ABSTRACT

BOYCE, ACEY KREISLER. Deep Gamification: Combining Game-based and Play-based Methods. (Under the direction of Dr. Tiffany Barnes.)

Gamification, the use of video game elements in non-game environments, is a very popular way to enhance existing systems. Traditional gamification focuses on the addition of points, leaderboards, and achievements in order to create gamified systems. This approach results in purely game-based systems and does not take advantage of play-based elements. Through a series of five studies we have incrementally gamified the educational tool Virtual Bead Loom using a variety of game elements to guide user behavior, resulting in a gamified system combining both game-based and play based elements with higher student learning and engagement compared to the traditional gamification approach of simply adding points and leaderboards. These studies have resulted in a new approach to gamification, Deep Gamification, which blends game-based and play-based elements to improve learning gains and system preference in a wider range of users compared to the traditional gamification approach.

Deep Gamification: Combining Game-based and Play-based Methods

by
Acey K. Boyce

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Computer Science

Raleigh, North Carolina

2014

APPROVED BY:

Tiffany Barnes, Ph.D. Committee Chair	Michael Young, Ph.D.
Kristy Boyer, Ph.D.	Eric Wiebe, Ph.D.

DEDICATION

For my mom and dad. Thanks for putting up with this long and crazy adventure.

BIOGRAPHY

Acey Boyce will graduate from North Carolina State University in August 2014 with his PhD in Computer Science focused on the incremental development and refinement of BeadLoom Game into a deeply-gamified educational system for teaching Cartesian coordinates, graphing, iteration, layering, and optimization to middle school children. Boyce led a team of 4 students from 2008-2012 to design, prototype, test, and develop the BeadLoom Game for deployment in middle school camps, schools, and after school programs while doing his Master's degree at UNC Charlotte. Boyce co-led and co-developed a 10-week curriculum for the Citizen Schools middle school community outreach program that introduces middle school students to Game Development careers. He has conducted 12 summer camps between 2006-2012 for middle school children to learn basic math and computing concepts through playing educational games and making games themselves. Soon-to-be-Doctor Boyce worked at Warner Brother games as an intern in 2013, where he and one other intern single-handedly changed the company's marketing strategies by applying data mining techniques to Twitch.tv and in-game data to better understand player behavior and choices.

ACKNOWLEDGEMENTS

First I want to thank my advisor Dr. Tiffany Barnes. Without her I would have never gone for my PhD and never had any of the exciting adventures that have happened as a result. Although I introduced her to Dominion so I guess we are even.

I also would like to thank my mom for putting up with my long academic career and supporting me all the while. I doubt she expected me to spend 10 years in college learning about video games. You are the best.

Finally, I thank my friends and colleagues Michael Eagle and Drew Hicks. You helped keep me sane these past years and introduced me to such vices as League of Legends and Hearthstone. You guys are awesome.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The purpose of this research is the improvement of educational video games through a novel process we call "deep gamification." Deep gamification seeks to create gamified educational systems that produce higher learning gains and are preferred by more users when compared to traditional educational software and traditional gamified systems. This approach combines the two types of games in order to appeal to a wider range of users and guide user interaction towards desired behavior and learning outcomes.

Early Educational Games - Edutainment

Games have been shown to have inherent motivational properties that allow them to be used for improving educational applications (Garris, Ahlers, & Driskell, 2002; Gee, 2003; Barnes, et al., 2008). The use of games for education is not a new area of research. The term "Edutainment" was coined in the early 1990s and was adopted by academia and industry for games targeted to children. The purpose of edutainment was "to attract and hold the attention of the learners by engaging their emotions through a computer monitor full of vividly coloured animations" (Okan, 2003). In his book Beyond Edutainment Egenfeldt-Nielson surveyed over 50 edutainment titles and a variety of edutainment research in order to define a set of characteristics of edutainment software (2005). He has identified a set of characteristics shared by edutainment software:

- Little intrinsic motivation
- No integrated learning experience
- Drill-and-practice learning principles
- Simple gameplay
- Small budgets
- No teacher presence
- Distributed and marketed differently than commercial games

A major problem with this early approach to educational games was that the learning elements were often completely separate from the game elements resulting in a lack of an integrated learning experience (Egenfeldt-Nielson, 2005). Game elements can be defined as "a set of building blocks or features shared by games" (Deterding, et al., 2011a). The game elements of edutainment games were often limited to colorful graphics used to lure in a player and did not tie into, or augment, the experience (Egenfeldt-Nielson, 2005). In the worst cases these edutainment games would have the player play a traditional game then engage with traditional learning content before being allowed to continue the game. For example, a user may engage in a traditional match-three puzzle game, but might be forced to answer math questions between moves or game rounds. Answering the math question, and thus learning, is completely separate from the game. This can result in "learners [who] will not be motivated to learn but just to play with the computer" (Okan, 2003). The learning component becomes an obstacle to be done as quickly as possible or even skipped completely in order to get to the game experience. Nielson found many examples of this type of behavior in his survey of edutainment software and studies including “skipping text about the pyramids [the learning objective of the game] and going straight for the mini-games located in the game universe” (2005).

Serious Games

As the popularity of edutainment games declined a new term and design philosophy emerged called serious games. Serious games are “games that do not have entertainment, enjoyment, or fun as their primary purpose” (Chen & Michael, 2005). As this definition suggests, there are a wide variety of different types of serious games including (Chen & Michael, 2005):

- Military Games - Goal is to train soldiers in safe game environment
- Educational Games - Goal is to teach a skill or knowledge
- Corporate Games - Goal is to incentivize or evaluate work performance
- Healthcare Games - Goal is to encourage exercise and other healthy habits
- Games with a Message - Goal is to make the player think about topics such as politics, religion or art

While it is good to know that games can be used for such a wide variety of purposes, using use one term for them all is problematic. For example, game elements and design principles that are effective for learning may not be effective for making a game with a message about child labor. This makes it very difficult to discuss things like design principles, effective practices, or the impact of game elements for serious games. It is for this reason that this research focuses only on the Educational Game category (Chen & Michael, 2005) of serious games.

It is also a challenge to create serious games. People may have a message, skill, or learning objective they want to express through a game but not have the game design skills or tools to create it (Djaouti, Alvarez, & Jessel, 2010). The inverse, having knowledge of game design and the tools to make games but not having an expert knowledge of a field, is also often true. In other words serious games require a partnership between experts in the field of the game's purpose and game designers, or game designers with knowledge of both areas.

A problem with both edutainment and serious games approaches is that they define an end product. The reason why this is problematic is that the end product is a game, a term that is itself very loosely defined. One definition of game is "a form of play with goals and structure" (Maroney, 2001) or put formulaically Game = Play + Goals + Structure (Nicholson, 2012b). An important distinction for educational games is the distinction between an educational game and an educational tool. Nicholson defines an educational tool as software with the goal of teaching with structure or formulaically Tools = Goals + Structure (2012b). Using simple substitution he obtains Game = Tool + Play. It would seem like the simple addition of "play" is all it takes to make a tool into a game, however even with these definitions the line between game and tool is extremely fuzzy (Deterding, et al., 2011b). Deterding et al. illustrate this fuzziness by looking at *Foursquare*. *Foursquare* is a smartphone application (app) that allows users to check in at restaurants and other locations.

Although the primary use of the app is to collect coupons and reviews, users can also compete to become the “mayor” of specific locations by having the most check ins over the last 60 days. Is *Foursquare* a game or simply a “gamified” system? What complicates this question is that the answer may depend on the user and how the user interacts with the system (Deterding, et al., 2011a). Many users enjoy “playing” *Foursquare* and attempting to become the mayor of their favorite hangouts while others use it only for reviews and deals. In a similar way, the line between educational tool and game is very difficult to define. In order to avoid this “is the system a game” debate we do not claim to be creating a serious educational game. Instead we will focus on the addition of game elements to existing tools.

Gamification

Gamification is a relatively new term that first appeared in 2008 and became widespread by 2010 (Deterding, et al., 2011a). Gamification is defined as “the use of video game elements in non-gaming systems to improve user experience and user engagement” (Deterding, et al., 2011c). Hamari et al. expand on this and state gamification is “a trending topic and a subject to much hype as a means of supporting user engagement and enhancing positive patterns in service use, such as increasing user activity, social interaction or quality and productivity of actions” (2014). In other words, gamification is the enhancing of a non-game systems or tasks through the addition of game design elements, resulting in a gamified system. The most famous examples are classic consumer rewards programs that add points to the non-game task of buying products (Nicholson, 2012a).

Although the definition of gamification suggests all game elements could be used, gamification tends to focus on the elements of points, badges, and leaderboards (Bogost, 2011; Nicholson, 2012b, Iosup & Epema, 2014). The system awards points for completing certain tasks and presents badges based on the number of points achieved. These scores are recorded on a leaderboard so that users can compare their performance to their peers. The theory behind this type of gamification is that if users have their scores beaten they are challenged to overcome the new best scores and are thus incentivized to replay the game. Liu et al. refer to this as the "gamification loop" as outlined in Figure 1 (2011).

Figure 1. The gamification loop (Liu et al., 2011)

The cornerstone of this approach is the point system. The simple addition of points can sharply increase user performance in a system or task. Mekler et al. investigated the impact of points and providing a meaningful frame on performance and motivation in an image annotation system (2013). They conducted their experiment with 172 participants (26 male, 123 female, 3 not specified; mean age 32.95 years, range 15-74) using a 2x2 between-subjects experimental design. Their independent variables were points (points vs. no points) and meaningful framing (framing vs. no framing). Participants were assigned randomly to one of the four pairings then asked to annotate one image without points and without a meaningful frame. The fourteen additional images were then annotated with the additional features of their group. Meaningful frame was established by telling the participants that their tags would help improve computerized image categorization. Points were awarded at 100 points per annotation. After finishing the image annotation activity, participants filled out the Intrinsic Motivation Inventory (IMI) to measure their motivation (Merker et al., 2013).

Merker et al. found that the single game element of points resulted in higher quantity of image annotations and in higher intrinsic motivation (Merker et al., 2013). They hypothesize that this could be a result of people's inherent need to accomplish tasks (Jung et al., 2010) and that points can provide individualized performance feedback that can fulfill this need. They also found that when combined with the meaningful frame there was an increase in time spent per tag ("a proxy for participants' effort to generate task-relevant labels") and tag quality (Meckler at al., 2013). Tag quality was measured both as a binary of sensible or nonsensical and using the Linguistic Inquiry and Word Count (LIWC) Analysis and with both measures the combination of points and meaningful frame resulted in significantly higher quality compared to other groups. Goh and Lee also gamified an image annotation system using both cooperative and competitive point systems and found the point based approaches to be especially motivating for users “with a high need for achievement” based on participant feedback (2011).

There has been a great deal of research on the gamification of image annotation systems because it represents one of the oldest gamified system types: the gamification of solving an unknown problem. This type of “game with a purpose” (von Ahn, 2006) (GWAP) predates the term gamification and is a popular form of “human computation” (von Ahn, 2009). GWAPS are examples of gamification since they add game elements including points and objectives to a non-game context. More recent research by Mekler et al. has begun referring these additions of game elements to create GWAPS as gamification (2013). The oldest example is an image annotation game called the ESP game (von Ahn, 2004). In the ESP game, players are matched with a partner and then both players are shown the same image. The goal is for both players to input the same word describing the image. When enough players have input the same word for an image it becomes a taboo word and players are no longer allowed to use that word for that image. This results in players having to input less obvious words in order to score for images that already have the easy or obvious tags. Another problem that has been gamified is the prediction of protein structures in the game Foldit (Cooper et al., 2010). The goal in Foldit is to take a 3D model of a protein and manipulate it so it takes up the smallest volume possible and creates multiple good bonds. Players are given points based on how well their folding meets these criteria and they are ranked against the other players' scores. Foldit was used to help decipher the structure of the Mason-Pfizer monkey virus, a problem which had been unsolved by scientists for 15 years (Khatib et al., 2011).

What distinguishes gamification from serious games? Gamification focuses on the process and elements used to create the gamified system rather than trying to define the end product. Although this process still leverages the motivational properties of games it does not concern itself with whether the final product is or is not a game. Examples of this distinction can be seen by examining various gamified systems for physical health. Some of these are referred to as "exergames" (Braumer er al., 2013) which are a type of serious game (Bergeron, 2006).

A very common approach to creating an exergame is to have an exercise task and assign points to the player for performing this task. This task can be anything from collecting apples (Brauner, 2013; Nickel et al., 2014), dodging asteroids (Nickel et al., 2012; Nickel et al., 2014), to standing instead of sitting while riding on crowded Japanese public transportation (Kuramoto et al., 2013). This use of objectives and points clearly falls in line with both the definition and the common focus of gamification (Brauner et al., 2013). However looking at a traditional fitness application like My Fitness Pal (myfitnesspal.com) the line between gamified systems and serious games becomes slightly less clear. My Fitness

Pal uses calories as points and additional game elements like leaderboards and achievements to help users achieve their fitness and weight loss goals. As with Foursquare, it is possible for some to engage with My Fitness Pal as if it was a game but it is not clear if it can be called a game. What is clear is that My Fitness Pal has enhanced the task of tracking nutrition and exercise with game elements, and thus has had some degree of gamification. Whether systems like these which have been enhanced with gamification are "games" or "gamified systems" is left to other research. Our research focuses on the impact of gamification on users.

Like exergames, many educational serious games are actually gamified systems. Many serious games are designed from scratch rather than beginning with an existing educational tool and adding game elements to them. However, many can be thought of as an educational tutor with some additional game elements. For example CopyMe is an educational serious game designed to teach emotion recognition to children with autism (Harrold et al., 2014). It was built from the ground up, but CopyMe can be thought of as an emotion recognition tutor with the added game elements of points and achievements. Players are shown the face of someone expressing one of six emotions. Players are then tasked with replicating the emotion with their own faces while being recorded using a camera. If they successfully replicate the face and emotion they receive a point. Figure 2 shows the game with the goal expression on the left, the image of the player being captured by the player on the right, and the score of zero being displayed at the top left (Harrold et al., 2014).

Figure 2. The educational serious game CopyMe (Harrold et al., 2014)

The Code of Everand is an educational massively-multiplayer online (MMO) game designed to teach road safety to preteens and teenagers (Dunwell et al., 2014). It can be thought of as a road safety tutor with many MMO game elements. These elements include a world map, turn based combat, levels, equipment, story, avatar customization, quests, and multiplayer communication. The creators constructed an “abstract metaphor for road crossings” called “spirit channels” within the game. While navigating the world the players carefully select a route, look both ways before crossing a spirit channel, and are encouraged to form peer groups for safety. When crossing the spirit channels they engage in turn based combat and defeat the enemies by looking left and right and using their accumulated items and strength (Dunwell et al., 2014). Other than looking left and right this combat aspect of the game does not contribute to the learning and is there only for the enjoyment of the players.

The Code of Everand and CopyMe are designed for audiences of teens and younger but serious educational games can also be used to teach college level material. Wu's Castle is a serious educational game designed to teach loops and arrays in college level introductory computer science courses (Eagle & Barnes, 2009). In it players visually walk through loops and nested loops by navigating corridors. Players also learn to use arrays by placing snowman heads on snowman bodies; each snowman body or full snowman represents the data that is held in a particular array position as shown in Figure 3. Wu's Castle was found to be more effective at teaching the concepts of arrays and nested loops than a traditional programming homework assignment (Eagle & Barnes, 2009). While this system does not have a traditional point system it does include a progress bar which acts as a reward as the player advances.

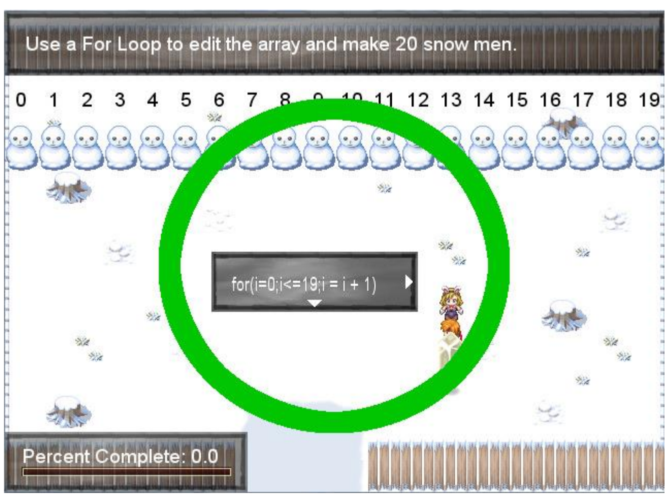


Figure 3. The educational serious game Wu's Castle (Eagle & Barnes, 2009)

Another highly researched game-based system is the reading tutor iSTART. iSTART (Interactive Strategy Training for Active Reading and Thinking) is a reading tutor developed by Danielle McNamara which was converted to the game-based tutor iSTART-ME in 2009 (Jackson et al., 2009). The gamified version of the system featured points, levels, purchasable rewards, mini-games, and customizable avatars (Jackson et al., 2009). The evolution of this system is an excellent example of typical long-term educational game research. Early work with the game was primarily focused on explaining the system (Jackson et al., 2009). This was followed by studies showing that the game-based system was rated as more fun than the non-game version (Jackson, et al. 2011) and that the game-based system was more motivational than the non-game version (Jackson & McNamara, 2011).

Work on iSTART-ME culminated in comparative studies comparing the learning gains of the game-based to the non-game based system. In both studies they found that the game-based system did generate significant learning gains but not ones that were significantly better than the non-game version (Jackson et al., 2012; Jackson & McNamara, 2013). Both studies had participants assigned to either the game or non-game group. After taking a pretest participants engaged with their specified system over a set period of time for multiple sessions. After completing all the sessions, participants took a posttest to evaluate learning gains and a survey to evaluate their enjoyment of the system. In both cases these surveys indicated that although no significant improvement in learning occurred between the two systems, participants enjoyed the game version more (Jackson et al., 2012; Jackson & McNamara, 2013).

This concurrent research to our own illustrates the potential of gamification and the importance of long term gamification research, but does not investigate the traditional gamification method. Instead it looks at the impact of an expanded form of gamification that includes avatars, points as currency, and mini-games. Their iSTARS-ME gamified system does generate more enjoyment than the original iSTART system and provides learning gains equivalent to those in the original system, but does not improve upon them. This calls into question about the effectiveness of purely traditional gamification approaches to achieve improved enjoyment and learning. The iSTART research demonstrates that a successful educational system can be gamified without reaching the full potential of gamification – to achieve both fun and improved learning gains. Their results failed to show that their gamified system could outperform their original non-gamified system in the area of learning gains. Our deep gamification approach seeks to improve both the enjoyment of the system and the learning gains from using the system compared to a non-gamified system.

Many serious game studies analyze the result of the system as a whole rather than investigating the impact of individual aspects of the system. The Code of Everand involved many game elements. Some of these were designed to teach the learning objectives and some were added to improve player retention. Although the Code of Everand was found to be successful at teaching road safety it is impossible to say which of its many game elements contributed to or hindered learning from the system and player retention. Similarly the iSTART-ME system featured six mini-games each designed to teach the learning objectives and included extra features like avatars (Jackson, et al., 2009). The iSTART-ME studies do not evaluate individual game elements separately or independently. This means it is difficult to attribute improvements in the system to particular features, mini-games, or design choices. Studies like this show a game can work but does not answer which game elements contribute to that success. Without understanding the impact individual game elements have on learning and player preference it is impossible to identify what game elements should or should not be used in gamification to maximize learning gains or player preference.

Gamified systems can be designed from the ground up with both domain experts and experts in game design working together. Similar results could theoretically be obtained by starting with an existing educational system designed to teach the topic and adding game elements to it. In fact the research presented in this dissertation began as serious games research. It was later that we realized that in creating an educational game by adding game elements to an existing educational tool that we had engaged in gamification. This allowed us to focus on the game elements and their impact rather than on the educational content of the system.

This ability to add on to an existing system represents one of the major strengths of gamification. In addition to not worrying about whether the final system is a game or not, gamification does not suffer from the two-domain experts problem. Gamification only requires an expert in game design and does not require an expert in the domain of the educational content or the objective task. This is because the gamification is expanding a pre-existing system that has already been developed in concert with domain experts. Although this means two experts are still technically required, there are already thousands of tools and systems that can be enhanced through gamification. All that is needed are people with the skills needed to implement game elements and a deeper understanding of the impact of those elements.

One of the fastest growing areas for gamification is in industry. There seem to be two subcategories in this growing trend. The first are companies designed to assist other companies in gamifying their traditional services (Hamari, 2014). These include Badgeville (Badgeville, 2014), NextBee (NextBee, 2014) and GameEffective (GameEffective, 2014). These companies gamify existing systems and companies with only an understanding of game design thus illustrating gamification's ability to circumvent the two experts problem. Like most gamification implementations the most well-known game elements these companies use are leaderboards, badges, and points (Bogost, 2011; GameEffective, 2014; Badgeville, 2014). These companies provide solutions that award points to customers and employees based on performing certain desirable tasks.

The second area of gamification in industry is the gamification of work-related tasks. This can be as simple as encouraging participation in company programs like social networks (Farzan, 2008), to as complex as giving virtual (green to red) colored blocks to assembly line workers based on task completion time that explode with "the intensity of visual feedback" being based on the number of green blocks obtained (Korn, 2012). The second example may sound farfetched but a simpler implementation of this system is already in use by the superstore Target (Crowley, 2009). Figure 4 shows the screen of a Target cash register after finishing a transaction. Each transaction is scored either a G or green for good or an R or red for needs improvement. Target cashiers are also presented with their last ten scores and an overall percentage of greens. These scores are based on the total transaction time relative to the number of items purchased. It is designed to encourage faster more efficient transactions. Employees have their own personal leaderboard and these scores impact things like raises, promotions, and who remains an employee (Crowley, 2009). This is clearly an example of gamification used more for the improving user engagement over user experience. By 2015 Gartner estimates that over 50% of all organizations managing innovation processes will gamify some aspects of their business (2011).

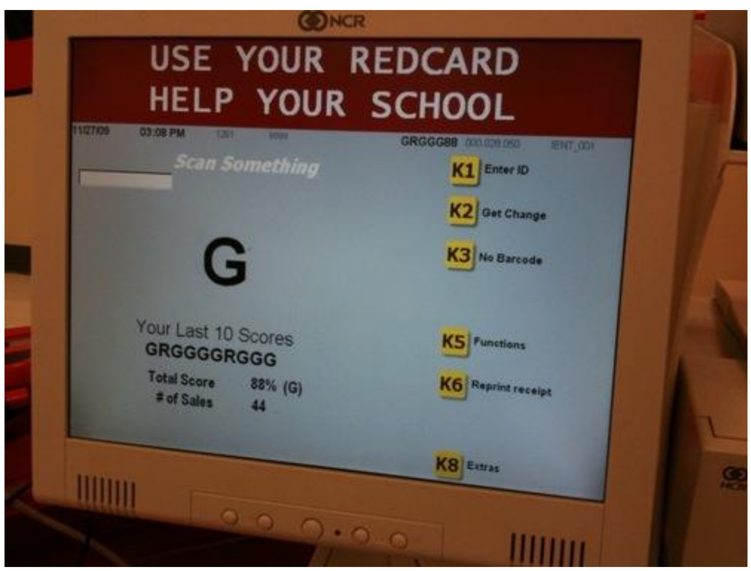


Figure 4. Gamification example from Target cash register (Crowley, 2009)

Similar to the gamification of the work environments many teachers and professors have begun to gamify their classrooms. One popular method is the restructuring of homework and tests into quests and grades into experience points like in traditional role playing games (Gillispie & Lawson, 2010). This approach lends itself extremely well to studying or working with medieval or fantasy environments although that is not required. In their report on the gamification of their technical university courses, Iosup and Epema identify seven core tools for successful gamification of the classroom (2014). There are the core game mechanics of point systems, levels, and leaderboards and the core dynamics of badges, onboarding, social engagement, and unlocking content (Iosup & Epema, 2014). Although this approach involves a few additional elements, classroom gamification, like most other gamified systems, is focused on the game elements of points, leaderboards, and badges.

Like customer reward programs, gamification of the classroom and workplace tend to rely heavily on external rewards in the forms of free items, grades, or payment. This aspect of gamification is often highly criticized (Bogost, 2011; Nicholson, 2012a; Nicholson, 2012b) because external rewards can reduce performance (Kohn, 1999) and internal motivation (Deci et al., 2001). However, not all gamified systems rely on external rewards. For example, gamified problem-solving systems feature no external rewards, with some like the ESP game being played without the user even being aware that they are performing a task (von Ahn & Dabbish, 2004). Exergames and gamified educational systems offer external rewards but only in the form of personal growth whether physical or mental. Regardless of rewards the goal of the gamified system remains, as the definition states, "to improve user experience and user engagement" (Deterding, et al., 2011c).

Gamification boils down to encouraging users to perform a task, and making the task itself more enjoyable through the use of game elements. The target task can range from solving problems, to learning material, to exercising, to doing your job. It is possible that game elements and gamification strategies that are effective for one task may be effective for another; however that is research left for future work. In this research we focus our objective on maximizing the effectiveness of gamification for a single educational software system.

Classifying Serious Games and Gamified Systems

Serious games and gamified systems come in variety of different forms. To make cataloging, searching for, and understanding these systems easier, work has been done on developing classification systems. One approach is "market-based" classifications which aim to classify the systems based on the market that uses the system (Djaouti et al., 2011). The serious game classifications listed earlier represent one such market-based system (Chen & Michael, 2005) although others exist (Zyda, 2005; Alvarez & Michaud, 2008). However, new markets are constantly being explored and discovered and market-based classification focuses on the use of the game rather than its content (Djaouti et al., 2011).

An alternative approach is "purpose-based" classification which aims to classify systems based on the intention they are designed to satisfy (Djaouti et al., 2011). An example of this would be Bergeron's seven purpose categories: Activism games, Advergames, Business Games, Exergames, Health and Medicine Games, News Games, and Political Games (2006). Although these classifications are similar to market-based approaches they have the important advantage of focusing on a game's content rather than the intended use. Focusing on the uses requires making assumptions about the intent of the creators while focusing on the content does not. However, neither of these classification approaches considers the game elements or structures of the system.

Sawyer and Smith combined these two approaches to create a multiple criteria classification system called the "Serious Game Taxonomy" (2008). It categorized serious games by both market (Government and NGO, Defense, Healthcare, Marketing and Communication, Education, Corporate, or Industry) and purpose (Games for Health, Advergames, Games for training, Games for education, Games for science and research, Production, or Games as Work). This system had overlaps between the market categories and purpose categories and also lacked some key purposes such as games with a message. Although there are multiple criteria in this system, it still did not incorporate the game elements of the system.

In their work in developing an improved classification system Djaouti et al. wanted to combine previous classification systems with the concepts of game elements and mechanics (2011). They first looked at game classification used by traditional games in the entertainment industry. However, they found industry standards of genre classification and mechanics were too subjective. The categories also had too much overlap and were in a constant state of flux as genres and mechanics were created and evolved (Djaouti et al., 2011). Their goal was to create a classification system that would be less affected by the evolution of games and work in the context of changes in game mechanics and principles.

Djaouti et al. believed “gameplay” would be a good way to classify the game elements and mechanics of these serious games and gamified systems. First they looked towards Portugal’s definition of “gameplay” as the combination of rules, input methods, space-related setup, time-related setup, and drama-related setup (2006). They chose to focus on the rules because rules lend themselves to formal and objective deconstruction.

“Ludology,” the study of games, acknowledges the existence of different categories of rules (Frasca, 2003). One categorization comes from Caillois’ work on play which defines two types of play: ludus and paidia (1962). Ludic play is structured and has a defined set of rules while paidic play is more freeform. Using this distinction gameplay can be separated into two categories that Djaouti et al. call play-based systems lacking defined goals and game-based systems with defined goals (2011). Since this classification looks at the style of play it is unaffected by the ever-evolving nature of game mechanics and principles. No matter what new game elements emerge, users will still engage with the system in ludic or paidic play making this game-based or play-based classification relatively unchanged over time.

Game-based and play-based classification can define the type of gameplay for entertainment games, serious games, and game-like systems. For example *SimCity* can be defined as play-based and *Pac-Man* as game-based (Djaouti, 2011). Note that this classification is based only on what the system presents you. The user is free to set their own goals in play-based systems. Even if a user decides to create a city with a million people in *SimCity*, this is a self-defined goal, so *SimCity* remains play-based.

We shall be using the terms game-based (featuring system defined objectives) and play-based (featuring user-defined objectives) systems and elements throughout this paper. Education research would refer to game-based systems as more guided learning and play-based systems as more open-ended learning or sandbox environments. Investigating the results of combining these elements is a primary area of investigation in this research.

Combining this with previous market-based and purpose-based classification systems results in the Gameplay/Purpose/Scope (G/P/S) model of game classification outlined in Figure 5 (Djaouti, 2011).

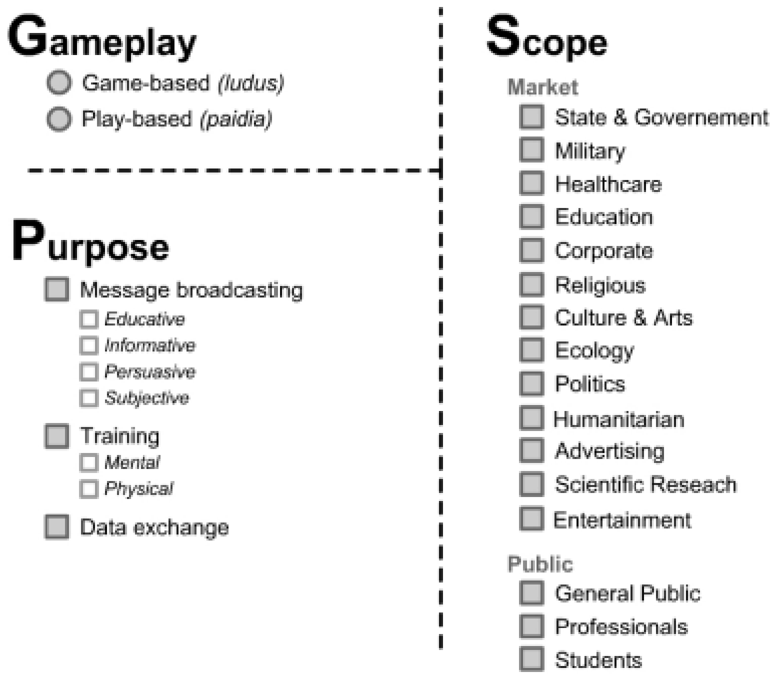


Figure 5. The G/P/S model of game classification (Djaouti, 2011)

It is important to note the binary nature of gameplay in this model. While a system can have multiple purposes and scopes its gameplay is defined as either game-based or play-based but not both. This means if a user prefers play-based experiences and the system only offers a game-based experience the user may not be satisfied. One of the primary areas of investigation for this research is questioning the binary nature of gameplay in the G/P/S model. In other words, we will investigate the impact of combining both game-based and play-based elements into a single system. For this research we will look only at a system with the purpose of mental training, the scope (or target audience) of public school students, and working towards a system with both game-based and play-based elements.

Problems with Traditional Gamification

The most typical approach to gamification is the addition of points, leaderboards, and badges ((Bogost, 2011; Nicholson, 2012b, Iosup & Epema, 2014). Nicholson calls this approach “BLAP gamification” (Nicholson, 2012b). BLAP stands for Badges, Levels and Leaderboards, Achievements, and Points. This approach falls perfectly in line with many research and industry implementations including Liu et al.’s previously discussed gamification loop (2011) and nearly every example covered in the overview of current gamification trends.

Many have questioned this focus on points and leaderboards. To illustrate this point some have suggested renaming gamification to “BLAP gamification” (Nicholson, 2012b), “pointsification” (Robertson, 2010), “Ludification” (Bouca, 2012), or even “exploitationware” (Bogost, 2011). The central argument for all of these name suggestions is that, since games are more than just game-based points and leaderboards, gamification should also be more than just adding points and leaderboards to systems. We will refer to the BLAP form of gamification as “traditional gamification.”

Ludic (another word for game-based) systems have well-structured system-defined rules and objectives. Thus as Bouca’s name “Ludification” (2012) suggests, traditional gamification’s reliance on points, achievements, and rules defining how to earn them, results in gamified systems that are classified as game-based or ludic using the G/P/S model. This means if a user does not enjoy game-based systems they may prefer the original non-gamified system to the gamified system. For example in Goh and Lee's work comparing gamified image annotation systems they found that some participants who used the non-gamified annotation system reported liking that system because "they did not need to challenge anybody" (2011). While traditional gamification and its focus on points and leaderboards is very appealing to players "with a high need for achievement" (Goh & Lee, 2011) not all players prefer game-based systems. Traditional gamification fails to draw users who may prefer play-based systems that do not focus on system-defined challenges.

In their literature review of gamification, Hamari et al. collected 24 empirical studies by the research community each analyzing whether gamification works (2014). They looked at what game elements were used in the gamification of several systems. Story/Theme was the only non-game-based element featured in these studies and it was only used in six of the studies (Hamari, 2014). Points, leaderboards, badges, or achievements were present in 18 of the studies (Hamari, 2014). This gravitation towards only a small subset of game elements may be because of how easy it is to add points, leaderboards, and achievements to existing systems (Nicholson, 2012b). The spectrum of game elements are not being fully explored in a majority of gamification research (Nicholson, 2012a; Nicholson 2012b; Hamari, 2014)

Since play is a central point of games some argue it should be a central point of gamification (Nicholson, 2012b). One of the central points of play is that it is voluntary (Brown, 2009; Caillois, 2005). If a task is required, it is difficult for a user to achieve play. This is an especially pressing challenge in the gamification of educational software that is intended for required use in a classroom environment. This is a difficult challenge but one that is important to address because play can encourage learning (Hirsh, 2008). Meaningful options that can lead to play should be incorporated into the gamification of educational software to maximize appeal while also improving learning gains (Nicholson 2012a; Nicholson 2012b). Traditionally-gamified systems lack these meaningful options and only present one (game-based) way to engage with the system. There are no other options for those students who do not want to engage with a competitive game-based system. Therefore, people not interested in game-based systems are unlikely to identify using that system as play. We believe play-based game elements can be used to provide meaningful options for students who do not see game-based elements as play.

Constructionist Learning Environments

While traditional gamification approaches often lack play, free play and exploration are of central importance in constructionist learning environments (Papert & Harel, 1991). In these systems, students experience discovery learning through the creation of artifacts. These systems are an extension of Piaget's constructivism knowledge theory which highlights the importance of play and suggests human learning occurs through a process of accommodation and assimilation (Piaget, 1955). Assimilation is taking in new information into a person's own current cognitive models without change, and accommodation is the changing of cognitive models to fit with new information (Piaget, 1955).

The first example of a more "play-based" constructionist learning environment software was the LOGO programing language (LOGO Foundation, 2011) where users control a triangular turtle's movement on the screen with programing commands. Other successful examples of (play-based or open-ended) constructionist learning environments linked to computational thinking include Scratch (Resnick, et al., 2009), Alice (Cooper et al, 2000), and Culturally Situated Design Tools (Eglash, et al., 2006). Ito refers to such systems as the "construction genre" of educational software (2009). Students are free to play and explore in these play-based environments and learning occurs organically. There are no objectives or rules. Although these systems are not games, if they were categorized using the continuum between game-based systems defined by rules and objectives, or a play-based system with only potential user-set goals we are using here, they would certainly be play-based systems and not game-based. It should be noted that this is not a rigid classification. For example many teachers when using Alice or Scratch will develop objectives for the class to perform and a way to evaluate those objectives. In doing so they have added "system" defined rules and objectives. This transforms the play-based environment with its user-defined goals into a more game-based environment with system-defined rules and goals.

Although these systems have all been found to be effective teaching systems there are some criticisms of constructionist learning environments and constructivism. Jonassen states that these constructionist learning environments may be ill-structured and that they are less effective for novices (Jonassen, 1997). These novices may be better taught using "well-structured" learning environments (Jonassen, 1997). Mayer also feels that these pure discovery environments are less effective for some learners than guided discovery environments (Mayer, 2004) citing multiple studies comparing pure discovery and guided discovery environments (Kittel, 1957; Gagne & Browne, 1961; Shulman and Keisler, 1966) including a study using the LOGO environment (Fay & Mayer, 1994). In all of these studies a constructionist learning environment was compared against itself with the experimental variable being guided discovery versus pure discovery. In the pure discovery groups participants only engaged with the constructionist learning environment without any outside influence from teachers or experimenters. In the guided discovery groups participants were given feedback and help from teachers or experimenters when they made mistakes in the constructionist learning environment. In all studies the pure discovery approach was found to be less effective at improving learning gains compared to the guided discovery. Pure discovery lacked the guidance and feedback some students needed in order to succeed.

We found an additional problem in our work with the constructionist learning environment of Culturally Situated Design Tools (csdt.rpi.edu) (CSDTs). In all of the CSDTs, there are two ways, one simple and one complex, to create the same artifact. For example, in Graffiti Grapher students can create virtual graffiti using either Cartesian coordinates or the more advanced polar coordinate system. In after-school outreach programs, we observed that students would almost always default to the simplest behavior in creating their artifacts and would not use the more complex systems without outside intervention from instructors. This meant that students could avoid portions of the learning objectives. We believe this behavior of defaulting to the simplest option and avoiding more complex learning objectives can occur in other constructionist learning environments such as using suboptimal coding practices in Scratch or Alice. The desire to encourage full exploration of existing learning environments and systems was one of the initial motivating factors of this research.

Gamification may be the solution to these shortcomings. Gamification provides strong encouragement to perform certain tasks and creates well-structured problems with immediate feedback, which constructionist learning environments lack. Constructionist learning environments' open-ended play-based structure may provide the free play that traditional gamification approaches lack. The separation between traditional gamification methods and constructionist learning environments is similar to the divide seen between game-based and play-based games. By combining game-based elements, elements that define rigid system based rules and outcomes, and play-based game elements, elements that do not impose system defined rules and encourage free play, it may be possible to achieve a system with the benefits of both gamified systems and constructionist learning environments.

Research Questions

Traditional gamification has shown great success in improving a variety of tasks and systems including educational software. Traditional gamification's focus on rules and points results in purely game-based systems which lack options that encourage play. These are major shortcomings of traditional gamification which must be addressed in order to maximize the effectiveness of gamified educational software. This research aims to explore new approaches to gamification through the incremental gamification of an existing educational system. These approaches will combine the success of traditional game-based approaches while giving players new play-based options. We refer to this new approach to gamification and the design methodology associated with it as “deep gamification.”

The success of our new approach will be measured by its ability to improve an educational tool in the areas of learning gains and user preference compared to the original system and a traditionally gamified version of the system. In order to make these measurements we will incrementally gamify an existing educational tool known as the Virtual Bead Loom (VBL). In doing so we will create a gamified system called BeadLoom Game (BLG). Both of these systems will be fully elaborated upon in their respective chapters. Each study we will refine our gamification methodology and further enhance the software and run comparative studies using previous versions for comparison. We believe that the deep gamification techniques developed in this research may suggest ways that designers can create gamified educational systems that outperform traditional educational software and traditional gamified systems. This will be evaluated by the gamified BeadLoom Game’s ability to outperform the non-gamified Virtual Bead Loom and previous incarnations of BeadLoom Game.

This research seeks to address the following research questions:

R1: Can a traditionally gamified educational system produce learning gains?

Since we will be incrementally gamifying the Virtual Bead Loom software it is important that the baseline involving the traditional gamification techniques is successful in generating learning gains. Similar questions looking at the effectiveness of serious games have been conducted (Eagle & Barnes, 2009; Jackson et al., 2012; Jackson & McNamara, 2013; Dunwell et al., 2014; Harrold et al., 2014), however none of these implemented the pure traditional or BLAP gamification approach of points, leaderboards, and achievements/badges without additional features. A rigorous evaluation of the learning outcomes from traditional gamification is a required first step in improving gamification. This question will be addressed in **study one** with the hypothesis of **the gamified system BLG will generate learning gains** in the areas of iteration and layering.

R2: Can the use of a traditionally gamified system outperform a non-gamified educational tool in the areas of learning gains and user preference?

As discussed in the literature review there is a lack of studies formally comparing serious games to non-game counterparts in areas other than user preference. It is important to ensure that traditional gamification results in systems that are both more popular and generate higher learning than non-gamified systems. A formal comparative analysis of a gamified to non-gamified system had yet to be conducted at the time of this research and understanding how traditionally gamified systems compare to non-gamified systems is an important step in the journey to improve gamification techniques. This question will be addressed in **study two** with the hypothesis of **the gamified system BLG will outperform the non-gamified system VBL** in the areas of iteration, layering, and user preference.

R3: Does a gamified system featuring both game-based and play-based game elements outperform a traditionally gamified system in the area of user preference?

Studies one and two evaluate and help inform design principles for the game-based elements associated with traditional gamification: points and achievements. This question addresses the binary nature of the Gameplay/Purpose/Scope (Djaouti, 2011) and seeks to understand what happens when you combine both game-based and play-based elements into a single gamified system. Does integrating both types of game elements into a single system improve user preference? This question is addressed in **study three** and although learning gains are also evaluated the primary hypothesis is the gamified system **BLG 2.0 (featuring both game-based and play-based features) will outperform the traditionally gamified BLG 1.0 and non-gamified VBL** in the area of user preference.

R4: Can play-based game elements be added to a traditionally gamified system to guide user behavior?

As discussed play-based environments can suffer from a lack of guidance. In educational software this may mean that users are not exposed to all of the learning objectives. This question seeks to understand if there is a way to design play-based game elements in such a way as to guide player behavior without impeding on the play-based nature. An evaluation of our initial play-based elements and the associated design principles is conducted in **study four**. It features two hypotheses. The first is that **content made in BLG 2.0 will be more creative and complex than content made in VBL** due to BLG's peer driven content rating system. The second hypothesis is that **content made in BLG 2.0 will be made with more iterations than content made in VBL** due to BLG's peer driven content rating system. This research question will be further addressed in **study five** with the hypothesis of **content made in BLG 3.0 will be made with more iterations than content made in VBL or BLG 2.0.**

R5: Can engaging only with the play-based elements of a gamified system produce learning gains?

Although the goal of deep gamification is the creation of educational gamified systems that combine both game-based and play-based elements it is possible for some students to choose to only engage with one type of element. In other words some students may use only the game-based features and others may use only the play-based features. Studies one and two evaluate the learning gains of the game-based elements alone but it is equally important evaluate the learning gains of only the play-based elements. **Study five** addresses this question with the hypothesis **the play-based mode of BLG 3.0 alone will generate learning gains** in the areas of Cartesian coordinates, iteration, and layering.

R6: What game elements and design principles can be used to create a gamified system that combines game-based with play-based mechanics?

There is not one particular study dedicated to addressing this particular question. Instead this question represents the culmination of the knowledge we shall gain from conducting studies one through five. With each study we will gain a deeper understanding of the impact of individual game elements which we incrementally add to the gamified system BeadLoom Game. In the end we will have a deeper understanding of the effect of select game-based and play-based game elements on learning and preference as well as design principles to help maximize these effects. What we learn from these studies will combine to form our deep gamification design methodology which we will present as our solution to this question.

As shown in Table 1, this research combines and expands upon previous work on educational video games, gamification, and Ludology to create a new and more effective gamification methodology.

Table 1. *Research Overview*

Media	Educational Tools & Educational Video Games
Methodology & Theoretical Framework	Gamification Ludology
Application	Educational game design
Goals	Improve learning gains Higher user preference
Future Directions	Application of new design methodology to existing gamified and non-gamified systems

Virtual Bead Loom

The Virtual Bead Loom (csdt.rpi.edu/na/loom/bl/beadloom.html) is the educational tool used as a baseline for this research. It is one of many Culturally Situated Design Tools used in middle school outreach by the STARS Alliance. Culturally Situated Design Tools are a suite of educational constructionist learning environments designed to teach math and computer science through the creation of cultural artifacts. These tools have been shown to be effective at teaching basic mathematical principles and were designed to be more appealing than traditional assignments especially for minorities that are traditionally underrepresented in the fields of computing and mathematics (Eglash, et al., 2006). Although these tools were designed to appeal to specific cultures, learning through use of the tools was found to be very effective at motivating students regardless of cultural background (Eglash, et al., 2006).

The Virtual Bead Loom is one of the most popular Culturally Situated Design Tools (Bert, et al., 2009). The Virtual Bead Loom is designed to teach Cartesian coordinates, geometry, and the basic computer science concept of iteration. Users use six different graphing functions to place virtual beads on a grid to create Native American bead art. The Virtual Bead Loom has shown positive results in teaching basic mathematical principles and users reported preferring to use the tool compared to a more traditional assignment (Eglash, et al., 2006).

In our STARS Alliance outreach sessions with the Virtual Bead Loom we observed similar results to those reported by Eglash et al. However we also observed that a majority of users did not use the more complex and challenging graphing functions available in the Virtual Bead Loom (Boyce, et al. 2009). This meant that the users were avoiding the game mechanic designed to teach the learning objective of iteration. The Virtual Bead Loom needed a way to guide user behavior and encourage users to explore and learn from all the functions available. This made the Virtual Bead Loom a perfect candidate for gamification.

Virtual Bead Loom Design

Before discussing the gamification of the Virtual Bead Loom it is important to understand its original design. The Virtual Bead Loom (Figure 6) allows users to create their own Native American bead art by placing colored dots (beads) onto a 41 by 41 (-20 to 20 along the x and y axis) Cartesian coordinate grid. This is done through the use of six different graphing functions: point, line, rectangle, triangle, linear iteration, and triangle iteration.

For each function the user must input values for variables, select a color, and then click the create button to plot beads onto the grid based on their input. The point function takes numbers for x and y and a color as inputs, and places a bead of the chosen color at the given Cartesian coordinate point. All possible designs can be created by using this function exclusively; however doing so would require one point function call for each of the 1681 cells of the grid. Students can speed up this process with the other functions that can place multiple beads per function call.

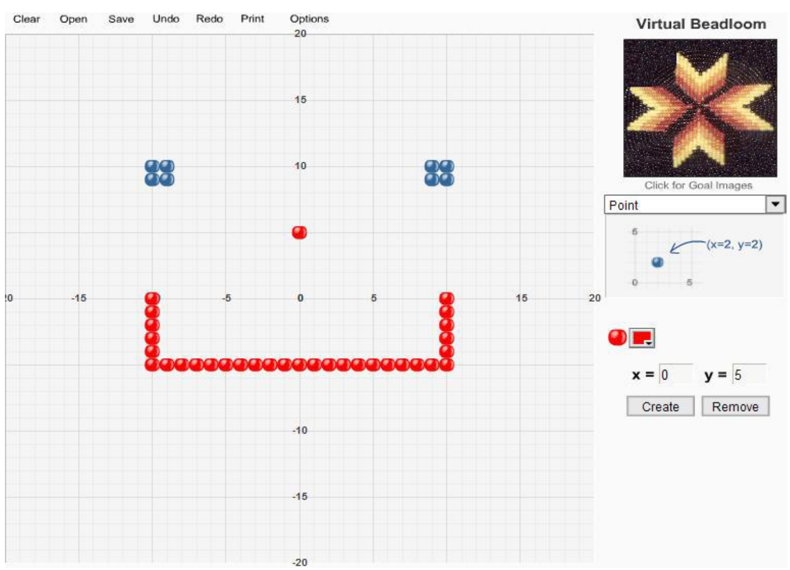


Figure 6. Screenshot of Virtual Bead Loom with a student-created pattern

The Virtual Bead Loom line and rectangle functions each take two Cartesian coordinate points (4 variables) and create a line or rectangle of beads of the given color between those points. Similarly, the triangle function creates a triangle of beads from three points (6 variables). These functions are more complex than the point function, but students tend to quickly learn how they work and use them in conjunction with the point function to make their own art. Ideally, the students would transition from these basic geometric functions into using the more advanced iterative functions.

As the name suggests, the iteration functions use iteration to place beads in complex patterns defined by variables A-I that are set by the user. The linear iteration function (Figure 7 Left) begins with a straight line of beads of a given length A at a given point (B.x, B.y). It then moves to an adjacent space in the specified direction and redraws the line adding or subtracting the specified number of beads from each side of the line (C and D) for E lines. The triangle iteration function (Figure 7 Right) begins with a single bead at a given point (F.x, F.y). Then, for G rows, the tool H beads to each side. It does this for I rows and in the specified direction.

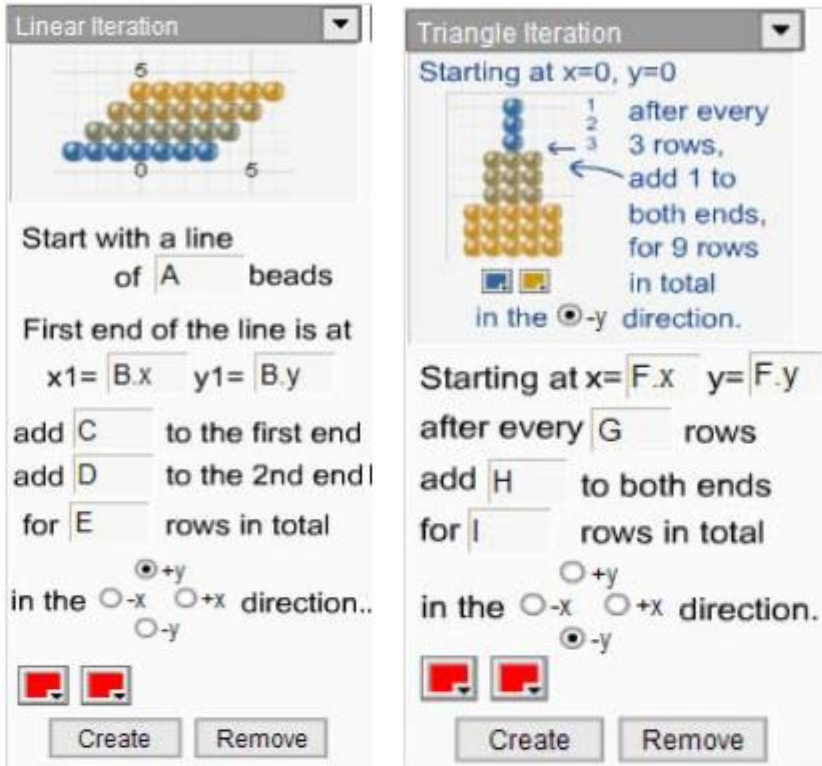


Figure 7. Linear and triangle iteration in the Virtual Bead Loom

The Virtual Bead Loom also uses layering with all functions. For example, if a user places a black bead at position (1, 1) and later places an orange bead at (1, 1) only the orange bead will be displayed. The Virtual Bead Loom does not blend bead colors. This can be seen as a simple implementation of the Painter's Algorithm, a key concept in computer graphics that calls for drawing objects in the distance before drawing closer ones. In other words, to create a background, it would be easier and more efficient to start with a single rectangle function call to cover the entire grid. If the user starts with the foreground design elements they will be forced to use multiple functions to place the background around the rest of the image. The use of more efficient functions and incorporating layering into the creation of designs helps introduce users to the concepts of efficiency and optimization that are corner stones of Computer Science.

BeadLoom Game Prototype

The Virtual BeadLoom was gamified into BeadLoom Game with the goal of creating a system that was fun and encouraged students to explore all of the system's learning objectives. The first step was to analyze what the Virtual Bead Loom did well. The Virtual Bead Loom was very effective at teaching the basic concept of the Cartesian coordinate system (Eglash, et al., 2006). This was possibly because the very act of creating the designs with the functions required understanding and mastering the Cartesian coordinate system. Students practiced this learning objective at every step of the creation process; thus learning was tied to, and not separate from, creation.

We wanted to integrate this successful element into the gamified version by ensuring there was **no separation between the learning element and the game element.** This solves the problem many edutainment titles suffered from by preventing the learning from being an obstacle to play. We wanted learning to occur from playing the game and not as a requirement to be fulfilled before being allowed to play the game. This is accomplished in BeadLoom Game by using the same six graphing functions (Point, Line, Rectangle, Triangle, Linear Iteration, and Triangle Iteration) as the Virtual Bead Loom. By using all six graphing functions students would be exposed to the learning objectives while they played. This also meant that BeadLoom Game would maintain all the interactions and successful elements of the Virtual Bead Loom. Skills from one system would be directly transferable from one to the other.

Where the Virtual Bead Loom had problems was getting the students to use all six functions. The system did not provide any reason to use the more complex strategies like iteration, optimization, or layering. Thus students could engage in the less desirable activity of only using the point and line functions or the desirable activity of using all the functions to create their designs in an optimal fashion. We hypothesized that **points and achievements** could be used to encourage full exploration of the system. However awarding points and achievements for superfluous activities would not guide user behavior. Therefore we needed to **integrate the points and achievements with learning objectives and desired behavior.** BeadLoom Game needed to reward players for engaging in desirable behavior and mastering the learning objectives.

The point system we designed had players recreating a given goal image in the fewest function calls or “moves” possible. Their score was how many moves it took them to solve a given goal image or “puzzle” with a lower score being better. Although every puzzle can be solved with the less desirable behavior or only using the point and line functions, the only way to get a better score is to engage in the desired behavior and master the learning objectives that are associated with that behavior.

Identifying the less desirable and desirable behaviors was key to designing this point system. Without this identification it would be impossible to use it to guide user behavior and thus it would lose its effectiveness as a tool for improving learning. Points could not just be randomly given out and needed to guide player behavior ensuring the only way to improve in game performance is through learning the material and performing the desired behavior.

An alternative design we could have used was to not allow users to use the simple functions and only allow users to use the iterative functions. In other words we could have prohibited the less desirable behavior and only allowed the desirable behavior. This was not ideal because less desirable behavior is still desirable behavior and thus should not be prohibited. In BeadLoom Game using the point and line functions is helping the students learn Cartesian coordinates which is one of the learning objectives. Additionally starting with the complex iterative functions would create a steep learning curve that would drive away many students. Understanding the Cartesian coordinate system is an important first step to using the iterative functions. We wanted players to start with the less desirable behavior. Once they are ready they can choose to transition to the more desirable behavior with the point system encouraging them to make that transition. This also increases the chances that the students will be able to identify the BeadLoom Game activity as play because choice and freedom are key elements to play (Brown, 2009; Caillois 2005). Therefore we felt it was important to **not prohibit less desirable behavior** and instead let the point system simply discourage it.

Before spending the time to develop BeadLoom Game, it was important determine whether the designed game mechanic and point system was fun and could potentially guide user-behavior. To accomplish this, a mixed-fidelity prototype combining the Virtual Bead Loom with paper and pencil worksheets was created. By integrating the pencil and paper version with the functioning custom puzzle mode the system could avoid the traditional problems with pure paper prototypes like abstractness and response lag (Lim et al., 2006). The worksheets had a goal image or “puzzle” created using the Virtual Bead Loom. Users were asked to recreate the image using the Virtual Bead Loom and record the functions used on their worksheets. This design has many drawbacks. Users have to manually log every move they make including the values for all variables. In addition, the user must determine when the puzzle is finished. We hypothesized that, if users still enjoy playing BeadLoom Game even with these shortcomings it would indicate that this game objective is effective and a full software version should be developed.

Initial Trial

We ran a simple trial to get player feedback on this prototype version of BeadLoom Game in a weekly after school apprenticeship program called Citizen Schools that lasted ten weeks. We worked with a group of eight middle school students who had voluntarily enrolled in a program designed to teach them how to create video games using the Game Maker software. At the beginning of the first session we walked the students through a 15-minute tutorial on the functions of the Virtual Bead Loom. At the start of sessions two through nine we had the students use the mixed fidelity prototype to run small BeadLoom Game competitions to see who could finish the given puzzle in the fewest number of moves.

We observed that students seemed to enjoy this opening activity and were attempting to use the advanced functions. We also asked for feedback on the game as they played. The general response was that the game was fun but that writing down the moves was “annoying.” Additionally students reported that it was difficult to know where individual shapes started and ended on the paper print out since there were no gridlines as shown in Figure 8. Due to time constraints of the apprenticeship we were unable to have a BeadLoom Game contest in the final session. This however gave us a positive indication about the fun of the prototype because many of the students requested to have a BeadLoom Game contest at the start of this session. Based on the results of this preliminary testing we began development of BeadLoom Game 1.0.

Figure 8. Example prototype puzzle with no gridlines

BeadLoom Game 1.0

BeadLoom Game 1.0 was developed by adding additional features to an existing Java implementation of the Virtual Bead Loom. The user is presented with the same 41 x 41 grid on which to plot virtual beads of various colors using the same six functions as in the Virtual Bead Loom: point, line, rectangle, triangle, linear iteration, and triangle iteration. A second grid was added with a completed design on it that acts as the goal image from the prototype. On both grids the user can hover over a point to see the Cartesian coordinates of that point. This prevents the problem we had in the prototype of not being able to see the grid lines and having to count individual beads to get the Cartesian coordinates. We also added a menu where users can change puzzle, restart the current puzzle, submit their solution for scoring, and a counter that tracks how many moves have been used so far. Figure 9 shows a screenshot of the game with an American flag as the goal image on the left and a blank solution grid on the right.

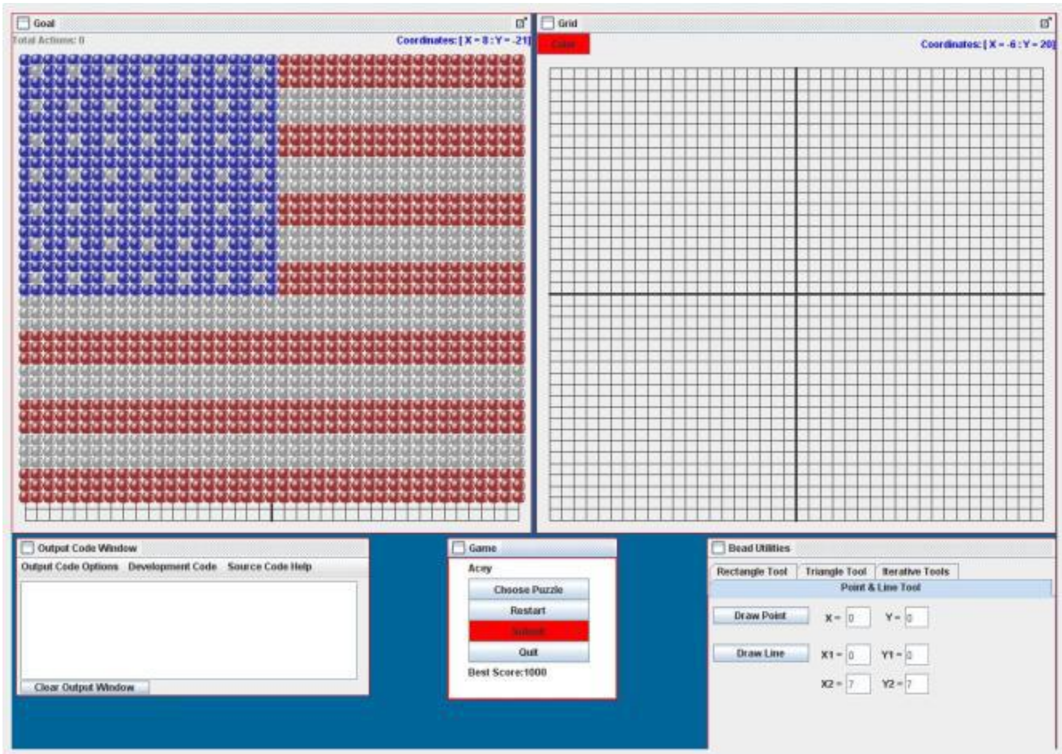


Figure 9. BeadLoom Game 1.0 screenshot

Like the prototype, the core game mechanic of BeadLoom Game 1.0 is the recreation of the given goal image or “puzzle” using the six available functions. Using any of the functions counts as a “move” and the goal is to solve each puzzle in the fewest moves possible. Every puzzle has an “ideal solution”: the lowest number of moves needed to complete the puzzle. When the player clicks submit the game compares their work to the goal image and pinpoints the locations of differences. This system solves the problem of the user having to manually check solutions and locate mistakes that was in the prototype.

To encourage finding the ideal solution, when the player matches the goal image exactly they are awarded a medal based on their performance. If the player finds an ideal solution they receive a platinum medal. A gold medal is awarded for solving it in up to two times the ideal solution, silver is awarded for up to three times the ideal solution and for just completing it they receive a bronze medal. For example, if a puzzle can be finished in 10 moves, a gold medal means the player finished in 20 or less moves, a silver medal means finishing in 30 or less moves, and a platinum medal means finishing in 10 or less moves. Or less is used for platinum because an algorithm for solving these puzzles was not developed. The best solution our developers and designers found is used as the ideal. If a student does find a solution better than the current best solution then the ideal solution is updated accordingly. This acts as a form of **tiered achievements**. The goal of this tiered structure was to help prevent discouraging the player by lowering the learning curve. If there was only one correct solution and the player could not find it they may become discouraged and not try again. The tiers allow them to explore less ideal solutions while showing them that more optimal solutions exist. This provides incentive to replay completed puzzles to get a better score and by extension of the integrated learning objectives better learn the material.

This implementation is a slightly modified version of traditional gamification. It has points and the medals can be thought of as badges and achievements, tied to the central learning objectives. It is lacking leaderboards. The reason for this is because when this early study and system were being designed gamification was a relatively new term and the specific elements that contributed to successful gamification had not yet been as firmly defined. The goal was not gamification but the conversion of the tool into a game. It was not until later that it was realized that this process had essentially been gamification. Using the Gameplay/Purpose/Scope model (Djaouti, 2011) this version of BeadLoom Game would be classified as a game-based system.

Study One: BLG (Boyce & Barnes, 2010)

R1: Can a traditionally gamified educational system produce learning gains?

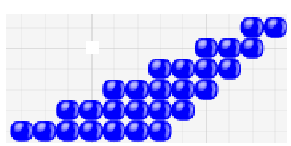
The goal of the first study was to determine if a traditionally gamified system could generate learning gains - an important starting point. Prior to this study there had been a lack of rigorous research on educational games and gamified systems that utilized pre and posttest analysis of learning objectives. Much educational games related research has focused on user preference. While user preference is an important measure, the ability to generate learning gains is extremely important for educational software. Previous work also did not focus on a minimal gamification approach of adding only points and achievements. It was also important to determine if the game elements and design decisions made for BeadLoom Game were effective before moving forward to deeper research questions. If there was a fundamental flaw in the system it was important to identify it early. We hypothesized that **the traditionally gamified system BeadLoom Game will generate learning gains in the areas of iteration and layering.**

Study One: BLG - Methods

For this study data was collected during a high school summer camp held at UNC Charlotte in 2009 with 18 students: 16 male and 2 female. This summer camp was a five-day video game and computer forensics camp where half the day was devoted to learning computer forensics and the other half to learning how to develop video games using the Game Maker software. BeadLoom Game was integrated into the video game portion of the camp as an example of an educational video game.

On the first day of camp we gave a one-hour introduction to BeadLoom Game. During the introduction, the students were walked through six tutorial puzzles. Each tutorial puzzle was solved using only one of the six functions and was designed to illustrate how one particular function works. Each day, 30 to 60 minutes were dedicated to an activity called the BeadLoom Game contest. During this activity the students would race to see who could complete an assigned puzzle in the fewest moves. Students who found a solution would get to write their names on the board along with their score. The winner of the event would be the student(s) who found the ideal solution and if no one found the ideal solution then whoever found the best solution the fastest. Although our original design for BeadLoom Game did not have the traditional gamification element of leaderboards, having the students write their names on the board acted like a leaderboard for the purposes of the study. It also allowed us to gauge the impact of a leaderboard to see if it should be added to BeadLoom Game.

At the beginning of the week before engaging with BeadLoom Game the students were given a short pretest (Appendix A) to gauge their understanding of Cartesian coordinates, iteration, and layering/optimization. The pretest consisted of 12 questions with 4 questions devoted to Cartesian coordinates, 4 questions devoted to iteration, and 4 questions devoted to layering and optimization. The Cartesian coordinates questions asked the students to draw given points on a graph with one point in each of the four quadrants. Two iterative questions asked the user to fill in the blank to create a given iterative pattern (Figure 10) and two asked the students how many beads would be present after iteratively adding beads for a certain number of steps. For the fill in the blank questions with multiple parts, partial credit was given for each correct part. The layering and optimization questions asked how many shapes it takes to recreate a given image. At the end of the five-day summer camp students were given an isomorphic posttest.



This shape was created using iteration.
It can be described by the iteration pattern:
Start with 7 beads

Every line add _____ bead(s) to the right and

add _____ bead(s) to the left

for _____ lines in total.

Figure 10. Example iteration question

We were forced to remove the questions regarding Cartesian coordinates from the analysis in this study. As mentioned this study was conducted at a summer camp dedicated to learning how to create video games using the Game Maker software. Game Maker uses a different coordinate system than the traditional Cartesian system with the point (0, 0) being the top left of the grid instead of the middle. As a result some of the students believed they were supposed to use the Game Maker coordinate system to answer the Cartesian coordinate questions. No elements of Game Maker we utilized during the camp pertained to iteration or layering so the changes from pretest to posttest for these questions were the result of playing BeadLoom Game.

Study One: BLG - Results

With the remaining questions we submitted the student's pretest and posttest scores, shown in Figure 11, to a 2-tailed matched paired t-test. Students performed significantly better, $t(17) = 3.29$, $p = .005$, $d = .77$, on the post test ($M = 3.57$, $SD = 1.14$) than on the pretest ($M =2.98$, $SD = 1$). Cohen's d indicates that students improved from pretest to posttest by .77 of a standard deviation.

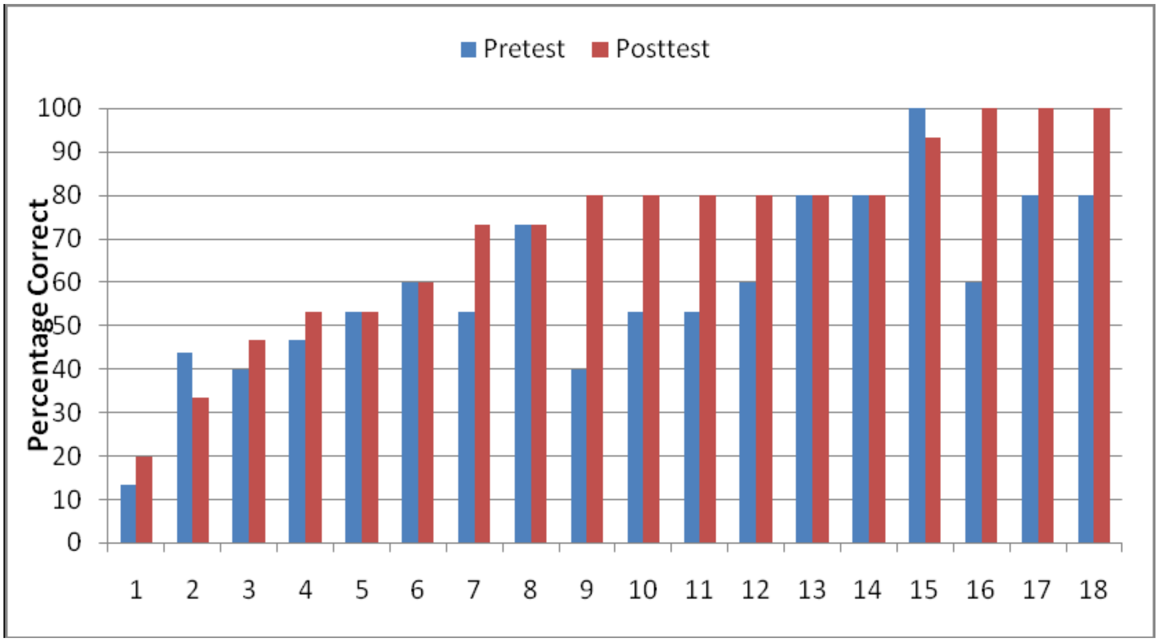


Figure 11. Study one pre to posttest scores for all participants, sorted by posttest score

In addition to the quantitative pre and posttest analysis we collected a few qualitative measures. On the posttest students were asked to rate the fun of BeadLoom Game on a scale from one to five with five being the most fun and one being the least fun. The students gave BeadLoom Game an average of 4.28 out of five. Additionally two of the participants requested a copy of the game on a CD to take home to play. They were committed to getting a platinum medal on all the puzzles in the game. This suggests that in addition to being educational this version of BeadLoom Game was fun for at least a majority of the students and extremely fun for at least two.

These results show that the BeadLoom Game can generate learning gains in the areas of iteration and layering. This confirmed our hypothesis and shows that **the traditionally**

gamified system BLG can generate learning gains. Additionally our subjective measures suggest that BLG can achieve these learning gains in a way that is fun for a majority of participants. BeadLoom Game was a successful piece of educational software but further evaluation was necessary to compare it to its predecessor the Virtual Bead Loom.

Study Two: BLG Vs VBL (Boyce et al., 2011a)

R2: Can the use of a traditionally gamified system outperform a non-gamified educational tool in the areas of learning gains and user preference?

The goal of the second study was to compare the learning gains from the traditional educational software Virtual Bead Loom and the gamified educational software BeadLoom Game. This study was also designed to determine if the students preferred to use the Virtual Bead Loom or the gamified BeadLoom Game. The goal was to see if the use of a traditionally gamified system can outperform a non-gamified educational tool in the areas of learning gains and user preference. We hypothesized that **the gamified system BeadLoom Game will outperform the non-gamified system the Virtual Bead Loom in the areas of iteration, layering, and user preference.** We needed to compare learning gains from both systems to test this hypothesis.

Study Two: BLG Vs VBL - Methods

For this study we used the same version of BeadLoom Game as in study one. We conducted experiments during two five-day summer camps in 2010. The first camp focused on the use of Culturally Situated Design Tools and had 21 middle school students with six females. The second camp was the 2010 version of the high school camp we used in study one which focused on computer forensics and game design. That year it had 20 students with seven females.

We used a “switching replications” study design as used by Eagle and Barnes (2009). This design involves dividing the participants into two groups: control and experimental. After exposing the groups to their condition participants are evaluated and then the groups are switched exposing the control group to the experimental condition and vice versa. This allows all participants to experience both conditions allowing them to provide feedback on things like preference comparing the two conditions.

At the start of both camps we gave the students a pretest to gauge their knowledge of Cartesian coordinates, iteration, and layering. We utilized the same test format as in study one with the exception of making it clear that the coordinates questions were not using the Game Maker coordinate system for those students in the high school summer camp.

After the pretest each camp was divided into two groups: the Game (experimental) group played the BeadLoom Game first and the Tool (control) group used the original Virtual Bead Loom tool first. Both groups were walked through a short demonstration of each of the six functions and then given 90 minutes to play BeadLoom Game or use the Virtual Bead Loom depending on their group. Afterwards, they were given an isomorphic posttest to evaluate their learning gains from using the system.

After using the Game or the Tool respectively, each group used the other software for 90 minutes. Following this activity, the middle school camp was introduced to other Culturally Situated Design Tools, while the high school camp learned how to use Game Maker. Since none of the activities with Culturally Situated Design Tools or Game Maker used iteration or layering, BeadLoom Game would be the only source of learning in these areas. A clear distinction between Cartesian coordinates and the coordinate system used in Game Maker was also made with the high school camp.

Each day after the first day began with a 30-minute BeadLoom Game contest like the ones held in study one with students competing to find the best solution for a given puzzle. Additionally the middle school group was given 30 to 60 minutes at the end of the day where they could choose from a selection of activities including BeadLoom Game, the Virtual Bead Loom, and the other Culturally Situated Design Tools introduced thus far in the camp. Each day one to two additional Culturally Situated Design Tools would be introduced except for the first day where only BeadLoom Game and the Virtual Bead Loom were used. Their selections during this time were recorded. At the end of each camp the students took a second isomorphic posttest (post2) to evaluate overall learning gains, and a short survey to determine which system they preferred to use.

Study Two: BLG Vs VBL - Results - Middle School Camp

There were no statistically significant differences between the Tool and Game groups in the middle school in the area of Cartesian coordinates. This is not surprising, as

BeadLoom Game and the Virtual Bead Loom use the same basic functionality to plot Cartesian coordinates. The success of the Virtual Bead Loom at teaching Cartesian coordinates is one of its strongest features and is one of the reasons it was selected for gamification.

Figure 12 shows the middle school students' pre, post, and post2 test scores on the eight remaining questions on iteration and layering. The Game group showed increases in knowledge during the first session with their system and continued to improve through the remainder of the summer camp. The Tool group showed no learning in their first session with their system but after being exposed to BeadLoom Game showed significant learning gains.

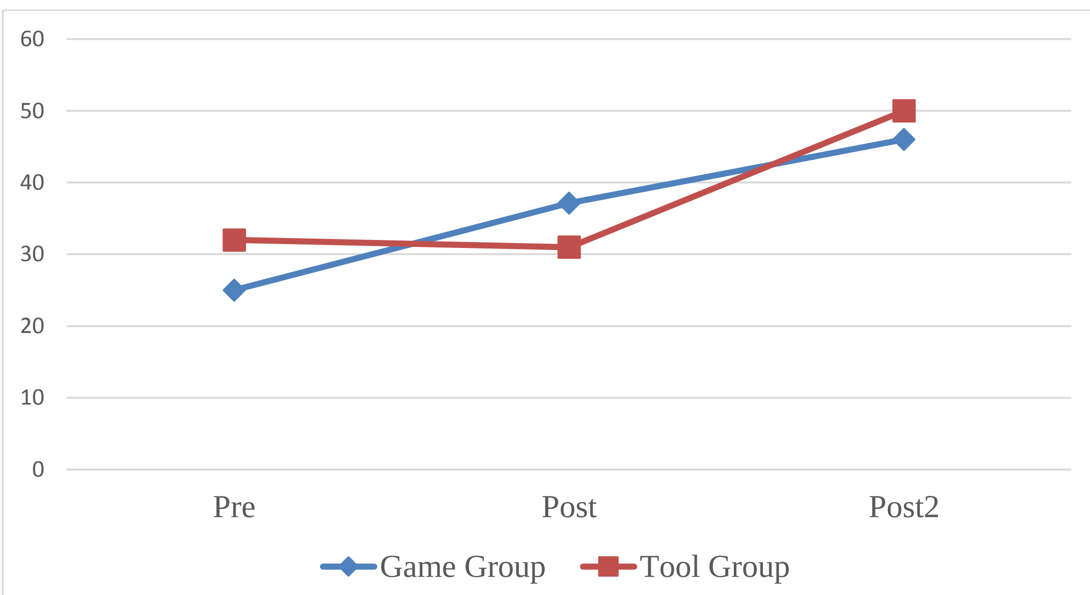


Figure 12. Average test scores (%) of middle school camp Game and Tool groups

Table 2 compares the changes in test scores for each group for the iteration and layering questions for the middle school camp. There was a statistically significant difference in the test scores changes from pre to post between the Game and Tool groups. This indicates that the Game group learned more from using the gamified system for 90 minutes than the Tool group did from the non-gamified system in the same amount of time. There was not a statistically significant difference in change between the two groups by the end of the camp indicating that the Tool group was able to reach the same level of learning gains as the Game group through exposure to BeadLoom Game.

Table 2. *Average percentage test score changes between middle school groups*

	N	**Post-Pre**	**Post2-Post**	**Post2-Pre**
Game	11	12.12	8.33	20.45
Tool	10	-1.67	18.75	17.08
Difference p value		.019	0.27	0.70
t-stat		2.57	-1.12	0.38

Table 3 shows the number of students who selected BeadLoom Game, the Virtual Bead Loom, and other Culturally Situated Design Tools for their activity during the end of day choose-your-own activity part of the camp Monday through Wednesday. This activity was not done on Friday due to a presentation to the student's parents and not done on

Thursday to allow for preparation time for the Friday presentation. Each day BeadLoom Game was the most popular choice.

Table 3. *Number of students who chose each activity by day*

	Monday	**Tuesday**	**Wednesday**
BeadLoom Game	16	19	13
The Virtual Bead Loom	5	0	0
Other	-	2	8

When asked which system they preferred between BeadLoom Game and the Virtual Bead Loom 16 selected BeadLoom Game and 5 selected the Virtual Bead Loom. There was also an area for the students to explain their preference. One student responded that he liked BeadLoom Game best “because we have to solve (puzzles) using our brains.” Over the course of the five-day summer camp the 21 middle school students earned a total of 183 Platinum medals.

Study Two: BLG Vs VBL - Results - High School Camp

Similar to the middle school camp, there were no statistically significant differences between the high school Tool and Game groups for scores on the Cartesian coordinate questions. The high school students in both groups had high scores on the pretest for this subject and had little change from pre to post and on to post2. These students already had a strong understanding of the Cartesian coordinate system.

Figure 13 shows the high school students' pre, post, and post2 test scores on the iteration and layering questions. Similar to the middle school group, the Game group saw improvements from pre to post and continued to improve into the post2 while the Tool group saw no improvement from pre to post. It was not until after these students were exposed to BeadLoom Game that their scores improved.

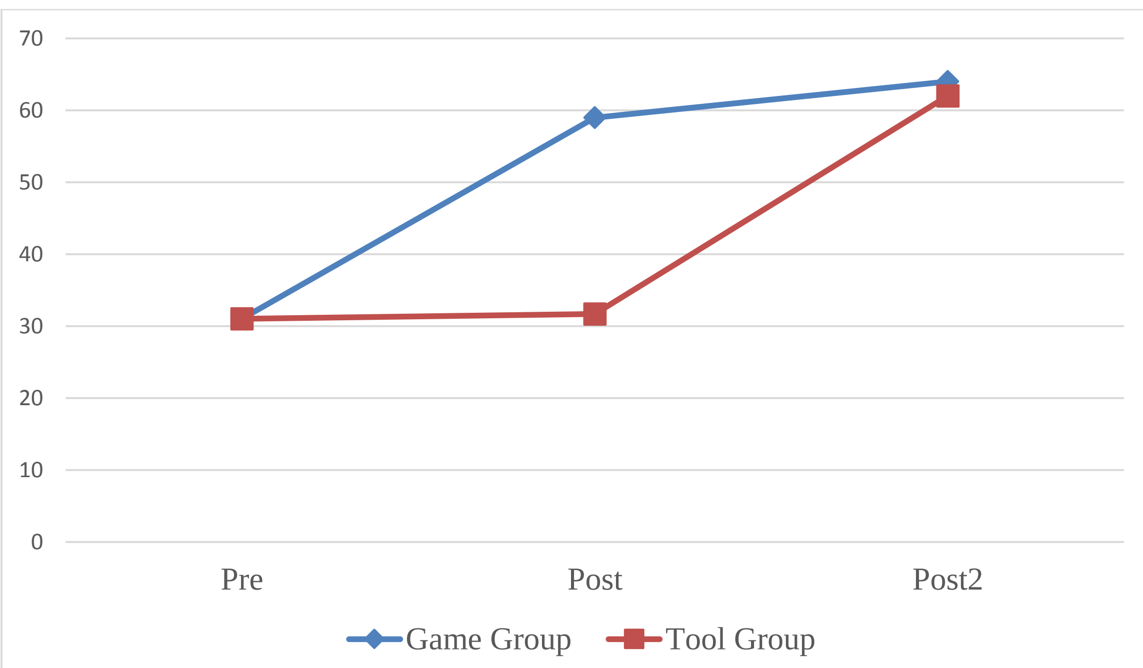


Figure 13. Average test scores (%) of high school camp Game and Tool groups

Table 4 compares the changes in test scores for each group for the iteration and layering questions for the high school camp. As with the middle school camp there was a statistically significant difference in the test scores changes from pre to post between the Game and Tool groups. Once again the Game group had a significantly larger change in test score resulting from the 90-minute session with the gamified system compared to the group that used the non-gamified system. After exposure to BeadLoom Game there was no statistically significant difference between the groups.

Table 4. *Average percentage test score changes between middle school groups*

	N	Post-Pre	Post2-Post	Post2-Pre
Game	10	27.50	8.33	35.83
Tool	10	.083	30.42	31.25
Difference p value		.0012	0.27	0.63
t-stat		2.57	-1.12	0.49

Comparing the middle school and high school scores reveals one more interesting observation. The middle school Game group outperformed the high school Tool group on the post test, averaging 37.12% on the iteration and layering questions while the high school Tool group averaged 31.67%. Although this is not a statistically significant difference, it is surprising to see the middle school students outperforming high school students after playing BeadLoom Game.

When asked which they preferred, 13 of the high school students chose the gamified BeadLoom Game and 7 chose the non-gamified Virtual Bead Loom. In the free response section one student wrote “the game forces me to use more complicated iterations to improve my score which makes it a fun activity and a good learning tool.” During the five-day camp the high school students earned 215 Platinum medals.

Study Two: BLG Vs VBL - Discussion

Prior to this research there was a lack of rigorous research comparing learning gains in educational software and a traditionally gamified version of that software. Using a switching replications experimental design this study was able to evaluate the learning gains resulting from playing both the original Virtual Bead Loom and the traditionally gamified BeadLoom Game while also allowing the students to select which system they preferred. In both the middle school and the high school groups the gamified BeadLoom Game generated higher learning gains from 90 minutes of play. Once the students that began with the Virtual Bead Loom had used BeadLoom Game they showed similar learning gains. In addition in both age groups a majority of students reported preferring BeadLoom Game. This verifies the hypothesis and shows that **the gamified system BLG outperforms the non-gamified system VBL in the areas of iteration, layering, and user preference.**

Although these results are a strong indicator of the potential strength of traditional gamification the results also illustrates one of its major weaknesses. While 70.7% of the students preferred BeadLoom Game this also means that 29.3% of the students preferred the original system. Analyzing these students' free responses suggests two major causes for this preference.

Some students stated they preferred the "creative freedom" of the Virtual Bead Loom. They stated they did not want to make the puzzles we had designed and instead wanted to create their own unique and interesting designs. They lacked the freedom to make a choice and thus were unable to see the activity as fun. The rules and game-based nature of BeadLoom Game were not appealing to these students.

The other common cause was a dislike of the competitive aspect of the BeadLoom Game contests. They did not enjoy having their performance compared to those around them especially when they underperformed. While some students praised the leaderboard element others cited it as a cause of discouragement and the reason they preferred the Virtual Bead Loom. The competitive element central to traditional gamification and a major component of game-based systems was encouraging some but strongly discouraging others.

Although BeadLoom Game was effective at teaching it was discouraging to some users. This is problematic because an educational system should be effective for as many users as possible. One of the biggest advantages of gamification is making users want to engage with the system and the traditional method does not do this for some users. BeadLoom Game and gamification needed elements that can appeal to users who prefer the less competitive and more creative environment of the Virtual Bead Loom. This realization led to the development of BeadLoom Game 2.0.

BeadLoom Game 2.0

Based on the results of the first two studies there were four very important elements that BeadLoom Game 2.0 needed. The first was that rather than a standalone program BeadLoom Game needed to be changed to run in a web browser. This allows easier access to the game and enables students who enjoy it to play it at home after the study. It also makes it easier to log player usage behavior.

The second element was the integration of **leaderboards** into the system. Rather than continuing to approximate the effect of leaderboards through the manual recording of participant scores BeadLoom Game 2.0 features leaderboards for every puzzle. It is ordered by moves used to solve the puzzle with time taken as the tiebreaker. Now BeadLoom Game would encourage competitive students to replay individual puzzles when their score is beaten by friends similar to what was seen in the first two studies with tracking names on the white board and groups of players playing together. There is also a leaderboard for “most medals earned” to encourage players to try more puzzles for a breadth of play. These were clear improvements that were supported by traditional gamification approaches.

The third element of **unlockable** content was relatively minor to implement but is important to note. The puzzles in BeadLoom Game 1.0 were divided into four levels of difficulty: Tutorial, Easy, Medium, and Hard. Tutorial puzzles used only a single function and illustrated the basic shapes each function could make. Easy puzzles combined the basic functions with some layering, medium puzzles required more iteration and layering, and hard puzzles required creative use of iteration and layering to solve optimally. The Hard puzzles were so challenging even college students who worked in our research lab found it tricky to get platinum medals. These Hard puzzles were designed specifically to require mastery of the material and address a common complaint of the original Virtual Bead Loom that “you cannot make anything cool with this.” Figure 14 shows both an easy and a hard puzzle. The easy puzzle can be solved in six moves and the hard one can be solved in 24 moves. However some students in the first studies would see the challenging but cool hard puzzles and attempt to play those levels first. Of course they did not possess the skills or knowledge to conquer them and became frustrated. These students were encouraged to start with the easier puzzles before moving to the hard ones. To prevent students from attempting content they were not prepared for a simple **unlock system** was added. Students could see all the puzzles from the beginning but had to complete all the puzzles of the previous difficulty to be able to play them. This meant the students would have to start with the tutorial puzzles, move to the easy puzzles, and then solve the medium puzzles before they could attempt the hard ones.

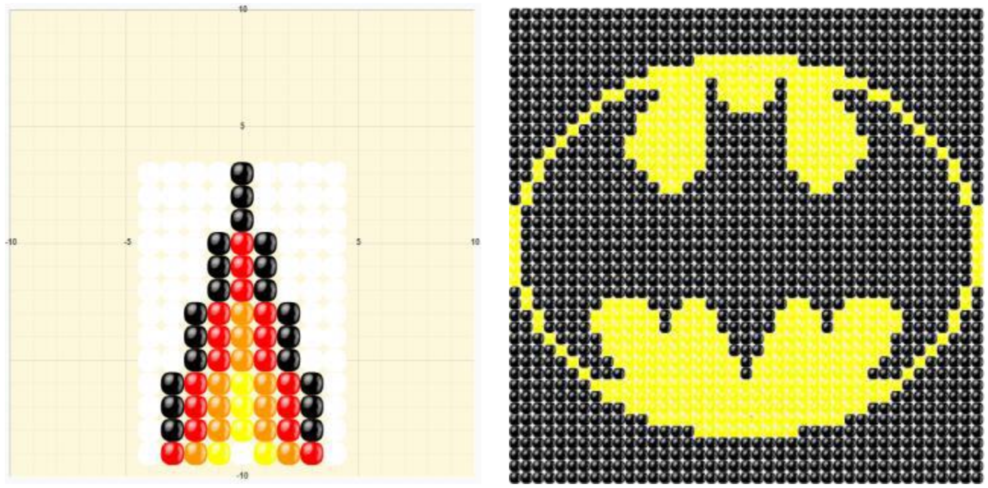


Figure 14. Easy (left) and hard (right) BeadLoom Game puzzles

The fourth element of **creative user-generated content** was needed to encourage use from students who disliked the game-based elements of BeadLoom Game and preferred the more play-based Virtual Bead Loom. The user-generated content or “Custom Puzzle” mode was created for these users. In this mode, users are given a blank puzzle grid and no objective grid similar to the interface of the Virtual Bead Loom. The user is then free to create any design they want. This play-based environment lacks objectives assigned by the system and leaves goal setting to the user. Unlike the original tool, players can publish their finished work online. Published puzzles can be played like the built in puzzles and have high score boards for each puzzle. These puzzles are also displayed in the “Custom Puzzle Showcase.” Here users can rate each custom puzzle from one to five stars with five stars being the best and one star being the worst.

At first glance, it seems like this Custom Puzzle mode would suffer from the same problems as the Virtual Bead Loom: players avoiding the complex functions in favor of creating designs point by point. Two game elements were added to discourage this behavior. These custom puzzles are displayed in the showcase in order of ranking. If a user wants their design to be actively viewed and appreciated by the most users it is in their best interest to make their designs complex and appealing. Simple or mundane designs should be quickly down voted and slip to the lower pages. Custom puzzles are also sorted by their creators when choosing a custom puzzle to play. If a user wants others to play and enjoy their work they must become known for creating tricky and complex designs. This should typically involve incorporating elements such as iteration and layering into custom puzzles. We believed this social element would overcome the issues of the Virtual Bead Loom and would allow for a play-based mode that appeals to users that originally preferred the Virtual Bead Loom. Combining it with the game-based puzzle-solving mode should result in a system that can appeal to, engage, and teach a wider range of users.

This combination has an additional benefit. It dramatically adds to the replayability of the game. In BeadLoom Game 1.0 if a player solved all the puzzles with a platinum medal they were essentially done. They could replay the same puzzles again for review but that is unlikely. With BeadLoom Game 2.0 there is a constantly growing number of puzzles, so the user can never run out. This is very important for an educational system, as replayability is an important component of successful educational games (Prensky, 2005).

Incorporating user-generated content into an educational game can result in a content creation and learning loop. As more competitive users complete more puzzles they will desire more complex and challenging puzzles which will push puzzle designers to come up with even more tricky and complex puzzles which push the limits of layering and iteration. This phenomenon of ever improving user-generated content has had great commercial success in games such as Little Big Planet. To date Little Big Planet 2 has had over 8.5 million levels created by its users (lbp.me). This improved replay value and content creation loop has great potential for improving educational systems and gamification.

BeadLoom Game 2.0 was a large improvement over 1.0. To determine the impact the integrated online leaderboards and the user-generated content creation environment had it was important to be able to enable and disable features during user testing. Therefore all these additional features were flagged with Booleans that enabled them to be turned on and off. Accounts on the website used to host BeadLoom Game would be flagged to enable and disable features on a per user basis as necessary. This allowed BeadLoom Game 2.0 to take on four different modes each giving different options to the user:

- Base BeadLoom Game - No integrated leaderboards or user-generated content. Essentially BeadLoom Game 1.0 only online.
- Leaderboards – Integrated leaderboards enabled but user-generated content disabled
- Custom Puzzle – User-generated content enabled but integrated leaderboards disabled
- Full BeadLoom Game – Both user-generated content and leaderboards enabled

Study Three: Custom Puzzles (Boyce et al, 2011b)

R3: Does a gamified system featuring both game-based and play-based game elements outperform a traditionally gamified system in the area of user preference?

The goal of the third study was to investigate the impact the new features in BeadLoom Game 2.0 have on user preference and ensure that these features do not detract from the learning gains of the system. More specifically this third study on custom puzzles investigates which of the previously listed modes are most preferred by the users. If these new features result in more users preferring the gamified system to the non-gamified system then that provides evidence that these elements should be incorporated into gamification approaches. The goal is to have as many users as possible prefer the gamified system to maximize the number of users who may engage with and learn from the system. We hypothesize that **BeadLoom Game 2.0 (featuring both game-based and play-based features) will outperform the traditionally gamified BeadLoom Game 1.0 and the non-gamified Virtual Bead Loom in the area of user preference.** Additionally BeadLoom Game 2.0 featuring only competitive leaderboards or user-generated content will outperform BeadLoom Game 1.0 with neither of these features.

Study Three: Custom Puzzles - Methods

This study was conducted with a total of six classes of middle school students in the fall of 2010. The school followed an A-day, B-day schedule where students alternate their courses. The students were divided into two groups based on if they were in the class on an A-day or B-day. The classes we worked with were required courses for the students. The sixth grade class was a keyboarding course and the seventh and eighth grade classes were a career and technical education course. Both A-day and B-day had a class at all three grade levels. In total the A-day group had 33 males and 37 females and the B-day group had 37 males and 34 females. Since this study was being performed in the classroom we had access to a larger and more diverse group of students than in the on-campus summer camp setting. This also meant more complications including student absences, teacher absences, and fire alarms. Students were allowed to access BeadLoom Game outside of class but only to those features that had been introduced in the study thus far.

The original plan for this study was to use an expanded switching replications design to expose all students to the different modes of BeadLoom Game 2.0. On the first day both groups were given the same pretest used in previous studies to gauge their understanding of Cartesian coordinates, iteration, and layering. They were then walked through a 15-minute tutorial on the various functions in the non-gamified Virtual Bead Loom. After the tutorial the students used the Virtual Bead Loom for 30 minutes.

In the next session with the A-day groups we administered an isomorphic posttest (post1) to evaluate the learning that occurred from using the Virtual Bead Loom. Afterward the students used the gamified BeadLoom Game 1.0 with no integrated leaderboards and no user-generated content. This same activity was planned for the B-day session however the teacher was absent and we were unable to engage with the B-day students. This resulted in the B-day group following an abbreviated schedule.

At the beginning of the third session, the A-day group was given another isomorphic posttest (post2) to evaluate the learning from BeadLoom Game 1.0. They were also given a survey asking them which system they preferred, Virtual Bead Loom or BeadLoom Game, and why. This group was then granted access to a version of BeadLoom Game with the user-generated content elements but not the integrated leaderboards. Students were allowed to play with this version and its features for in class 45 minutes. In the B-day group students were introduced to the BeadLoom Game with the integrated leaderboards but not the user-generated content elements. Students were allowed to play with this version in class for 35 minutes. No tests or surveys were administered here because of a fire alarm at the beginning of the session.

The two groups began following the same schedule again with the fourth session. Here both groups were first given an isomorphic posttest (post3) to once again evaluate their understanding of Cartesian coordinates, iteration and layering. They were also given a survey asking which version of the software of those that they had seen they preferred and why. Afterwards the students were given access to the full BeadLoom Game 2.0 with integrated leaderboards and user-generated content. Students were given 45 minutes with this version in class.

The final session was a short follow up session where we administered a final posttest (post4) to evaluate overall learning from the study. The students were also given a survey asking whether they preferred the Virtual Bead Loom or BeadLoom Game 2.0 and why. They were also asked which feature (main game, leaderboards, creating custom puzzles, or solving custom puzzles) was their favorite and which was their least favorite. A summary of this process is shown in Table 5.

Table 5. *Study three session schedule*

Session	A-Day	B-Day
Session 1	Pretest Virtual Bead Loom	
Session 2	Post1 test BeadLoom Game 1.0	- - -
Session 3	Post2 test and survey BeadLoom Game 1.0 with custom puzzles	BeadLoom Game 1.0 with integrated leaderboards
Session 4	Post3 test and survey BeadLoom Game 2.0	
Session 5	Post4 test and survey	

In addition to the surveys and tests a logging system was used to analyze the usage of BeadLoom Game and the Virtual Bead Loom. Within the puzzle-solving environment and in the Virtual Bead Loom, the logs are updated any time a user uses a function or submits their solution. This allowed for monitoring of play behavior outside the classroom. Any break in the log longer than 15 minutes was removed under the assumption that the user had taken a break from the system and later returned. Due to an error in the custom puzzle mode logger only the action of submitting a finished puzzle was logged with time stamps. This meant that incomplete designs and exact play time for this feature could not be tracked. However the functions used to create these designs were logged.

Study Three: Custom Puzzles - Results – Tests

Although the primary purpose of this study was to determine how the new integrated leaderboards and user-generated content features impacted student preference it was important to ensure that these new features did not detract from the learning gains the previous version achieved. Figures 15 and 16 show the average scores for each group (A-day and B-day). These scores are out of the twelve potential points on the tests.

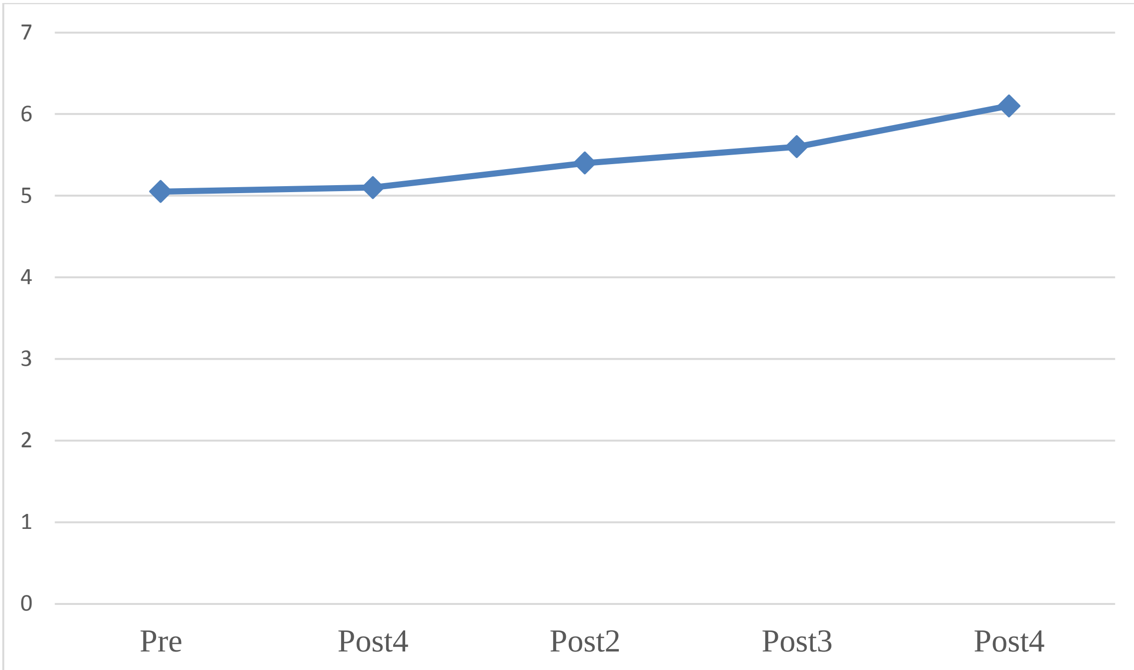


Figure 15. Average A-day group test scores out of twelve

A paired t-test was performed on the data from both groups. The mean grade on the posttest (M=6.12, SD = 1.98, N=51) of the A-day students that completed both the pre and post4 test was significantly greater than the pretest scores, t(50)=2.34, two-tail p = .02. A 95% confidence interval about average score increase is (.09, 1.2). Similarly the mean grade on the posttest (M = 6.48, SD=1.79, N=60) of the B-day students that completed both the pre and post4 test was significantly greater than the pretest scores t(59)=5.68, two tailed p < .001. A 95% confidence interval about average score increase is (.85, 1.77).

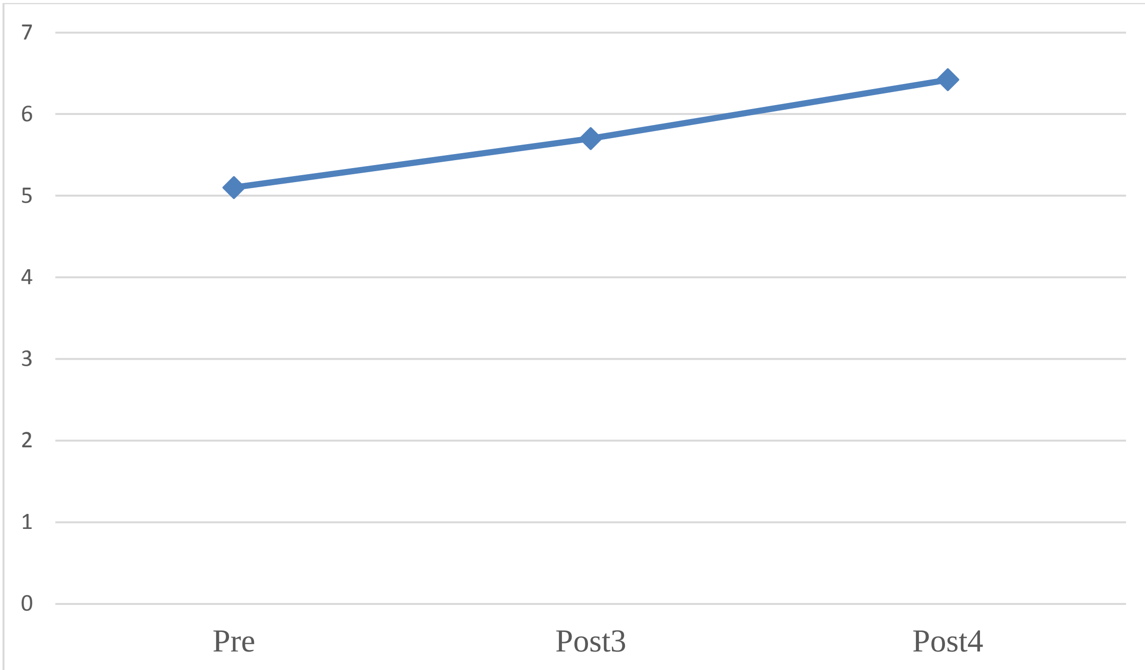


Figure 16. Average B-day group test scores out of twelve

These increases are however not the important element to take away from this study. The tests were given to ensue none of the new features or builds hindered learning. As figures 15 and 16 show, both groups had steady increases in test scores from session to session. The A-day group had a very small increase from pre to post1 after using the non-gamified Virtual Bead Loom. Larger gains occur after utilizing the gamified BeadLoom Game. This study cannot identify which features are contributing to these learning gains but the continual increase in scores from playing the game over the course of multiple sessions does support the need to make gamified systems with as much replay value as possible.

Study Three: Custom Puzzles - Results – Game Metrics

Figure 17 shows the combined logged play time of the students in the A-day group between each test. The logged play time for both the Virtual Bead Loom and the Bead Loom Game was significantly lower for Post2-Post3 compared to Post1-Post2. This was the session when the custom puzzle feature was given to the students. As a result of the bug in custom puzzle mode logging time spent in this mode is not factored into the BeadLoom Game play time. What is important though is the fact that from this point forward there is no play time in the non-gamified Virtual Bead Loom. This suggests that the custom puzzle mode was accepted as an alternative to the original tool in students looking for creative freedom or who preferred more play-based environments.

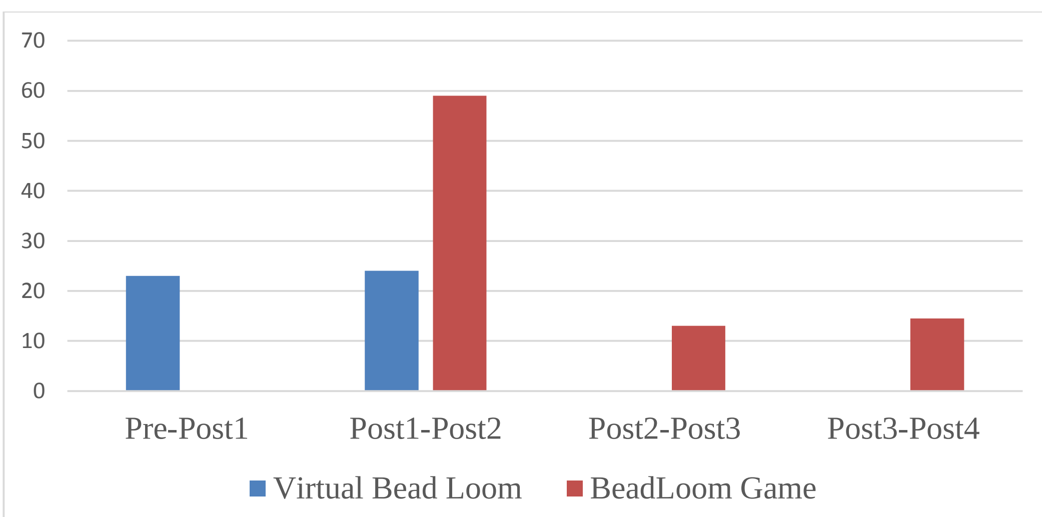


Figure 17. A-day group play time in hours

Figure 18 shows the combined logged play times of the students in the B-day group between each test. Post1-Post3 was the session where the integrated leaderboards were introduced to the B-day group. During this time they did not have access to the custom puzzle mode. Although use of the Virtual Bead Loom dropped during this time it did not go to zero. This indicates that, similar to the previous studies, there were some users who preferred the Virtual Bead Loom to the traditionally gamified BeadLoom Game. However just like in the A-day group once the custom puzzle mode was introduced play of the Virtual Bead Loom dropped to zero. We also see an increase in play of the logged portion of BeadLoom Game.

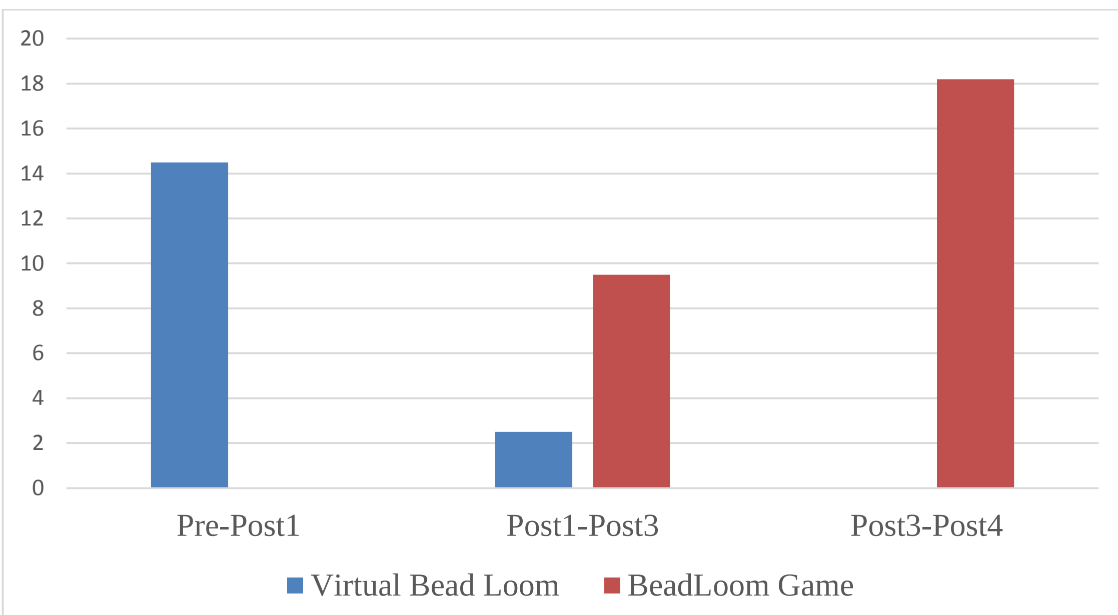


Figure 18. B-day group play time in hours

During the sessions the A-day group created and submitted 69 custom puzzles and the B-day group created and submitted 58. Based on observation of the group many more were started but were either discarded or not submitted. Figure 19 shows some of the highest rated puzzles created by the students, as voted on by their peers in the custom puzzle showcase.

Figure 19. Top rated custom puzzles

In-game logging continued over their two-week winter break in December and January and two weeks of school in January. We were interested in seeing if the new BeadLoom Game 2.0 was engaging enough to have students play it outside the controlled environment of the sessions. During this time the students logged a combined total of over 23 hours of gameplay excluding time spent in the custom puzzle mode. The students submitted an additional 65 new custom puzzles during this time. This shows that both the play-based and the game-based elements had enough appeal for students to get them to play and learn in their spare time even when at home for winter break.

Study Three: Custom Puzzles - Results – Surveys

During session 3 the A-day group was asked which they preferred Virtual Bead Loom or BeadLoom Game and to explain why. Here 28.8% of the students preferred the non-gamified Virtual Bead Loom and 71.2% preferred the gamified BeadLoom Game. A major factor in system preference was preference of creative freedom ("I like the freedom to create what I want") versus directed goals and challenges ("Because it allows me to unlock different thing[s] which make[s] me have a goal"). This parallels the two styles of play and two types of games. Saving some puzzles as unlockable content prevented students from attempting puzzles they were not prepared for, while providing them with content to work towards. This was very appealing to some students.

In session 4, after having the play-based user-generated content environment added the students were once again asked for their preference. Now 86.0% of students preferred the

BeadLoom Game and 14.0% preferred the Virtual BeadLoom. It may seem odd for some students to still prefer the Virtual Bead Loom considering the custom puzzle mode had many similar features but some of the students provided legitimate explanations. Some students were concerned with their art being judged harshly by others. This is similar to the fear expressed by some students in previous studies regarding leaderboards but the number of affected students was smaller in the summer camps, perhaps because of the less formal evaluation system (on the whiteboard). Other students preferred Virtual Bead Loom because it features more color options. It allows users to use any RGB value for beads while BeadLoom Game only allows for preset colors because different shades of colors may be hard to differentiate from one another for puzzle re-creation. Admittedly, BeadLoom Game 2.0 was missing a few key colors, most notably brown.

In session 4 the B-day group had their first survey that asked them if they preferred the BeadLoom Game with the integrated leaderboards or the Virtual Bead Loom and why. 76.1% of the students preferred BeadLoom Game and 23.9% preferred the Virtual Bead Loom. Responses for preferring the Virtual Bead Loom were very similar to those we saw in the first study. Many preferred the creative freedom of the Virtual Bead Loom. Others cited disliking the competitive aspects and feared being picked on for their performance. This illustrates the problem with traditional gamification approaches where the game-based mechanics that are such a strong draw for some users can be a strong deterrent for others. This is further supported with the differences in preferences at this stage between the two groups. 86% of the A-day group preferred BeadLoom Game with its single additional feature of custom puzzles while only 76.1% of B-day group preferred BeadLoom Game with its single feature of integrated leaderboards. This suggests the need for different game styles to appeal to more students.

The final survey in session 5 for both groups asked them to compare the full BeadLoom Game 2.0 with integrated leaderboards and user-generated content to the non gamified Virtual Bead Loom. 85.1% of the students in the A-day group and 88.1% of those in the B-day group preferred the BeadLoom Game 2.0. Reasons for preferring the Virtual Bead Loom centered mostly on the greater freedom available especially in color selection and interestingly 20% of those that preferred Virtual Bead Loom cited their preference of the word "Virtual." Table 6 shows the summary of preferences over the course of the study for the two groups.

Table 6. *Percent of groups that prefer BeadLoom Game compared to the Virtual Bead Loom*

Group	**BeadLoom Game 1.0**	**One Additional Feature**	**BeadLoom Game 2.0**
A-Day content creation first	71.2%	86.0%	85.1%
B-Day Leaderboards first	- - -	76.1%	88.1%

The final survey also asked the students to select their favorite and least favorite feature of BeadLoom Game 2.0 from the main game of solving existing puzzles, creating custom puzzles, solving other peoples' custom puzzles, or the leader boards. It also asked them to explain their reasoning for both. Table 7 shows the students response to favorite feature. While there is a blend of favorite features, the B-day group more highly favored creating custom puzzles. This could be because this was the newest feature given to these students so they were most interested in this new feature. Similar to previous open responses competition was the most popular reason for liking the leaderboards and creative freedom was the most popular reason for liking creating custom puzzles. Interestingly, students who preferred solving the existing puzzles and those that preferred solving custom puzzles both explained their preference by saying that feature had the better puzzles. Some students preferred the content we had developed and others preferred the content their peers made.

Table 7. *Favorite BeadLoom Game 2.0 feature*

Favorite feature	**A-Day**	**B-Day**
Main Game (creating target pictures)	23.4%	18.6%
Creating Custom Puzzles	19.1%	32.5%
Solving Custom Puzzles	14.9%	7.0%
Leaderboards	45.5%	41.9%

Table 8 shows how players responded to which feature was their least favorite. In the free response to this question there was a wide variety of dissenting opinions. Solving custom puzzles and the main game were criticized by some players for being too hard compared to other modes and criticized for being too easy by others. Similarly, leaderboards were described as too competitive by some, while this competition was what many students cited as the reason for it being their favorite feature. Creating custom puzzles was also dismissed by some students as boring because of a lack of goal or challenge. There contrasting opinions on what makes for an enjoyable game experience highlights the need to include both game-based and play-based game elements in a gamified educational system.

Table 8. *Least favorite BeadLoom Game 2.0 feature*

Least favorite feature	**A-Day**	**B-Day**
Main Game (creating target pictures)	25.5%	33.3%
Creating Custom Puzzles	38.3%	6.0%
Solving Custom Puzzles	17.0%	44.0%
Leaderboards	41.8%	16.0%

The final analysis of the surveys focused on the differences in responses by the females in the two groups. During the session 4 survey only 7.7% of the females in the A-day group preferred the Virtual Bead Loom compared to BeadLoom Game with user-generated content. This is much lower than the 37.5% of females in the B-day group that preferred the Virtual Bead Loom compared to the BeadLoom Game with integrated leaderboards. The males in the two groups did not show this difference between groups with 17.4% of the A-day males and 11.4% of the B-day males preferring the Virtual Bead Loom. This difference went away in the final survey once both groups had been exposed to the full BadLoom Game 2.0. The much higher percentage of females in B-day group who preferred Virtual Bead Loom suggests that either the competitive element of leaderboards, or the lack of creative freedom, or both are particularly discouraging for females. This further supports the need for play-based game elements to help gamified systems appeal to a wider arrange of students.

Study Three: Custom Puzzles - Discussion

This study is the largest study so far using the BeadLoom Game software. The survey the A-day group took in session 3 reconfirms the results from the previous studies. The traditionally gamified BeadLoom Game 1.0 outperforms the non gamified Virtual Bead Loom in the area of user preference with 71.2% of students preferring BeadLoom Game. This is close to the 70.7% who preferred BeadLoom Game in study two. Although this survey could not be given to the B-day group, based on the consistent nature of these results we can assume a similar percentage would have selected BeadLoom Game there.

With the addition of a single game element, as hypothesized, the percentage of students who preferred BeadLoom Game went up. In the B-day group who were introduced to the integrated leaderboards there was a small increase to 76.1%. This is an understandably small increase because of the classroom environment of the study. Students who were already interested in creating competition could challenge their friends in person and make their own pseudo-leaderboards similar to the behavior seen in study one. This would, however, allow for these students to be more engaged in an online environment. Another explanation for the small increase is that it does not fundamentally change the game experience. With or without leaderboards this version is game-based and will appeal to those that enjoy that style of game.

When the A-day group was introduced to the user-generated content environment of the custom puzzle mode, the percentage that preferred BeadLoom Game to the Virtual Bead Loom rose from 71.2% to 86.0%. Students that preferred a more creatively free play-based game environment were brought to the BeadLoom Game preference with this addition. Once the B-day group was given this feature it similarly raised to 88.1% BeadLoom Game preference. The custom puzzle mode provides a play-based game environment that some students prefer. By looking at the preferences in the groups by gender, it seems that this play-based element or at least an alternative to the game-based element is especially important for female students. These results confirm our hypothesis that a gamified system featuring competitive integrated leaderboards or play-based user-generated content will outperform the non gamified version of that software and the gamified system with neither of these features in the area of user preference. Additionally a gamified system with both features will have the highest user preference of the available systems. This also demonstrates that **BeadLoom Game 2.0 (featuring both game-based and play-based features) outperforms the traditionally gamified BeadLoom Game 1.0 and the non-gamified Virtual Bead Loom in the area of user preference.** Additionally BeadLoom Game 2.0 featuring only competitive leaderboards or user-generated content outperforms BeadLoom Game 1.0 with neither of these features. To maximize user preference it appears that gamification approaches should incorporate both game-based and play-based game elements. This can result in a system that is so engaging that students will use the system outside the classroom even over their winter breaks.

Observation of students and analysis of the logs for the out-of-classroom use revealed an interesting play pattern that we had not accounted for. Many students were engaging with only one mode or the other. In other words, some students would only solve puzzles in the game-based main game mode and others would only create puzzles in the play-based custom puzzle mode. This study showed that exposure to BeadLoom Game results in learning gains in the areas of Cartesian coordinates, iteration, and layering, however in this study students were required to experience both the main game and the custom puzzle mode. If students could engage with only a single mode then it is important both modes are capable of guiding user behavior. This is a unique challenge for gamification of educational software but one that is very important to address to successfully integrate play-based game elements.

Study Four: Showcase (Boyce et al., 2011c)

R4: Can play-based game elements be added to a traditionally gamified system to guide user behavior?

The goal of this study was to evaluate how well the play-based user-generated content creation environment or custom puzzle mode guided user behavior towards desirable actions. Before investigating whether this mode could result in learning gains on its own, it was important to make sure the underlying design resulted in desirable play behavior. We believe that there are two key components to the desirable behavior in the play-based custom puzzle mode.

The first component of desirable behavior is the creation of creative and complex designs. Since these puzzles are being added to the game for other users to play, it is important that the created puzzles are complex and creative. Complex and challenging puzzles are most likely to require replay to master, and the more creative and interesting designs are more likely to get students to select them to solve. Easy puzzles will not push students to learn the material and students will not choose to play a boring puzzle even if the puzzle is well suited for teaching iteration. We hypothesize that **content made in BLG 2.0 (with a peer driven content rating system) will be more creative and complex than content made in VBL.** In other words the custom puzzle showcase and the ratings given by peers will result in the creation of higher quality puzzles.

The second important component of desirable play is that the students are using layering and iteration when creating their designs. Since some students play only the custom puzzle mode it is important that they be exposed to the learning objectives of iteration and layering. If they are not using and learning the concepts of iteration and layering then the custom puzzle mode and the play-based game mode must be changed to encourage this behavior. However since these puzzles are being peer-reviewed, students should rate interesting and challenging designs higher. Therefore we hypothesize that **content made in**

BLG 2.0 (with a peer driven content rating system) will be made with more iteration than content made in VBL.

Study Four: Showcase - Methods

During the middle school camp of study two, screenshots of student work in the Virtual Bead Loom were saved for the final presentation to their parents. To do this, the student would raise their hand during the Virtual Bead Loom activity, and the counselor would save their work as an image. At the time, the students did not know these images would be used in the final presentation and were told they could take the images home with them at the end of the week on CDs. Therefore the designs being saved were considered done by the students but they did not make or save them with the intention of them being seen or evaluated by others. These were designs created without the influence of a peer-driven content rating system and would be used as the control group in this study.

For the experimental group, a selection of custom puzzles were selected from the ones created during study three. The selection process began by first removing any puzzles with less than five functions to create. These were most often experimental designs where the student was learning how to use the system. Added on to designs were also excluded. This occurred when a student submitted their work and then came back later, added additional details, then resubmitted the design usually with a number added on to the end of the design name. This behavior was a result of a lack of a save feature and would be addressed with the addition of saving. Submissions whose name suggested works in progress were also removed as they were also an artifact of a lack of a save feature. A majority of students submitted at least one custom puzzle during the time in study three that required them to engage with the Custom Puzzle mode. Thus not all the submitted custom puzzles were from dedicated content creators. After this required time ended some went on to make multiple designs during their free time illustrating their dedication to content creation. All the designs for the Virtual Bead Loom control group were from students who chose to continue using the Virtual Bead Loom during their free play time and were thus dedicated to creating designs. To accommodate this difference in subjects, only designs from students who submitted at least three custom puzzles were used in the experimental group. The experimental group was selected at random from the remaining designs. A total of 54 Virtual Bead Loom and BeadLoom Game designs were collected for the study.

Five BeadLoom Game experts were asked to rate the designs on a 7-point Likert scale in the areas of creativity and complexity. Experts were chosen based on previous work developing or using BeadLoom Game in outreach. Expert judges had a thorough understanding of the intricacies of creating designs in BeadLoom Game. These intricacies are not readily apparent to novice users. For example the simple geometric shape of a circle appears simple to make but no function in BeadLoom Game or the Virtual Bead Loom is able to create a circle. It requires multiple rectangles or the carving of negative space to create. Expert judges are also skilled at identifying where iterative patterns and layering could be used for optimal efficiency. Experts were told that creativity would be defined as the originality of content and implementation. For example concentric squares are an easy yet attractive design many students implement and would have a low creativity score. A house is also a common design but if it was designed with high levels of iteration it could have a high creativity score. Complexity was defined as the difficulty of finding the ideal solution. When evaluating the designs, the experts were blind to which group the design was from.

The analysis of the use of learning objectives was much more straightforward. The custom puzzle mode saves the steps used to create the designs. Instead of saving the color of all 441 (21x21) beads of a design it stores the designs as the steps used to create the design. This means for each custom puzzle it is possible to calculate how much iteration is being used in the creation process. The number of iterations used per design was used as the primary measure of engagement with the advanced learning objectives

Study Four: Showcase - Results

The designs created in BeadLoom Game were found to be significantly more creative and complex as evaluated by the panel of five BeadLoom Game experts. The BeadLoom Game average creativity score ($M=4.51$, $SD=0.91$, $N = 28$) was significantly greater than that for VBL ($M = 3.75$, $SD = 1.25$, $N = 26$) using two sample t-test for unequal variances, $t(46) = 2.57$, $p \leq .014$. The BeadLoom Game mean complexity score ($M = 4.1$, $SD = 0.91$, $N = 28$) was significantly higher than that for the Virtual Bead Loom ($M = 3.36$, $SD = 1.24$, $N = 26$) using two sample t-test for unequal variances, $t(46) = 2.46$, $p \leq .018$. Pearson product moment correlations yielded a coefficient of .90, indicating that as complexity of the designs increased, so did their creativity. Analysis of the logs showed that the experimental group's

BeadLoom Game designs made during study three each used, on average, less than one iteration per design.

Study Four: Showcase - Discussion

This study confirmed our first hypothesis that **content made in BLG 2.0 (with a peer driven content rating system) is more creative and complex than content made in VBL**. This suggests that such a rating system is capable of improving the overall quality of created content in user-generated content environments. Knowing that peers will see and evaluate the content and that it will be displayed in order based on these ratings drives content creators to make better content which in turn can help those who will engage with that content.

The average use of iterations per puzzle disproves our second hypothesis. The peer-driven content rating system did not encourage the users to engage with the learning objectives of the system. Although their designs were complex and could be solved using advanced iteration and layering, students were not using these techniques in their creation. Thus this implementation of a play-based game mode based on creating additional content for the system was not fully successful. It was adding replay value for those that engaged with the game-based elements and impacted system preference in those that preferred play-based systems, but was not encouraging those creative users to fully explore the learning content. Since the logs of study three showed some students chose to only engage with the play-based custom puzzle mode it was important to ensure these students received the same educational content as those that engaged with the puzzle solving game-based elements. The new goal was the design of a play-based mode that was equally engaging for users but could guide user behavior towards learning objectives.

BeadLoom Game 3.0

The primary goal of BeadLoom Game 3.0 was the improvement of the play-based custom puzzle mode. These improvements needed to guide user behavior while maintaining its play-based aspects that improved user preference and the quality of created content. Minor improvements both for and based on other studies were also incorporated. This included the addition of requested colors like brown to the bead colors and automating the puzzle completion process. In BeadLoom Game 2.0 students needed to click a submit button when they believed the puzzle was completed, but in 3.0 after each move was made, the solution was checked. This was done after analysis of logs revealed that students would sometimes leave the goal state of the puzzle without realizing they had found the solution (Eagle et al., 2013). An integrated tutorial system for the main game-based mode was also developed to replace the need for the instructor led tutorial that began all of our studies. While not the focus of BeadLoom Game 3.0, this tutorial system was found to lower tutorial completion time while improving in-game performance and was used as the basis for a framework on effective game tutorial systems (Shannon et al., 2013).

To better guide user behavior in the play-based custom puzzle mode new restrictions were added to the puzzle creation process. Like the previous version the player is tasked with creating a design on a 41x41 Cartesian grid using the six functions of BeadLoom Game. Each of these functions is assigned a cost: five points for point function, four points for line function, three points for rectangle function, two points for triangle function, and one point for the iterative functions. The total number of points used is tracked and when the player submits their design it is placed in one of four showcases. One showcase is for puzzles 25 points and under, one for 26 to 50 points, one for 51 to 75 points, and one for greater than 75 points. Although game-based elements like points and scoring have been implemented it still remains play-based. It is similar to population in *SimCity*. It is a measure the system provides and regards as good but specific goals are still left to the player. This new custom puzzle mode incorporated a series of effective strategies for guiding behavior in play-based user-generated content creation environments that we developed (Boyce, 2012). What follows is a list of the design decisions and principles made for the Custom Puzzle mode of BeadLoom Game 3.0 and discussion on how these strategies help maximize learning and guide behavior for students engaging with this mode.

Effective Strategies for Guiding User Behavior in User-Generated Content Creation Environments

1) Integrate learning objectives into content creation

As the previous study showed some users are only going to want to engage with the play-based content creation system and avoid the game-based mode. This means learning must occur in the content creation environment or the content creation environment must be unlocked by engaging with the game-based mode. However by locking the play-based elements away this creates a system very similar to the edutainment software that layered traditional games between educational content. This just makes the learning a roadblock to fun for players that wish to be creative. The learning must be integrated with the play, so for students that only want to engage with the content creation, the content creation process must incorporate the learning objective. The BeadLoom Game custom puzzle mode accomplishes this in a similar manner to the game-based puzzle solving mode by having the user plot points and utilize the functions in creating their designs.

2) Provide multiple levels of success

Users will not start out having mastered either the learning objectives or the software user interface. Therefore, the software needs to have multiple ways to succeed within the system. There should be a simple option that only requires a basic understanding of the integrated learning objective and a better, more challenging approach that requires a deeper understanding to perform. This allows the user to start with the simple strategy and transition to the more complex one. This can create a sense of freedom and voluntary choice that is very important to perceiving the activity as fun (Brown 2009; Caillois 2005). The alternative would be requiring users to only use the more complex and preferred method for content creation, which would create a steep learning curve, remove choice, and discourage play.

BeadLoom Game 3.0 accomplishes this in two ways. The first way is by allowing users to create their designs through only the use of the point function. Although this function does not teach all the learning objectives it provides a simple way to become familiar with the environment and puzzle creation process. Users can then transition to the more efficient tools when they are prepared. This is better than only providing the iterative functions that would require a large amount of upfront learning from the user to solve puzzles or create content.

The second example of multiple levels of success in BeadLoom Game 3 is the tiered showcase. Although users will strive for the 25-point showcase, if they reach 24 points and are not completely finished they can submit what they have and then roll into the 50-point showcase. They can even come back later and find a more optimal solution to their design to have it moved to a more prestigious showcase. This system does not punish new users and provides multiple avenues for improving performance. An earlier design did not have these rolling limits or the ability to replay a submitted design. Instead, points counted down and when the user ran out of points they could not add additional details to the design. This proved to be discouraging because it meant having to restart the design to complete it. The rolling limits avoid this problem while maintaining distinct levels of success.

3) Implement limitations

By providing the option to create content using both simple and complex approaches there are both desirable and less-desirable actions the user can take. Since the less-desirable action is inherently easier, a majority of users will engage in this less-desirable activity. This is demonstrated by the fact most user-generated puzzles averaged less than one iteration per puzzle. Therefore it is necessary to implement limitations on the creation process. These limitations should be designed in a way that rewards the more desirable behavior and punishes or provides less reward for the less desirable behavior. However, these limitations must not negate the user's ability to perform the less desirable action. More simple interactions are needed and are important for novice users and not having them would violate the second strategy. Instead, the system should make it clear that other options are available, and that exploring those options will result in a better overall outcome.

This was the main strategy that BeadLoom Game 2.0's content creation environment lacked. Although all the functions were available, there were no limitations placed on the user so users defaulted to using the point and rectangle functions to create their designs. There was little reason to use the harder functions over the basic functions. By assigning each function a different point value and limiting how many points could be used for each showcase, the system illustrates which behavior is the desirable option without removing the important simple functions. The users are not forced to use iteration but quickly realize it will be advantageous to use it on designs for the 25 point showcase. As Figure 20 (left) shows using only the point tool will allow the user to create a crude smiley face for 25 points. Using iteration it is possible to create highly complex designs like the user-made cell phone design on the right.

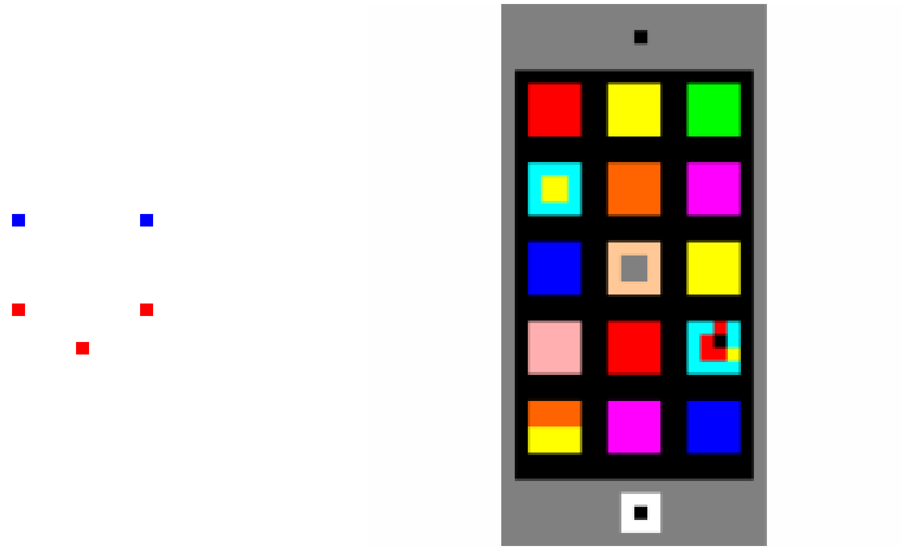


Figure 20. Two 25-point custom puzzle designs

4) Ensure the limitations do not limit creativity

When investigating why users enjoyed the custom puzzle mode or original Virtual Bead Loom, the most common response was always the creative freedom. Users needed the freedom to create what they wanted. It is therefore extremely important that the limitations do not limit user creativity. Ideally, all content that could be created without the limitation should be creatable with the limitation. It is acceptable and even preferable to require the user to work harder to create that content with the limitations. It should be clear that nearly anything is possible if the user is willing to work hard and learn the material. The earlier designed system that did not feature rolling point limits revealed that users get most frustrated when they have an idea but after working on it for some time realize it is unobtainable because of the arbitrary rules set forth by the system. This makes them dislike the system and its limitations rather than working with the limitations.

The general showcase for designs with over 75 points ensures that all content BeadLoom Game 2.0 content can also be made in BeadLoom Game 3.0. However, a vast majority of the 75-point and over puzzles can be created in under 25 points. This is because most functions can be replicated using the iterative functions. Points, straight lines, rectangles, and most triangles can all be defined iteratively. Thus a user can save many points through the use and mastery of iteration. For example, creating three rectangles will cost nine points but only three points when created with iteration. Over 90% of all the custom puzzles that were created in BeadLoom Game 2.0 could be created with 25 points or less.

5) Peer driven content rating system

Study four showed that a content creation system featuring a peer-driven content rating system results in the creation of more complex and creative designs. Although this system does not by itself guide user behavior the benefits it provides should not be overlooked. It also acts as a way for users to get ideas for content to create. For example, in Study three, after one student made a Christmas tree, a pseudo Christmas tree war began in the showcase. Christmas trees began as crude triangles but students continued to make more and more complex trees featuring lights, presents and stars. The behavior of taking an existing idea and improving on it is further enhanced with the point limitation. By “saving” points users will have more left over at the end to add on these extra details. Inspiration from the showcase can result in finding "cheaper" ways to create a design to add on that extra little something that will make it rise up the ranks of the showcase. Similar peer driven content rating systems have been used to great success in educational systems such as Scratch (Resnick et al, 2009) and should be a staple of content creation environments in play-based gamified systems.

With these strategies and the new BeadLoom Game 3.0 designed, a study was developed to test the effectiveness of the new custom puzzle mode and the principles used to create it.

Study Five: Point Limits (Boyce et al., 2012)

R4: Can play-based game elements be added to a traditionally gamified system to guide user behavior?

R5: Can engaging only with the play-based elements of a gamified system produce learning gains?

The goal of the final study involving BeadLoom Game was to evaluate the effectiveness of the newly designed play-based custom puzzle mode. Like study 4, study 5 was designed to measure how well the custom puzzle mode and the design choices behind it can guide user behavior towards desired learning objectives. In this regard we hypothesize that **Content created in BLG 3.0 (with its improved user-generated content creation mode) will be made with more iterations than conent made in the VBL or in BLG 2.0.**

Since some users will want to only engage with the custom puzzle mode of BeadLoom Game it is important that it also generates learning outcomes like the game-based elements. Therefore this study will also evaluate the learning gains that occur from playing the custom puzzle mode without exposure to the game-based puzzle solving component. For this part of the study we hypothesize that **the play-based mode of BLG 3.0 alone will generate learning gains in the areas of Cartesian coordinates, iteration, and layering.** Students that play only the custom puzzle mode will still experience significant learning gains in the areas of Cartesian coordinates, iteration, and layering.

Study Five: Point Limits - Prototype Methods

This study involved two experiments to evaluate the effectiveness of the updated custom puzzle mode with limitations. The first experiment was designed to see if the use of our strategies in the custom puzzle mode could guide player behavior. It utilized a mixed-fidelity prototype using BeadLoom Game 2.0 and a paper worksheet. With this system, players create the custom puzzle on the computer using the old system and track function usage and points on the worksheet (Appendix B). The worksheet also listed the point value for each function, had a place to put the name of the puzzle, and an area for putting the point limit for the activity. This prototype featured the hard point limits that would later be removed based on user feedback.

This mixed fidelity prototype was used based on the success of the mixed fidelity prototype from early in BeadLoom Game's history. Just like the first prototype for the game, it was important to get initial user feedback on design choices and ensure the underlying mechanics were effective before developing the system. As a result we were able to learn about the downsides to the hard point limit without having to waste development time.

This system was used at four five-day summer camps in 2011. All four camps were for middle school students with one camp focused on the use of educational tools and three focused on video game design using Game Maker. Two of these camps were held at the University of North Carolina at Charlotte campus and the other two were held at the Discovery Place children's museum. In total there were 57 participants. On the first day of camp the students were introduced to BeadLoom Game 2.0 and walked through a 30 minute tutorial on how to use each function. They were then given an hour and a half to play the custom puzzle mode with no limitations.

On days two, three, and four daily custom puzzle contests were held using the mixed fidelity prototype. Since there was only the one showcase in the mixed fidelity prototype these daily challenges were used to approximate the effect of the different showcases. During these contests, students had one hour to create a design while tracking the moves used to make it on their worksheet. At the end of the contest, camp counselors gave a small prize from the dollar store for the best designs. Students who did not win a prize by the end of the week then received a participation prize. The first contest used a 100-point hard limit. The hard limit meant that the puzzles were not allowed to use more than 100 points. The second contest had a 50-point hard limit. To encourage players who thought this was too few points an announcement was made informing everyone that all shapes except irregular lines and irregular triangles can be created using the iterative functions. The final contest had a 25-point hard limit.

Students were also given an hour each day where they were allowed to play the custom puzzle mode without limitations. Students were also allowed and encouraged to play with BeadLoom Game at home once the camp was over. Logs were saved of the functions used to create the puzzles with limitations, without limitations in the camp, and without limitations after the camp. On the final day of camp students were given a survey asking their opinions on various aspects of the custom puzzle mode with limitations. These questions included “which point limit was your favorite?” and “do you feel the point limit motivated you to use the advanced functions like iteration more than the free play mode?”

Study Five: Point Limits - Prototype Results

In total the students created 279 puzzles during the camps and an additional 34 during the two weeks following the camp. Figures 20 (right) and 21 show some of the designs the students created. The designs were grouped by when they were created and what limitations were in place for analysis.

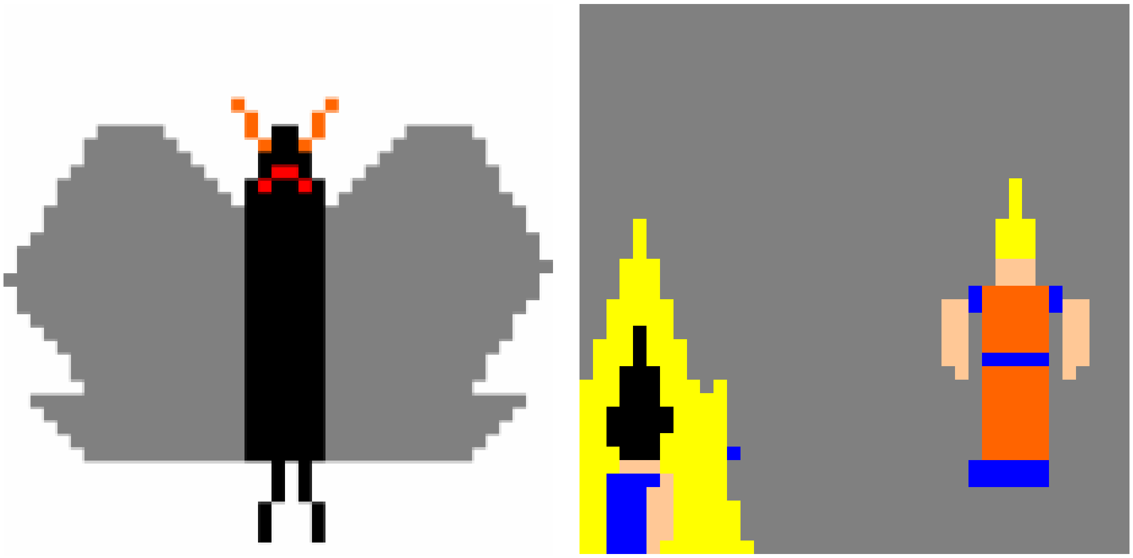


Figure 21. 100 point (left) and 50 point (right) custom puzzles created by students

Designs made on the first day with no limitations averaged .725 iterations per puzzle. The following days saw no significant change to the average iterations per puzzle made with no limitation (1.3 on day two, .957 on day three, and .56 on day four). Puzzles made with no limitations after the camp similarly averaged .79 iterations per puzzle. Puzzles made during the second day with the 100 point limit averaged 5.85 iterations per puzzle. The 50 point and 25 point limit puzzles averaged 7.80 and 9.80 iterations per puzzle respectively. These averages exclude two puzzles from the 25 point limit, one from the 50 point limit, and one from the 100 point limit because they were begun during the one hour time frame but were completed during the free play time. These results are summarized i1n Figure 22.

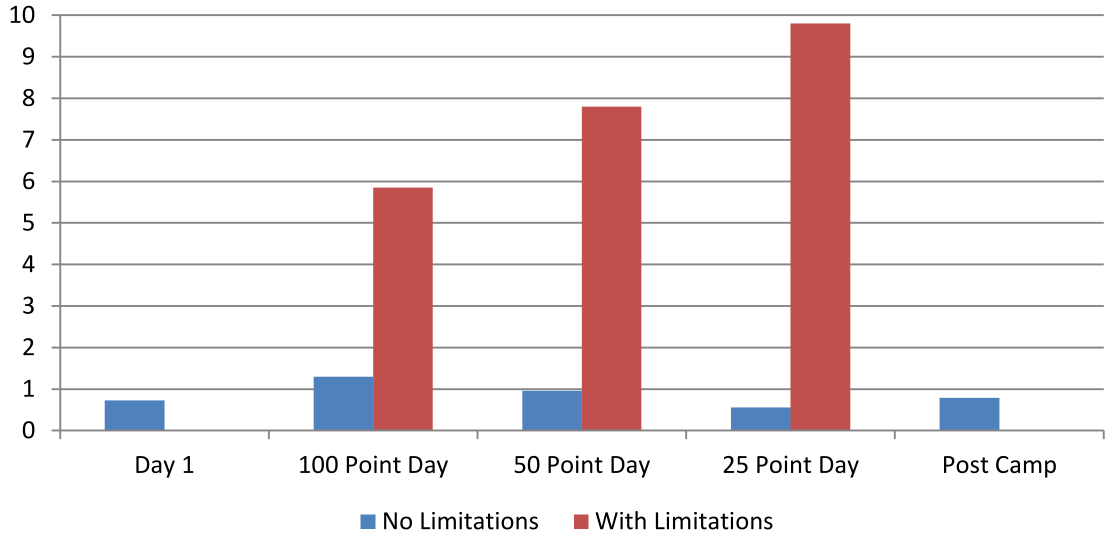


Figure 22. Average iterations per puzzle

When asked about their favorite point limits 24 students preferred the 100 point limit, 17 preferred the 50 point limit, and 16 students preferred the 25 point limit. A majority of students (46/57) reported that they felt the point limit motivated them to explore and use the iterative functions.

Study Five: Point Limits - Prototype Discussion

Even though this prototype required the students to track their moves with pencil and paper, the point limitations were effective at guiding student behavior. As we saw in our previous study, when students had no required limits, their designs averaged approximately one iteration per puzzle. With the addition of the point limitations this rose significantly. Additionally the average iterations continued to grow each day despite decreases in the total allowed points. This increasingly restrictive limitation could have resulted in less complex designs as students simply did less with less available points. This did not happen and instead the desired behavior of exploring iteration to create the designs they wanted within the tighter limitation occurred.

This mixed fidelity prototype also revealed that the hard point limit could be discouraging and did not provide enough freedom for some users. Some students would reach the point limit but still have additional features they wanted to add. They were unhappy when they were told they could do so only by restarting the puzzle and creating it more optimally. Students suggested the soft limit and the ability to go back and redo a created design in fewer moves to improve its classification. We also observed that 100 points seemed slightly too high for the upper limit so the limit was adjusted to 75. The prototype revealed that the improvements designed for the custom puzzle mode were effective and worth implementing. With this new knowledge, BeadLoom Game 3.0 was developed to build in these effective showcases and soft point limits for the second part of the study.

Study Five: Point Limits - BeadLoom Game 3.0 Study Method

The second half of this study focused on determining if using only the custom puzzle mode that follows our strategies resulted in learning gains. This is a very important objective to allow students who only want to engage with the play-based custom puzzle mode to do so. To accomplish this, the game-based puzzle solving mode of BeadLoom Game 3.0 was disabled for the experiment. A secondary objective was to ensure that learning and the guiding of player behavior could occur in a shorter classroom environment as opposed to a week long summer camp setting.

The experiment was conducted with three middle school computer and business education classes in fall 2011. One class was 6th grade students, one was 7th grade students, and one was 8th grade students. Students were given the 12 question pretest used in previous studies to evaluate their initial understanding of Cartesian coordinates, iteration, and layering. After the pretest students were led through a thirty minute demonstration of the various functions. Unlike previous tutorial sessions this one was done in the custom puzzle mode since the game-based mode was disabled. In the next session two days later the students were given 40 minutes to use the BeadLoom Game 3.0 custom puzzle mode with point limitations. They were told they could aim for any showcase they wanted but were encouraged to go for the 25 or 50 point showcase. After playing with the custom puzzle mode the students were given the isomorphic posttest. In total 44 students were present both days of the activity.

Study Five: Point Limits - BeadLoom Game 3.0 Study Results

42 puzzles were submitted during the 40 minute activity with many more left unfinished. Although the primary purpose of this part of the study was to examine the learning gains of the play-based user-generated content mode iteration usage was examined to further verify the results from the first experiment. For the published puzzles created by participants, the average number of iterations per design was 5.38. This is very similar to the value seen on the first day using the mixed fidelity prototype. Even in a shorter classroom

environment, with students free to select their own point limit targets, the custom puzzle mode designed with our strategies was effective at guiding player behavior. With prolonged exposure it is likely that these students' designs would begin using even more iterations similar to what happened in the summer camp groups.

A 2-tailed matched paired t-test was performed on the pre and posttest data. Students performed significantly better, $t(43) = 3.86$, $p<.001$, $d=.82$ on the posttest ($M=5.69$, $SD=1.72$) than on the pretest ($M=4.18$, $SD=1.96$). Cohen's d indicates that students improved from pretest to posttest by .82 of a standard deviation. Thus the use of only the play-based custom puzzle mode for a short class period length of time resulted in significant learning gains for the user.

Study Five: Point Limits - Discussion

The first iteration of the play-based custom puzzle mode in study four had mixed success. While it was effective at improving students' preference of the system and improved the quality of the created content it was unable to guide user behavior towards learning material. This meant that students who only engaged with the custom puzzle mode would not get the same learning gains as those that engaged with the game-based puzzle solving mode. However by applying our strategies for effective play-based content generation environments we were able to overcome this problem and make BeadLoom Game 3.0 effective with both its play-based and game-based elements.

The results confirm our hypothesis that **content made in BLG 3.0 (with its improved user-generated content creation environment) is made with more iterations than content made in the VBL or in BLG 2.0**. The combination of limits with peer review results in highly creative, complex content created in a way that maximizes learning. This means that play-based elements can be added to traditionally gamified systems to guide user behavior. Like the game-based elements play-based elements are capable of guiding user behavior if limitations and peer review systems are used to encourage the desirable behavior and discourage less desirable behavior. This is done without limiting the creative freedom those players who like play-based systems desire and without forcing game-based interactions that they dislike.

The second part of the study validated the teaching potential of the play-based content creation mode and confirmed our hypothesis that **the play-based mode of BLG 3.0 alone can generate learning gains.** This was also the first study looking at the use of the user-generated content mode in a single shorter classroom setting. Even with the short exposure students saw significant learning gains and significantly increased use of iteration. Clearly having a system that is enjoyable over a long period of time helps with replay and mastery but it is important for an educational system to be able to generate immediate results in a classroom environment and time frame.

This study verified that a play-based mode in a gamified system could be effective at guiding behavior and generating learning gains. Combining traditional gamification techniques with the best practices learned from developing BeadLoom Game results in a system that appeals to a wider audience. We call this approach Deep Gamification.

Deep Gamification

R6: What game elements and design principles can be used to create a gamified system that combines game-based with play-based mechanics?

Deep gamification, our new approach to gamification and the design recommendations associated with it, is the final result of the studies on BeadLoom Game. It combines the benefits of traditional game-based gamification with the optimized play-based user-generated content creation environment. The use of this approach results in gamified systems that outperform the original tool and traditionally gamified systems in the areas of learning gains and user preference. User preference is especially important to encourage the use of the system outside the classroom environment to allow for replay and continued learning and mastery of the learning material. This section outlines the principles of deep gamification and discusses the benefits each principle has been shown to produce based on the studies with BeadLoom Game.

The first part of deep gamification mirrors traditional gamification with the inclusion of **points and achievements**. This forms the basis of the **game-based mode**. In studies one and two it was shown that just the addition of points and achievements results in a gamified system that outperforms a non-gamified system in the areas of learning and player preference. However, points and achievements alone are not enough to create successful gamified systems. We **integrated the point and achievement systems with learning objectives**. The points and medals earned in BeadLoom Game are directly tied to demonstrating an understanding of the learning material. Adding points and achievements to a system will not improve the system's educational potential if the points and achievements are not tied to and encourage engagement with the learning material. Contrast this to a potential serious game or gamified system that has students avoid ghosts in a haunted house while solving trivia questions. If the points and achievements are tied to the ghost dodging element then the points and achievements will not improve scores in the learning component. Integrating the points and achievements with the learning objectives or desired behavior ensures that the only way to improve in game performance is by learning the material or performing the desired behavior. BeadLoom Game's game-based mode gameplay centers on the use of the learning material. The only way to improve on a score is with better application of the learning material.

Another important component is ensuring that **the game element and learning element are not separate** but are instead completely integrated with one another. This is in contrast to earlier educational and serious games with little to no connection between the learning material and game elements for example playing a traditional game with trivia questions between the rounds. This does not leverage the motivational properties of game elements towards learning. Even if the scores and points are awarded based on knowledge of the material if there is a separation between the learning and the gameplay the learning will be seen as an obstacle to the game play. Learning should be the result of engaging with the system not something that must be done to get to the part of the system that is fun. The core interaction in BeadLoom Game is the plotting of points and the use of iteration and layering. The game play is centered on the learning objective.

The purpose of these achievements is to provide the students who enjoy game-based systems goals supplied by the system to work towards. However it is important to remember that these students will not begin as masters of the learning objectives or system's user interface. Therefore it is important to provide **tiered achievements** so that there are goals to strive for at all skill and knowledge levels. This guides students toward higher achievements and hence toward mastering the material. It also helps prevent students from becoming discouraged early in using the system as there are still achievements they can obtain with a low knowledge of the material and system. This also improves the replay value of the system by encouraging students to work towards increasingly difficult achievements. BeadLoom Game's medal system acts as the tiered achievements granting bronze medals for base level performance up to platinum medals for perfect iterative and layering mastery.

It is also important to **not prohibit less desired player behavior**. The system should instead provide rewards for engaging in the desired behavior. This allows novices to start with less optimal behavior and transition as they become comfortable with the system and material. With tiered achievements the students will be encouraged to make this transition themselves without requiring the system to force them to. This improves the sense of freedom the students feel with the system. Both this and the tiered achievements help contribute to the play and fun for the student. Choice and freedom are critical components of play (Brown, 2009; Caillois 2005) and by not prohibiting behavior and having tiers of success you leave some aspects of goal selection to the user. While players who enjoy game-based systems do like being given goals it is important to let them choose a goal to increase the chance they see the activity as fun. Gamified systems are more interesting and effective when there is not just a right and wrong answer, but instead there is a continuum of behavior, at one end rewarded for the attempt and at the other rewarded for mastery. This lets the users set their own goals and sub goals, lowering the overall learning curve of the system, and increases the chance of engaging in play with the system.

BeadLoom Game 2.0 showed two additional important elements for maximizing the effectiveness of the game-based mode in a gamified system. The first was the **integrated leaderboards**. Many students enjoy the competition this feature brings to the gamified system. If this element is not present students will often simulate it by sharing scores, but by incorporating it the system can harness the competitive spirit of students even when they are not in the same room. For certain students, having their score beaten by friends can be very encouraging for replaying and mastering content. It is an effective element used in traditional gamification and has been shown to be worth including in deep gamification for the more competitive students.

The second element for game-based systems BeadLoom Game 2.0 demonstrated was the need for **unlockable content**. This was originally added to prevent new players from attempting content they were not yet prepared for. To ensure replay for more expert students the gamified system will have hard and easy challenges as encouraged by the tiered achievements. If a student attempts challenging content before having a basic understanding of the learning objectives and user interface then they may become frustrated and stop using the system. An unintended benefit of this addition was that many students were driven to unlock the locked content. This was seen as another challenge that many students picked as the one they wanted to achieve. So this element both prevents new user frustration and is a goal some students will self-select and strive for. In BeadLoom Game 2.0 all puzzles are given a difficulty and are locked until all the previous difficulty puzzles are completed.

In nearly every open feedback about BeadLoom Game in all its forms one comment was always made. Many students did not like having to input the Cartesian coordinates by hand. They wanted to be able to click on a bead and have the coordinates automatically input into the graphing function. It is true that this user interface would be faster and more user-friendly. However it would also bypass one of the fundamental learning objectives. If students could click on a bead to put in its coordinates then students could use the functions without understanding the Cartesian coordinate system. Students will make mistakes without the point and click functionality especially when learning the difference between positive and negatives on the axis but this is an important concept in the learning objectives of the system. Good game design says this should be point and click but this is one of the areas a gamified system dramatically differs from a game. The user interface and user experience should of course be improved as much as possible but the system **must sacrifice ease of use when that ease of use will promote less desirable behavior or bypass learning objectives**.

Following these design principles resulted in a more effective game-based gamified system. The gamified BeadLoom Game produced higher learning gains and a higher percentage of user preference compared to the non-gamified Virtual Bead Loom. However in study two 29.3% of users still preferred the non-gamified Virtual Bead Loom for its creative freedom and lack of system defined goals. Not appealing to these users is ok in many gamification settings but when gamifying an educational tool it is important to appeal to and engage with as wide an audience as possible. It is because of this that deep gamification incorporates the play-based game element of a user-generated content creation mode.

This play-based user-generated content creation environment provides many benefits to the gamified system that are lacking in traditional gamification. The biggest of these is the impact on user preference. Study three showed that gamified systems with play-based content creation systems have higher user preference than gamified systems without them. In every survey 85% or more of students preferred BeadLoom Game when it involved a play-based user-generated content creation mode. Having the play-based game element be a content creation environment has the additional benefit of providing additional content to the gamified system. With the importance of replay for learning in an educational system it is important there be enough content for users. With new content being created by users it is possible to develop gamified systems in less time by letting users create additional content for the system rather than the developers. As the content the students in the studies show, sometimes user created content even surpasses that of the content made by developers. Over the course of the BeadLoom Game studies over 500 custom puzzles have been created far out numbering the original 25 designed for BeadLoom Game 1.0.

Although study three showed the impact on user preference the play-based content creation system has, study four and its analysis of the created content revealed some problems. Students were not engaging with the learning material when creating the custom content. They were instead engaged in the undesirable activity of avoiding iteration. This indicates that not all play-based user-generated content creation environments are effective at guiding player behavior. Therefore deep gamification incorporates design principles for effective play-based user-generated content creation environments as well.

The first design principle is that the **learning objective must be integrated with the content creation**. The play-based content creation mode must be able to impact learning on its own. As the studies show some students will not want to engage with the game-based elements and will prefer to only use the play-based content creation mode. If the system makes creative users engage with the game-based mode before being allowed to create content, we have simply replicated an edutainment system with a traditional game interrupted by trivia questions. This causes students to view the educational component as a roadblock to fun. This means that both the high and low end learning objectives must be utilized in content creation. Students directly utilize Cartesian coordinates and are capable of using iteration and layering in creating their custom puzzles in BeadLoom Game.

The downfall of the original custom puzzle mode was that the system provided no incentives to engage with the more complex learning material. Students could make their designs bead by bead. This is an important step in the learning process, but students also need an incentive to excel beyond this basic strategy. Traditional gamification obtains this incentive through the use the game-based elements of points and rigid system defined goals. The play-based mode is free of such goals and therefore must instead **utilize limitations** to encourage desired user behavior. These limitations do not set the goal; they simply show the user that there are options in how to reach their own goals and that some of those options are better or more desirable than others. In BeadLoom Game the cost associated with the different functions acts as this limitation. Students are still free to use any function but they are shown that some are more costly and less desirable than others. Study five showed this strategy to be effective at increasing the use of iteration compared to the custom puzzle mode without limitations.

Similar to the game-based mode, the play-based mode needs more than one way to create content. This would result in a steep learning curve and limit the user's freedom thus lowering the chance for them to engage in play. Therefore there must be **multiple levels of success** in the creation of the content. By allowing the students to start with single point graphing and transition to the more complex functions, the system lets the student to become familiar with the material and realize the benefit of the harder functions on their own. BeadLoom Game 3.0 creates a slightly more rigid level of success in the form of the four different showcases. Although it is grouping the created content, the system never assigns a goal of which showcase to strive for. The goal of what to create and under what limitations is left to the user but the limitations are offered by the system. This maintains the essential play-based element of the content creation system.

Although it alone was not effective at guiding players to use iterations the **peer-driven content rating system** was effective at increasing the overall complexity and creativity of the created content. While exploring others' work students were often inspired to create more complex versions of existing designs to reach higher in the showcase. The peer-driven content rating system allows for evaluation of the student work without the system imposing goals on the user. Unlike system assigned goals the user is free to try to climb up the rankings in the showcase or simply ignore it and work towards their own objectives.

It is important to remember that this play-based content creation environment is meant to appeal to the students who dislike the game-based modes lack of creative freedom. Although it is important to implement limitations that illustrate the desired behavior these **limitations must not limit creativity or prohibit less desirable behavior**. Students may integrate the systems' limitations as their own but not if those limitations completely prohibit something they want to do. As study five showed, students often strived to accomplish designs within the point limits of the showcase but would become frustrated when they reached a hard point limit. This would pull them out of play and make them dislike the limitations that a moment before they had been working with and enjoying. Anything that can be made without the limitations must be creatable with the limitations; otherwise the system may break the play of some students. The soft bounds on the different showcases and the ability to use any function ensures that anything that could be made in BeadLoom Game 2.0's custom puzzle mode could also be made in BeadLoom Game 3.0 although the student will need to work harder to get that same design into the 25 point showcase.

Finally just like in the game-based elements the play-based elements **must sacrifice ease of use when that ease of use will promote less desirable behavior or bypass learning objectives**. It would once again be an easier user interface and make for a better game environment if when creating content you could click on points to automatically fill in the bead placement functions. This would completely bypass the learning. Since this is an educational system nothing should allow users to completely bypass the learning without it impacting their in-system evaluation, even if that means slightly harder to use interfaces.

Overall Results

Over its five years of research BeadLoom Game has been used to answer a variety of questions. Here we answer six that pertain to the process of deep gamification.

R1: Can a traditionally gamified educational system produce learning gains?

Yes, a traditionally gamified educational system can produce learning gains. While traditional serious game and gamified system research mainly focused on the impacts of user preference, study one investigated the impact BeadLoom Game has on learning gains. With only the implementation of basic points and achievements, the core elements of traditional gamification BeadLoom Game was able to generate learning gains in student understanding of iteration and layering while not sacrificing the learning of basic Cartesian coordinate principles.

R2: Can the use of a traditionally gamified system outperform a non-gamified educational tool in the areas of learning gains and user preference?

Yes, a traditionally gamified system can outperform a non-gamified educational tool in the areas of learning gains and user preference. Serious game and gamification research also tends to only evaluate the system in a vacuum when it is important to compare its effectiveness to the original system. Study two showed that not only does BeadLoom Game generate learning gains, it produces higher learning gains than the non gamified version, and a majority of students prefer the gamified BeadLoom Game. Although a majority of the students preferred BeadLoom Game, nearly 30% still preferred the non-gamified Virtual Bead Loom. These students did not like the lack of creative freedom and harsh competition of the game-based puzzle solving mode. For these students, the play-based custom puzzle mode was made and effective strategies for designing play-based user-generated content creation environments were investigated and incorporated into deep gamification.

R3: Does a gamified system featuring both game-based and play-based game elements outperform a traditionally gamified system in the area of user preference?

Yes, a gamified system featuring both game-based and play-based game elements outperforms a traditionally gamified system in the area of user preference. This was the main focus of study three. Although the traditionally gamified system outperformed the non gamified system, and the gamified system with either integrated leaderboards or the play-based elements outperformed both of them, the system featuring both game-based and play-based elements was the most popular. Additionally, these features did not detract from the learning gains from prolonged use of the system.

R4: Can play-based game elements be added to a traditionally gamified system to guide user behavior?

Yes, play-based game elements can be used to guide user behavior. The play-based custom puzzle mode in BeadLoom Game 2.0 was shown to be effective at guiding students to create more complex and creative designs through its use of the peer driven content rating system. However this was not enough to encourage the use of iteration and discourage the creation of these designs using the simple functions. Through the use of limitations and the other design principles for effective play-based user-generated content creation environments the system was able to guide students to also engage in desirable behavior while creating their content. Using these design principles it is possible to make a play-based environment that guides user behavior similar to the game-based environment.

R5: Can engaging only with the play-based elements of a gamified system produce learning gains?

Yes, engaging only with the play-based elements of a gamified system can result in learning gains. This was a very important question to answer because some students who did not like the game-based elements wanted to only engage with the play-based content creation mode. If the mode on its own was unsuccessful generating learning gains, then it would be unsuccessful as an educational system. However the design principles for play-based user-generated content creation environments created a mode that generated learning gains and successfully guided user behavior. Additionally this was shown to work in a short classroom setting as opposed to the prolonged summer camp setting other studies had used. This validated the system as an educational system that can generate results in a short classroom session but is enough fun to make students elect to play it at home in their free time. Doing this for the widest group of students possible is ultimately the goal of a gamified educational system. Our design principles created a system that accomplished just that.

R6: What game elements and design principles can be used to create a gamified system that combines game-based and play-based gameplay?

As summarized above and in Table 9 deep gamification with its combination of game-based and play-based game elements and the design principles surrounding those elements can be used to create a gamified system that effectively blends the two types of games, increases learning gains, increases total content of the system, and maximizes user preference of the system. This process has been sculpted and driven by the five year incremental improvement of BeadLoom Game. It is the result of the feedback and desires of hundreds of student participants and is the next step in gamification. Students who prefer play-based games can now be offered the same benefits that traditional gamification allows while contributing content for those who prefer the traditional game-based approach. While it is slightly more complex to design systems that feature both game styles the player generated content lowers the amount of packaged content that must be included in the game-based system.

Deep gamification calls for a shift in the traditional G/P/S model of serious game and gamified system classification. To maximize user preference it is important to include both game-based and play-based elements in the gamified systems. Doing so maximizes user preferences, creates gamified systems that produce higher learning gains than the original non gamified system, and greatly adds to the total content of the system. Following the design principles outlined in deep gamification ensures both the game-based and play-based modes are effective at teaching and promoting play for users.

Table 9. *Deep gamification design principles*

<table>
<tr><th colspan="2">Element</th><th>Benefit</th></tr>
<tr><td colspan="2">Points and achievements
(game-based)</td><td>Motivates users who want a game-based experience</td></tr>
<tr><td></td><td>Integrate points/achievements with learning objective/desired behavior</td><td>The only way to improve performance is through learning/ desired behavior</td></tr>
<tr><td></td><td>No separation between
learning element and game element</td><td>Prevents learning from becoming an obstacle to play</td></tr>
<tr><td rowspan="2"></td><td>Tiered achievements</td><td rowspan="2">Allows setting sub goals
Encourages play
Lowers learning curve</td></tr>
<tr><td>Do not prohibit less desirable behavior</td></tr>
<tr><td></td><td>Locked content</td><td>Motivates some users
Prevents discouragement</td></tr>
<tr><td></td><td>Competitive leaderboards</td><td>Motivates some users</td></tr>
<tr><td></td><td>Sacrifice ease of use to avoid bypassing learning objectives</td><td>Prevents avoiding learning</td></tr>
<tr><td colspan="2">User-generated content creation environment
(play-based)</td><td>Motivates users who want play-based experience
Provides content for game-based mode</td></tr>
<tr><td></td><td>Integrate learning objectives with content creation</td><td>Users engage with learning objective through content creation</td></tr>
<tr><td></td><td>Implement limitations</td><td>Encourages full exploration of learning objective and desired behavior</td></tr>
<tr><td rowspan="2"></td><td>Multiple levels of success</td><td rowspan="2">Does not discourage user
Does not limit content creation</td></tr>
<tr><td>Limitations do not limit creativity or prohibit behavior</td></tr>
<tr><td></td><td>Peer driven content rating system</td><td>Encourages more complex and creative content generation</td></tr>
<tr><td></td><td>Sacrifice ease of use to avoid bypassing learning objectives</td><td>Prevents avoiding learning</td></tr>
</table>

Application

Although this research has focused on the single educational system of BeadLoom Game, the deep gamification approach is applicable to any educational system that could be gamified. To verify this claim other research projects within our research lab have already begun utilizing the deep gamification approach in their development. The most heavily researched and published of these projects is the educational game BOTS. BOTS is a gamified programing tutor that teaches programing logic such as loops. It features both a game-based and play-based modes similar to BeadLoom Game with the play-based mode being a level creator. Like BeadLoom Game it has been found to be an effective educational system that students both learn from and have fun engaging with (Hicks, 2012; Hicks, et al., 2014). As BeadLoom Game and BOTS show deep gamification can and should be widely used in the improvement of educational software and education as a whole. With systems like these it is possible to teach students in a classroom setting the desired material all while the student enjoys the activity. Perhaps most exciting of all is the prospect that students may continue to play and learn in their spare time. If BeadLoom Game can get middle school students to play and learn at home over their winter break then why can’t all educational software do that?

This approach has been specifically designed to maximize the learning gains in an educational system but gamification can be applied to a wide array of systems other than educational ones. Deep gamification is an extension of the successful gamification approaches that combines the two types of games to appeal to more users and maximize learning. It is not a stretch to imagine that all existing gamified systems and all systems that could be gamified could be further enhanced by following the guidelines and design principles of deep gamification.

Limitations and Future Research

The greatest limitation to this research has been its focus on a single system. However this has also been one of its greatest strengths. Its continual focus on one system has allowed for incremental improvements and comparative evaluation to each of its previous incarnations. This allowed us to see what impacts each change had and rigorously evaluate the different versions of the system. This limitation is also slowly being reduced by the research of others in the field. The success of systems such as BOTs further validates the results we have seen in BeadLoom Game that deep gamification is a powerful tool for augmenting educational systems.

A limitation of the process itself lies in trivia based learning objectives. BeadLoom Game, BOTS, and all the culturally Situated Design Tools are skill based learning where the user is learning how to perform a skill such as graphing or creating loops. For systems like these the deep gamification approach is straightforward. The points and achievements center around performing those skills and the user-generated content tasks them with creating content with those skills. For systems that focus on pure knowledge (like state capitals for example) it would be more difficult to design the play-based user-generated content environments. The user-generated content would typically need to be checked by an expert or the system would need to be created in such a way that it could evaluate the correctness of the created content. This is not needed in skill-based systems, because, if they are designed well, the very act of creating the content using the desired skills demonstrates that the content can at least be replicated.

Another area of future research in this area is investigating ways to improve the quality and learning potential in user-generated content in educational systems. It is one thing to ensure that the created content can be recreated or solved in the game-based mode but it is another to have the created content be as effective as content made by experts. It is important to design systems that can guide users to create content that is good at teaching targeted concepts (Hicks, et al., 2014). In BOTS this means having users create content that encourage the use of loops. In BeadLoom Game it could be having users create puzzles that are especially effective at promoting the use of iteration. It is important to ensure that the content users are creating is effective at teaching those that engage with it in the game-based modes. The principles discovered in this research will further improve the play-based user-generated content creation environments and the quality of the content available in the game-based modes in deep gamification.

Another important area of research will be investigating the effectiveness of deep gamification in areas other than education. Deep Gamification was specifically designed to maximize learning and improve user preference. All gamified systems want their gamified system to be preferred to the non-gamified system and many involve altering or guiding user behavior. Therefore, as discussed, it is likely that deep gamification is capable of further enhancing gamified systems in areas other than education but developing and testing such systems to verify these beliefs is an important next step.

Conclusion

BeadLoom Game began as a traditional gamification of the educational tool the Virtual Bead Loom. The BeadLoom Game project had a simple goal: encourage students to fully explore the learning content of the Virtual Bead Loom. It accomplished this goal through the use of traditional gamification approaches. However this revealed that although effective for a majority of students the very elements that encouraged some students discouraged others. By including both game-based and play-based game elements BeadLoom Game was able to overcome the shortcomings of traditional gamification and create a system that appealed to a wider range of students. Because these play-based features were content creation it also added replay value on to the game-based element. **Through a series of five experiments we have shown that BeadLoom Game 3.0 and its deep gamification outperforms the non gamified system Virtual Bead Loom as well as the traditionally gamified BeadLoom Game 1.0 in the areas of learning gains and player preference.** This marriage of game-based and play-based elements has the potential to invigorate gamification in educational software and result in the creation of new gamified systems that are more effective at teaching and more fun for a wider range of students.

This deep gamification approach can be used to augment current gamified systems and non-gamified systems. Although this research focuses on educational systems, it is possible this approach can be effective in all the areas where traditional gamification works; but play-based elements could potentially increase the appeal to certain users and enhanced replay value. Games are more than points and leaderboards. Games do not have to be game-based. Deep gamification acknowledges and incorporates the strongest elements of the play-based games to appeal to a wider audience. To appeal to the widest audience gamified systems should not be play-based or game-based as the G/P/S model suggests. Instead gamified systems should incorporate the strengths of both game types and be both. That way no matter which style of game the user prefers the gamified system has something for them.

REFERENCES

Alvarez, J., & Michaud, L. (2008). “Serious Games: Advergaming, edugaming, training and more.” France, IDATE.

Badgeville (2014). “Game Mechanics” http://badgeville.com/products/gamification Accessed May 4, 2014.

Barnes, T., Powell, E., Chaffin, A., Lipford, H. (2008). “Game2Learn: Improving the motivation of CS1 Students.” In *Proceedings of the 3rd International Conference on Game Development in Computer Science Education* (GDCSE ’08). ACM, 1-5.

Bergeron, B. (2006). Developing Serious Games. USA, Charles River Media

Bert, C., King, K., Brown, A, & Watkins, D. (2009). “Standardizing Culturally Situated Design Tools.” Accessed May 23, 2013. www.cra.org/Activities/craw_archive/dmp/awards/2009/Bert/Website/FinalReport.pdf

Bogost, I. (2011, August 9). Gamification is bullshit. The Atlantic: Technology. Retrieved from http://www.theatlantic.com/technology/archive/2011/08/gamification-is-bullshit/243338/

Bouca, M. (2012). “Mobile Communication, Gamification and Ludification.” In *Proceedings of the 16th International Academic MindTrek Conference (MindTrek ’12)*. ACM, New York, NY, USA, 295-301.

Boyce, A., Finkelstein, S., Clark, S., Kinsey, H., Doran, K. (2009). “Culturally Situated Design Tool games.” Poster presented at *STARS Celebration 2009,* August 9-12, 2009, Tallahassee, FL.

Boyce, A., Barnes, T. (2010). “BeadLoom Game: Using Game Elements to Increase Motivation and Learning.” In *Proceedings of the 5th International Conference on Foundations of Digital Games* (FDG ’10). ACM, 25-31.

Boyce, A., Campbell, A, Pickford, S., Culler, S, Barnes, T. (2011a). “Experimental Evaluation of BeadLoom Game: How Adding Game Elements to an Educational Tool Improves Motivation and Learning.” In *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education* (ITiCSE ’11). ACM, 243-247.

Boyce, A., Doran, K., Campbell, A., Pickford, S., Culler, D., Barnes, T. (2011b). “BeadLoom Game: Adding Competitive, User Generated, and Social Features to Increase Motivation.” In *Proceedings of the 6th International Conference on Foundations of Digital Games* (FDG ’11). ACM, 139-146.

Boyce, A., Doran, K., Campbell, A., Pickford, S., Culler, D., Barnes, T. (2011c) “Social User Generated Content’s Effect on Creativity in Educational Games.” *In Proceedings of the 8th ACM Conference on Creativity and Cognition* (C&C ’11). ACM 317-318

Boyce, A., Campbell, A., Pickford, S, Culler, D., Barnes, T. (2012) “Maximizing Learning and Guiding Behavior in Free Play User Generated Content Environments.” In *Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education* (ITiCSE ’12). ACM, 10-15.

Brauner, P., Valdez, A., Schroeder, U., Ziefle, M. (2013). “Increase physical fitness and create health awareness through exergames and gamification: the role of individual factors, motivation and acceptance.” *SouthCHI 2013.* 349-362.

Brown, S. (2009) “Play: How it Shapes the Brain, Opens the Imagination, and Invigorates the Soul” New York: Avery.

Caillois, R. (1961). “Man, Play and Games.” USA, Free Press of Glencoe

Caillois, R. (2005) "The Definition of Play and The Classification of Games." The Game Design Reader: A Rules of Play Anthology. Ed. Katie Salen and Eric Zimmerman. New York: MIT P, 2005. 123

Chen, S., & Michael, D. (2005) *Serious Games: Games that Educate, Train, and Inform*. USA, Thomson Course Technology.

Cooper, S. Dann, W., Pausch, R. (2000) “Alice: a #-D tool for introductory programming concepts” In *Proceedings of the Fifth Annual CCSC Northeastern Conference on the Journal of Computing in Small Colleges* (CCSC ’00). John G. Meinke (Ed.) Consortium for Computing Sciences in Colleges, USA, 107-116

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic. (2010). “Predicting protein structures with a multiplayer online game.” *Nature* 466 756-760.

Crowley, D. (2009) "Screenshot from a checkout terminal at Target (Milford MA)" http://dpstyles.tumblr.com/post/260890668/screenshot-from-a-checkout-terminal-at-target Accessed May 3, 2014.

Csdt.rpi.edu. "Culturally Situated Design Tools." Accessed: August 13, 2014.

Csdt.rpi.edu/na/loom/bl/beadloom.html. "Virtual Bead Loom." Accessed: August 13, 2014.

Deci, E., Koestner, R., Ryan, R. (2001). "Extrinsic rewards and intrinsic motivations in education: Reconsidered once again." *Review of Educational Research, 71*(1). 1-27.

Deterding, S., Dixon, D., Khaled, R. N., & Nacke, L. (2011a). "From game design elements to gamefulness: defining gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9-15). ACM.

Deterding, S., Dixon, D., Khaled, R. N., & Nacke, L. (2011b). "Gamification: Toward a Definition." In *CHI 2011 Gamification Workshop Proceedings*.

Deterding S., Sicart, M., Nacke, L., O'Hara, K., Dixon, D. (2011). "Gamification: Using game-design Elements in non-gaming contexts". *In CHI'11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, 2425-2428.

Djaouti, D., Alvarez, J., Jessel, J. P. (2010). "Can 'Gaming 2.0' help design 'Serious Games': a comparative study." In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (pp. 11-18). ACM.

Djaouti D., Alvarez, J., Jessel, JP. (2011). Classifying Serious Games: The G/P/S Model. Felicia, P. (ed). Handbook on Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches, 118-136 IGI Global

Dunwell, I., de Freitas, S., Petridis, P., Hendrix, M., Arnab, S., Lameras, P., Stewart, C. (2014) "A game-based learning approach to road safety: the code of everand." In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (CHI '14). ACM, New York, NY, USA, 3389-3398.

Eagle M., Barnes, T. (2009) "Experimental evaluation of an educational game for improved learning in introductory computing." In *Proceedings of the 40th ACM technical symposium on Computer science education* (SIGCSE '09). ACM, New York, NY, USA, 321-325.

Eagle M., Johnson, M., Barnes, T., Boyce, A. (2013) “Exploring Player Behavior with Visual Analytics” In Proceedings of the 8th International Conference on Foundations of Digital Games (FDG ’13).

Egenfeldt-Nielson, S. (2005). *Beyond Edutainment: Exploring the Educational Potential of Compute Games*. IT University of Copenhagen, Department of Innovation, 84-85.

Eglash, R., Bennet, A., O’Donnell, C., Jennings, S., & Cintorino, M. (2006). “Culturally Situated Design Tools: Ethnocomputing from Field Site to Classroom.” *American Anthropologist* 108(2), 347-362.

Farzan, R., DiMicco, J., Millen, D., Dugan, C., Geyer, W., Brownholtz, E. (2008). “Results from deploying a participation incentive mechanism within the enterprise.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’08). ACM, New York, NY, USA, 563-572.

Fay, A. L., & Mayer, R. E. (1994). “Benefits of teaching design skills before teaching LOGO computer programming: Evidence for syntax-independent learning.” *Journal of Educational Computing Research*, 11, 187-210.

Frasca, G. (2003). Simulation versus Narrative: Introduction to Ludology. In Wolf J.P. & Perron B. (Eds.), The Video Game Theory Reader, USA, Routledge.

Gagne, R. M., & Brown, L.T. (1961). “Some factors in the programing of conceptual learning.” *Journal of Experimental Psychology*, 62, 313-321.

GameEffective (2014). “Narratives” http://www.gameffective.com/products/narratives/ Accessed May 4, 2014.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). “Games, motivation and learning: A research and practice model.” *Simulation & Gaming*, 33(4), 441-467.

Gartner, “Gartner says by 2015, more than 50 percent of organizations that manage innovation processes will gamify those processes.” Gartner.com/newsroom/id/1629214, April 12, 2011 Accessed May 4, 2014.

Gee, J. (2003) “What video games have to teach us about learning and literacy.” *ACM Computers in Entertainment. 1*(1), 20-20.

Gillispie, L., Lawson, C. (2010) *WoWinSchool – a hero's joruny- A middle grades language arts adventure.* http://wowinschool.pbworks.com/f/WoWinSchool-A-Heros-Journey.pdf Accessed May 4, 2014.

Goh, H., Lee, C. (2011). "Perceptions, quality and motivational needs in image tagging human computation games." *Journal of Information Science* October 2011 37: 515-531

Hamari, J., Koivisto, J., Sarsa, H. (2014). "Does Gamification Work? – A Literature Review of Empirical Studies on Gamifcation." In *Proceedings of the 47th Hawaii International Conference on System Sciences*. Hawaii, USA, January 6-9, 2014.

Harold, N., Tan, C., Rosser, D., Leong, T. (2014). "CopyMe: an emotional development game for children" In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '14). ACM, New York, NY, USA, 503-506.

Hicks, A. (2012) "Creation, evaluation, and presentation of user-generated content in community game-based tutors." In Proceedings of the 7th International Conference on Foundations of Digital Games (FDG '12). ACM, 276-278.

Hicks, A., Cateté, V., Barnes, T. (2014). "Part of the game: Changing level creation to identify and filter low quality user-generated levels." In *Proceedings of the 9th International Conference on Foundations of Digital Games* (FDG '14).

Hirsh-Pasek, K., Golinkoff, RM. (2008). "Why play=learning." In: Tremblay RE, Boivin M, Peters RDeV, eds. Encyclopedia on Early Childhood Development [online]. Montreal, Quebec: Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development; 2008:1-6. Available at: http://www.child-encyclopedia.com/documents/Hirsh-Pasek-GolinkoffANGxp.pdf. Accessed 1/15/2013.

Iosup, A., Epema, D. (2014). "An experience report on using gamification in technical higher education." In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education* (SIGCSE '14). ACM, New York, NY, USA, 27-32.

Ito, M. (2009). *Engineering play: a cultural history of children's software.* The MIT Press, Cambridge, MA.

Jackson, G., Boonthum, C., McNamara, D. (2009). "iSTART-ME: Situating extended learning within a game-based environment." In *Proceedings of the Workshop on*

Intelligent Educational Games at the 14th Annual Conference on Artificial Intelligence in Education. 59-68.

Jackson, G., Davis, N., McNamara, D. (2011). “Student’s enjoyment of a game-based tutoring system.” *Artificial Intelligence in Education*. Springer Berlin Heidelberg. 475-477.

Jackson, G., McNamara, D. (2011). “Motivational impacts of a game-based intelligent tutoring system.” *FLAIRS 2011*.

Jackson, G., Boonthum-Denecke, C., McNamara, D. (2012). “A comparison of gains between educational games and a traditional ITS.” *FLAIRS 2012*.

Jackson, G., McNamara, D. (2013). “Motivation and Performance in a game-based intelligent tutoring system.” *American Psychological Association*

Jonassen, D.H. (1997) “Instructional design models for well-structured and ill-structured problem solving learning outcomes.” *Educational Technology Research and Development* 45 (1): 65-94.

Jung, J., Schneigder, C., Valacich, J. (2010). “Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaborative environments.” *Management Science 56, 4*, 724-742.

Khatib, F., Dimaio, F., Cooper, S., Kazmierczyk, M., Kryzywda, S., Zabranska, H., Pichova, I., Thompson, J. (2011). “Crystal structure of a monomeric retroviral protease by protein folding game players.” *Nature Structural & Molecular Biology* 18 (10): 1175-1177.

Kittel, J. E. (1957). “An experimental study of the effect of external direction during learning on transfer and retention of principles.” *Journal of Educational Psychology*, 48, 391-405.

Kohn, A. (1999). *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Boston: Houghton Mifflin.

Korn, O. (2012). “Industrial playgrounds: how gamification helps to enrich work for elderly or impaired persons in production.” In *Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (EICS ’12). ACM, New York, NY, USA, 313-316.

Kuramoto, I., Ishibashi, T., Yamamoto, K., Tsujino, Y. (2013). "Stand up, heroes!: gamification for standing people on crowded public transportation." In *Proceedings of the Second International Conference on Design, User Experience, and Usability: Health, Learning, Playing, Cultural, and Cross-Cultural User Experience – Volume Part II* (DUXU'13), Aaron Marcus (Ed.) Springer-Verlag, Berlin, Heidelberg, 538-547.

LBP.me. "Little Big Planet Community." Accessed Feb. 15, 2014.

Lim, Y., Pangam, A., Periyasami, S., Sheweta, A. (2006) "Comparative Analysis of High- and Low- fidelity Prototypes for More Calid Usability Evaluations of Mobile Devices". *NordiCHI 2006.*

Liu, Y., Aleandra, T., Nakajima, T. (2011). "Gamifying intelligent environments." *In Proceedings of the 2011 International ACM Workshop on Ubiquitous Meta User Interfaces*, *Ubi-MUI '11* (pp. 7-12) ACM, New York, NY, USA.

Logo Foundation (2011). "What is LOGO." Retrieved from: el.media.mit.edu/logo-foundation/logo/index.html. Accessed June 21, 2014.

Maroney, K. (2001). "My entire waking life." *The Games Journal*, May. Retrieved from http://www.thegamesjournal.com/articles/MyEntireWakingLife.shtml

Mayer, R. (2004). "Should there be a three-strike rule against pure discovery learning? The case for guided methods of instruction." *American Psychologist*, 59, 14-19.

Mekler, E., Tuch, A., Bruhlmaan, F., Opwis, K. (2013). "Disassembling Gamification: The Effects of Points and Meaning on User Motivation and Performance." *In CHI'13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*

Myfitnesspal.com. "My Fitness Pal." Accessed May 3, 2014.

NextBee (204). "Gamifcation Platform." http://www.nextbee.com/buzz-marketing.htm Accessed May 4, 2014

Nicholson, S. (2012a). "A User-Centered Theoretical Framework for Meaningful Gamification." *Games Learning Society 2012: Madison*.

Nicholson, S. (2012b). "Strategies for meaningful gamification: Concepts behind transformative play and participatory museums." Presented at *Meaningful Play 2012*. Lansing Michigan

Nickel, A., Kinsey, H., Heidi, H., Pendergrass, M., Barnes, T. (2012). "Interval training with Astrojumper." In *Proceedings of the 2012 IEEE Virtual Reality* (VR'12). IEEE Computer Society, Washington, DC, USA, 161-162.

Nickel, A., Payton, J., Barnes, T., Wikstrom, E., Burgos, M., Wartell, Z., Pendergrass, M., Blanchard, N., Thomas, B. (2014). "Balancing physical and cognitive challenge: A study of players' psychological response to exergame play." In *Proceedings of the 9th International Conference on Foundations of Digital Games* (FDG '14).

Okan, Z. (2003). "Edutainment: is learning at risk." *British Journal of Educational Technology*, 34(2), 255-264

Papert, S., Harel, I. (1991). *Constructionism*, Ablex Publishing Corporation.

Portugal, JN. (2006). "Le Rapprochement du Jeu et de l'Apprentissage." Paper presented at the Serious Game Summit Europe 2006, Lyon, France.

Prensky, M. (2005) "Computer games and learning: digital game-based learning." *Handbook of Computer Games Studies.* Cambridge MA: The MIT Press.

Resnick, M., Maloney, J., Monroy-Hernandez, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., Kafai, Y. (2009) "Scratch: Programming for all." Comm. ACM 52, 11, 60-67.

Robertson, M. (2010). "Can't play, won't play.Hide & Seek: Inventing New Kinds of Play." Retrieved from http://www.hideandseek.net/2010/10/06/cant-play-wont-play/

Sawyer, B., & smith, P. (2008). "Serious Game Taxonomy." Paper presented at the Serious Game Summit 2008, San Francisco, USA

Shannon, A., Boyce, A., Gadwal, C., Barmes, T. (2013) "Effective Practices in Game Tutorial Systems." In *Proceedings of the 8th International Conference on Foundations of Digital Games* (FDG '13).

Shulman, L. S. & Keisler, E. R. (1966). *Learning by discovery*. Chicago: Rand McNally.

von Ahn, L., Dabbish, L. (2004). "Labeling images with a computer game." In *Proceedings of the SIGCHI Conference on Human Factors in Computing* (CHI '04). ACM, New York, NY, USA, 319-326.

von Ahn, L. (2006). "Games with a purpose." *Computer*, vol 39, no 6, 92-94 June 2006.

von Ahn, L. (2009). "Human computation." *Design Automation Conference 2009, DAC '09. 46th ACM/IEEE*, 418-419.

Zyda, M. (2005). "From Visual Simulation to Virtual Reality to Games." Computer, 38(9), 25-32.

APPENDICIES

Appendix A: Pretest for BeadLoom Game Studies

PRE TEST

Name: _______________

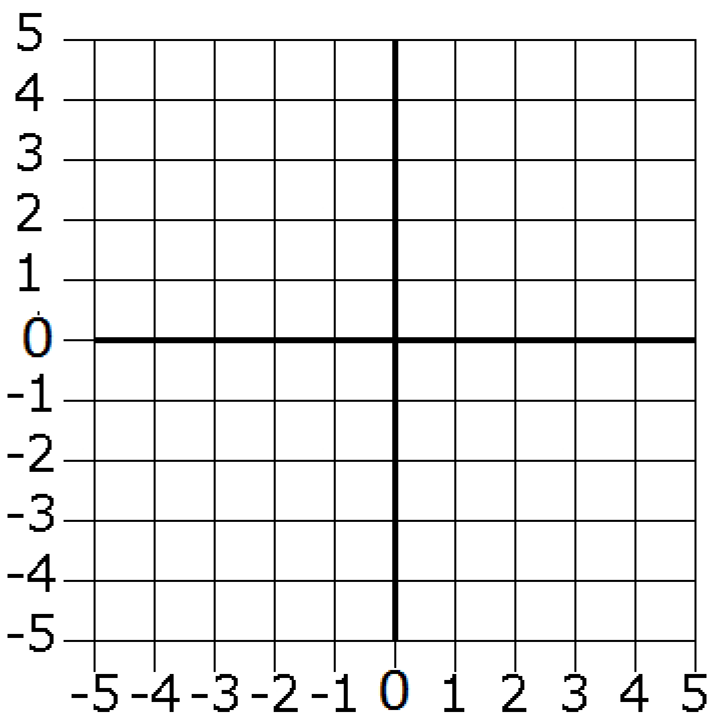
Grade: _______________

1) Have you ever worked with the Virtual Bead Loom or the BeadLoom Game before? (If you do not know what these are then please circle No)

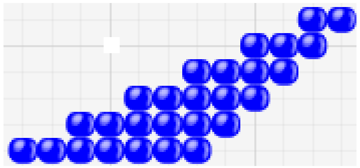
YES NO

2) Draw a bead on the grid for each ordered pair. Color the beads, or write the name of the color next to the bead

White -2,3 Black -3,-3 Gray 4,1 Spotted 1,-2



3)



This shape was created using iteration. It can be described by the iterative pattern:

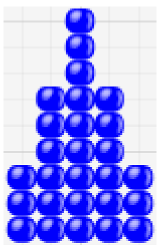
Start with _7___ beads

Every line add _____ bead(s) to the right and

add _____ bead(s) to the left

for _____ lines in total.

4)



This shape was created using iteration. It can be described by the iterative pattern:

Start with __1__ bead

Every _____ line(s)

add _____ bead(s) to each side

for _____ line(s) in total.

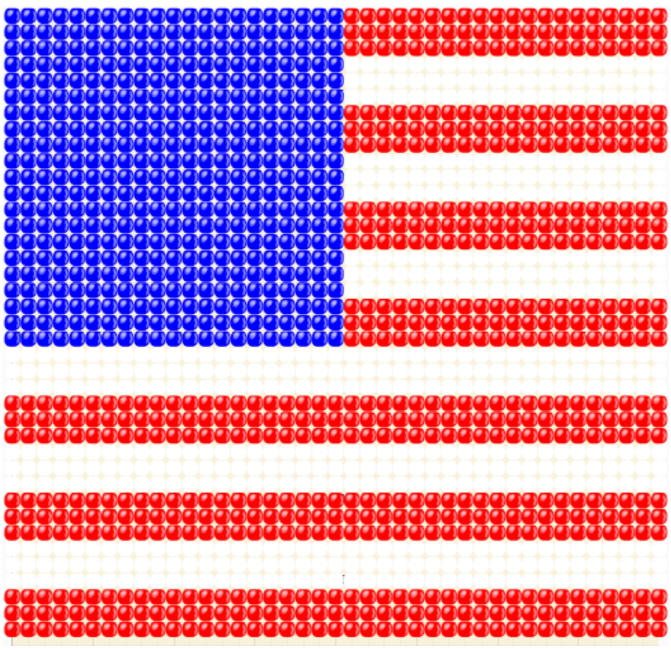
5) You have a line of 10 beads. Every iteration, you add 3 to the left side and remove 1 from the right side. How many beads are in the line after 5 iterations if the first line is iteration 0?

____

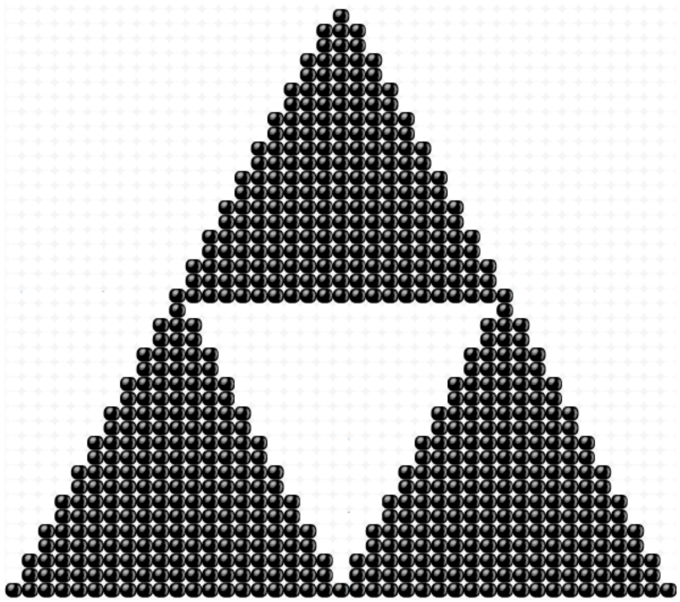
6) You have a line of 20 beads. Every iteration, you remove 2 to the left side and remove 3 from the right side. How many beads are in the line after 3 iterations if the first line is iteration 0?

____

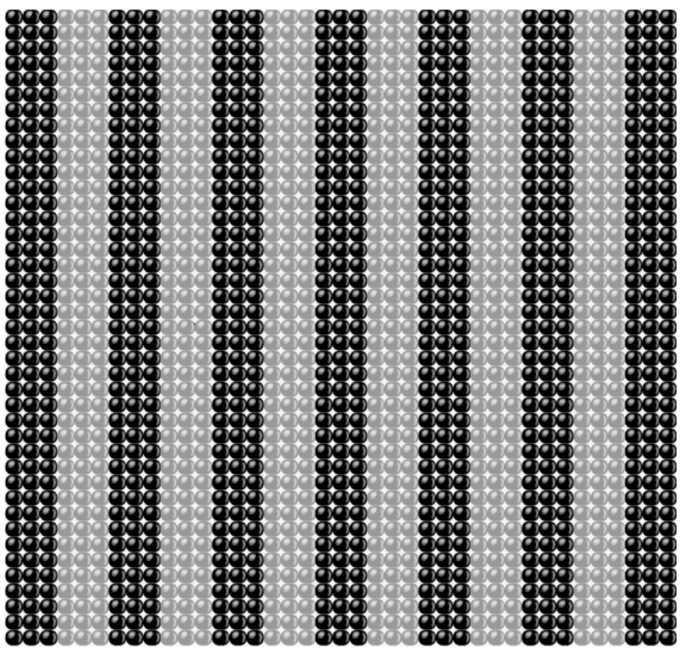
7) Below is a picture of the American Flag. How many rectangles does it take to recreate this picture (including the white stripes. The top and bottom stripes are red.) _________



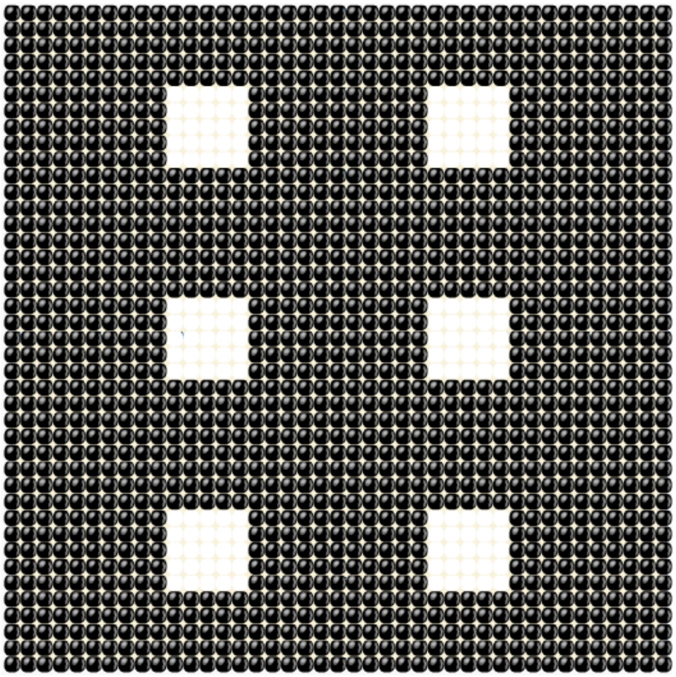
8) What is the minimum number of triangles needed to make the picture below (Do not include the background): _______



9) What is the minimum number of rectangles needed to make the picture below:_____



10) What is the minimum number of rectangles needed to make the picture below:_____



Appendix B: Paper Component of Mixed Fidelity Prototype for Study Five

Name: ____________________

Puzzle Name: _______________

Point Limit: 25 50 100

Functions		**POINTS**
Point Functions:_____________________	x5 =	________
Line Functions:_____________________	x4 =	________
Rectangle Functions:_________________	x3 =	________
Triangle Functions:__________________	x2 =	________
Linear Iteration Functions:_____________	x1 =	________
Triangle Iteration Functions:___________	x1 =	________
TOTAL:	**=**	________