups at the end of all three chapters. It should be noted, however, that the C group reported higher mean fun ratings than the N group. These differences were not statistically significant. The second question inquired about how difficult the chapter was. Independent samples *t*-tests found no significant differences between the C and N groups at the end of each of the three levels. The third question was about the subjects' desire to continue the game after completing a chapter. No significant difference was found between the C and N groups at the end of chapter one. A significant difference was found between the two groups at the end of the second chapter (t = -2.00, p = .047). The C group reported having a higher desire to continue (*M* = 4.32, *SD* = 1.098) compared to the N group (*M* = 3.88, *SD* = 1.409). However, no difference was found at the end of chapter three, although the average desire of the two groups was quite high at the end of this chapter, with the C group reporting the highest desire to continue (*M* = 4.51, *SD* = .952) compared to the N group (*M* = 4.06, *SD* = 1.045). A fourth question asked about subjects' interest in the game. An independent samples *t*-test found no significant differences between the C and N groups at the end of each of the three chapters. The fifth and final motivation question asked about the perceived helpfulness of the NPC guide. An independent samples *t*-test found no significant difference between the two groups at the end of the three chapters as measured by self-report. However, it should be noted that the C group reported higher helpfulness scores on average when compared chapter by chapter with the scores reported by the N group; however, the differences were not statistically significant.

The third aim of this study was whether or not the choice mechanic would influence in-game performance. Three measures of in-game performance were used: 1) the total number of completed levels, 2) the total time (in seconds) per level, and 3) the number of incorrect answers submitted for chapters one, two, and three. An independent sample *t*-tests found no significant difference in the total number of levels solved by each group or the average number of seconds spent per level. In terms of the number of incorrect answers submitted, no significant difference was found between the two groups after the first chapter. However, a significant difference was found in the second chapter (t = -0.130, p = .044) with the C group averaging more incorrect answers (*M* = 24.45, *SD* = 15.633) than the N group (*M* = 24.06, *SD* = 24.06). No significant difference was found for the third chapter.

The same three research questions were also asked of the assessment mechanic embedded in the game: informative versus elaborative feedback. To examine how feedback type influenced learning the researchers first examined whether the two groups differed significantly in their pre-test score. An independent samples *t*-test found no significant difference between the two groups. They also did not differ significantly on their post-test scores. There was, however, a significant change between the pre-test and post-test score of the E group (t = 3.128, p = .003). A significant change pre-to-post was also found for the I group (t = 2.086, p = .041). The change between the pre- and post-test between the two groups was not statistically significant.

The second research question asked if feedback type would influence subjects' engagement self-reports. No significant differences were found for any of the measures between the two groups on the examined chapters one through three.

Finally, three measures of in-game performance were compared across the two feedback groups. No significant differences were found for the number of levels completed and the average time spent per level. In addition, no significant differences were found between the two feedback groups as measured by the number of incorrect answer submitted in chapters one, two, and three.

Thus far, the analysis examined two variables, choice type and feedback type, independently. These independent analyses show some significance in terms of pre-to-post gains between the C and N

conditions. However, no significant difference was found between the informative and elaborative feedback types. This result suggests that having a choice of NPC character impacts student learning while feedback type does not. With this in mind, the researchers examined the impact of both variables using a two-way factorial analysis. The results of this analysis show that when examined together neither choice nor feedback were significant predictors of students' learning.

Discussion

The first goal of this study was to provide a concrete example of how educational games can be thought of in terms of distinct mechanics that work together to create a fun yet educationally valuable experience. By thinking of games and their effectiveness in terms of game, learning, and assessment mechanics, educational game designers have more powerful lenses through which to reflect on why games work or do not. The researchers and the G4LI feel this is a valuable contribution to the field.

The second goal was to examine the effectiveness of two specific candidate mechanics: the learning mechanic of choice and the assessment mechanic of feedback. The results show that providing players with a choice of NPC positively influences learning outcomes, as well as aspects of motivation and in-game performance.

Perhaps the most intriguing finding of the study is that although students in the choice condition answered statistically more problems incorrectly than the no-choice group, their average reported interest and desire to continue were higher than the no-choice group. In other words, despite submitting more incorrect answers, the choice group reported having higher levels of motivation for the game. This is rather counter-intuitive in that one might expect submitting more incorrect answers to elicit greater negative affectation. In this case, however, it seems that the choice of NPC offset or protected against the negative experience of answering incorrectly. This is rich area for further study.

Another important area of discussion is the study's instantiation of the choice mechanic. Recall that learning mechanics are by definition theoretical and must implemented concretely within a game's ecosystem. The current study chose to do this through the use of a NPC character selection screen presented before game play began. This is, of course, but one way to instantiate choice; there are many other possibilities worthy of exploration. How else can the choice mechanic be operationalized within a game context? Are some instantiations more effective than others? For example, what if players could choose a new NPC guide at the end of each chapter of the game? How would this impact learning, motivation, and in-game performance?

Finally, the researchers hypothesized that different types of feedback would influence students' learning, motivation, and in-game performance; this turned out not to be the case. This does not mean that feedback cannot or should not be used as an assessment mechanic in educational games. Indeed, feedback has a long and well-argued history in education. The lack of a significant effect in this study is likely to have more to do with how the mechanic was operationalized rather than some inherent issue with feedback itself. For example, perhaps the two types of feedback were not different enough to elicit any change. Another possibility is that the elaborative feedback simply wasn't elaborate enough to help the target audience. Clearly, other explanations exist and more research is needed to find the best ways to implement feedback into games for learning. The point is that distinguishing between game, learning, and assessment mechanics is a useful approach to organizing and implementing iterative games for leaning research.

References

Bandura A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W.H. Freeman.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development, 53*(1), 86-107.

Bjork, S., & Holopainen, J. (2005). *Patterns in game design*. Hingham, MA: Charles River Media, Inc.

Adams, E. (2009). *Fundamentals of game design*. Berkley, CA: New Riders Pub.

Butler, A. C., Karpicke, J. D., & Roediger, H. L., III. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, *13*(4), 273–281.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245-281.

Collins, B. E., & Hoyt, M. F. (1972). Personal responsibility-for-consequences: An integration and extension of the "forced compliance" literature. *Journal of Experimental Social Psychology, 8*(6), 558-593.
Cordova, D. I., Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 19*(88), 715–730.
Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215-225.
Fullerton, T., Swain, C., & Hoffman, S. (2008). *Game design workshop: A playcentric approach to creating innovative games.* Burlington, MA: Elsevier Morgan Kaufmann.
Heaton, T. (2006). A circular model of gameplay. *Gamasutra.com.* Retrieved from http://www.gamasutra.com/features/20060223/heaton_01.shtml
Hyland, P. (2000). Learning from feedback on assessment. In P. Hyland & A. Booth (Eds.), *The Practice of University History Teaching* (pp. 233-247). Manchester, UK: Manchester University Press.
Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.
Leotti, L. A., Iyengar, S. S., & Ochsner, K. N. (2010). Born to choose: Biological bases for the need for control. *Trends in Cognitive Science, 14*(10), 457−463.
National Governors Association Center for Best Practices (2010). *Common Core State Standards for Mathematics.* Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
Plass, J. L., Homer, B. D., Kinzer, C. K., Chang, Y. K., Frye, J., Kaczetow, W., Isbister, K., and Perlin, K. (in press). Metrics to assess learning and measure learner variables in simulations and games. In M.S. El-Nasr, A. Canossa, A. Drachen, & K. Isbister (Eds.), *Game Telemetry and Metrics: Maximizing the Value of your Data*. New York, NY: Springer.
Plass, J. L., Homer, B. D., Kinzer, C. K., Frye, J., & Perlin, K. (2011a). Learning mechanics and assessment mechanics for games for learning. G4LI White Paper # 01/2011. Version 0.1 September 30, 2011. Available online at g4li.org.
Plass, J. L., Homer, B.D., Hayward, E.O., Frye, J., Biles, M., Huang, T.T., & Tsai, T. (2011b). The effectiveness of different game mechanics on motivational and educational outcomes in a middle school geometry game. *Submitted for publication*.
Salen, K., & Zimmerman, E. (2004). *The rules of play.* Cambridge, MA: MIT Press.
Shaffer, D. W., Squire, K. R., Halverson, R., and Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan, 87*(2), 104-111.
Topping, K. J. (1988). *The peer tutoring handbook: Promoting co-operative learning*. Cambridge, MA: Croom Helm Ltd.
Turkay, S., Hoffman, D. L., Gunbas, N., & Chantes, P. (2011). Investigating video game characters as role models in science, technology, engineering and mathematics. Paper presented at the Teachers College Educational Technology Conference, New York, NY.
Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). On the importance of self determination for intrinsically-motivated behavior. *Personality and Social Psychology Bulletin, 4(3)*, 443-446.

Acknowledgments

This work was funded in part by Microsoft Research through the Games for Learning Institute. The content and opinions herein are the author's and may not reflect the views of Microsoft Research, nor does mention of trade names, products, or organizations imply endorsement. The authors would also like to acknowledge Murphy Stein who was instrumental, both intellectually and technically, in creating the original version of *Noobs vs. Leets: the Battle of Angles and Lines* upon which this research is based.