EXPLORING DIGITAL GAMES AS A RESEARCH AND EDUCATIONAL PLATFORM FOR REPLICATING EXPERIMENTS

Steven C. Sutherland, Northeastern University, Boston, Massachusetts, 02115, st.sutherland@neu.edu
Casper Harteveld, Northeastern University, Boston, Massachusetts, 02115, c.harteveld@neu.edu
Gillian Smith, Northeastern University, Boston, Massachusetts, 02115, gillian@ccs.neu.edu
Joseph Schwartz, Northeastern University, Boston, Massachusetts, 02115, j.schwartz@neu.edu
Cigdem Talgar, Northeastern University, Boston, Massachusetts, 02115, c.talgar@neu.edu

ABSTRACT

Having students create their own experiments is an important way to teach science inquiry and research methods. Additionally, having individuals create experiments to replicate previous research creates a better understanding of the concepts derived from previous findings. Replication is often difficult due to limited access to the necessary tools. However, with the increased prevalence of digital games and the ease with which new games can be modified for research, games offer the promise that individuals will be able to recreate experiments previously difficult to replicate. In this paper we discuss a new digital game, *Mad Science*, created specifically for novice users with no programming skills to create their own experiments and to recruit and collect data from participants. Our findings suggest that replication of previous findings can be accomplished using this platform and will be a valuable tool for teaching concepts and methods from the decision science literature.

Games as research method, games as instructional method, replication

INTRODUCTION

The decision sciences are interested in better understanding the decisions that people make when there is uncertainty or risk involved, including: whether to select one risky gamble over another, whether to seek advice from a costly expert, and whether we accept some amount of money now or wait for a larger amount later? Much of the research in the decision sciences has involved survey-style paper and pencil tasks. While these tasks are good at providing the details about the decision space, they are often hypothetical decisions with no real, or minimal, consequences. Due to a host of ethical issues, it is often not possible to put individuals in risky scenarios and provide real consequences, relevant to the decision being made, for many situations. As researchers, we hope that individuals make decisions in our tasks that are consistent with their decisions in the real world. The question that we are faced with is how we can create more realistic studies with real consequences while still providing some control over the variables of interest. We argue that digital games and virtual environments, in general, are a promising option for studying decision making and teaching decision-making concepts.

Digital games have come a long way in how realistic they are and their accessibility to non-traditional gamers for playing games and for non-programmers for creating games or changing content. Researchers have begun using games in a number of settings and for reasons beyond their original purpose. For instance, off-the-shelf games have been shown to improve performance on cognitive tasks or to affect behavior outside of the games [15]. Additionally, some researchers have started to modify variables and consequences in games, using the games as research tools and to create more immersive scenarios and environments [46]; [45]. Because researchers can manipulate and control variables of interest while allowing players to navigate their environments and control their characters, games offer researchers a tool to better understand behavior. For decision scientists, games give the option to look at both single-shot decisions and repeated decisions, such as those used in experience-based decision-making. Games also allow researchers to easily assign participants randomly to scenarios and to randomly sample stimuli from a pre-established set of variables or stimulus space, increasing the external validity of the study and the ability to better identify the functional relationships between variables [44]. Based on these capacities, we argue that games can also be leveraged further to teach students how to conduct research and about past research studies in a controlled but interactive manner.

Researchers maintain that games can be a valuable tool for education [16]. In 2006, the [8] published a report describing the potential impact of games and concluded that games may be especially effective in teaching higher-order skills such as decision making because "in games, players are making decisions continually, in contrast to low levels of decision-making in traditional learning" (p. 43). More recently, Constance [39], the previous Senior Policy Analyst in the Executive Office of the President of the United States, who advised on national policy decisions about the impact of games, led a White House Academic Consortium on Games for Impact. This consortium concluded that there are five affordances that make games a powerful educational tool: (i) they are pervasive: a broad and diverse audience plays nowadays anytime and anywhere; (ii) they include elaborate models and simulations of the real world, which allows for safe experimentation of any kind of situation; (iii) they are highly engaging: games have the ability to captivate and sustain the attention of users; (iv) they have proven to make a positive impact on cognition and behavior: although more evidence is warranted, various studies have found conclusive evidence on the benefits of playing games [4,14,47]; and (v) they enable a data exhaust as each and every possible action can be tracked and monitored in a controlled environment, allowing for assessment and feedback to players and their teachers [35,36]. Many have realized this educational potential, as educational games have appeared in various disciplines and context, from teaching about cancer [22] to world history [37]. For the decision sciences, it has this potential as well, as games are about making decisions [18].

In this paper, we explore this potential of using digital games for research and education in the context of the decision sciences using a game that we developed called

Mad Science. This game allows players to create their own experiments and to participate in other researchers' experiments. Eventually it will be able to be utilized by general audiences and will be a tool that allows us not only to crowd-source results regarding how people respond to different manipulations in our own studies, but also to crowd-source results about the types of studies players create and the results of their studies. For the tool to be successful, a crucial requirement is that it teaches players how to perform research. In addition, we envision that this game can be successfully integrated into classrooms – to teach research methods but also to illustrate relevant experiments to students in an engaging, participatory manner. In the classroom, teachers will be able to both have students play through previously programmed scenarios and to create their own scenarios, requiring that they work through the experimental design process. In fact, implementing the game into the classroom is our first step in the development process and in this paper we are sharing our first insights of using the game in such a context, which has ramifications for the use of games for both research and education.

Our focus in this exploration is the replicability of existing experiments in this game. We have implemented three well-known and extensively-researched phenomena (framing effect, decoy effect, and emotional contagion) using *Mad Science* and one of our first steps in our design process was to find evidence that would suggest replication of the original findings. The results of this design step and the discussion around what may or may not explain the replicability are the main contributions of this paper. Before we present our results and provide for explanations, we first describe the design and vision surrounding the *Mad Science* project in more detail as well as detail the implementation of it in the classroom, to provide the necessary context of our work and the results presented.

Mad Science: Enabling Participatory Crowdsourced Research

Mad Science is a digital game where players join the fictitious corporation Mad Science Inc. as one of their new "mad" scientists—people who are intrinsically curious and show that curiosity in every aspect of their behavior. Mad Science Inc.'s mission is to "understand why people do what they do." Before players can contribute to this mission, they start on the (tutorial) island Madland where they are familiarized with the core rules of Mad Science Inc. and the tools available for them to create playable scenarios. Players also get experience with the corporation's proprietary machinery to study human behavior. Once players are familiar with the core rules and tools (achieving a score illustrative of understanding), they can go to Madworld, a world where players can design and perform research, participate in the research of other players, and share effective strategies for designing research scenarios. This scaffolding is intended to foster a basic understanding of how research scenarios are created before creating their own experiments.

Our long-term goal is to accomplish through *Mad Science* what we have coined "participatory crowdsourced research." This term refers to having large numbers of players authentically and collaboratively participate in creating and experiencing scientific research. While in this section we will explain our concept and design of *Mad Science* in more detail, we will first set the foundation by discussing educational gaps that this game is meant to address.

Educational Gaps

In formal education, research is integrated in various degrees of involvement. Typically, in most content-focused courses research is passively experienced; teachers explain to students what research has taught us about a topic or students will read significant research as part of the required textbooks. For example, in a decision-making course students will become familiar with the availability heuristic through a lecture or a reading on the work of Tversky and Kahneman [40]. Although this passive approach of consuming research is by no means wrong and is sometimes necessary to cover all of the material, students likely learn better through "experiential learning" [24]. Content is better retained when students have engaged with a phenomenon such as the availability heuristic and then have reflected on it, rather than just taking note of it. This realization has pushed much of the educational innovations in the past decades such as project-based learning, the flipped classroom, and more recently the use of games [23].

In some courses, such as an introduction course in psychology, students are required to participate in actual research. From an educational perspective the idea here is that students learn about research through their participation. Studies have supported using participation in psychology experiments as a means to teach students about how research is conducted in psychology and about the breadth of content that is researched by psychologists [13]; [32]. Although participants are debriefed about the purpose and motivation of the study, they usually do not receive any feedback on the results. Such feedback is essential for learning [12], hence they do not have a chance to reflect on their experiences. In addition, the debriefing is still a passive manner to convey what the research is about and students have no incentive to learn about this. Debriefing is also not provided as a learning tool, but rather to inform participants about what they consented to. Also, they do not have a say in the design of the studies.

Additionally, research methods courses typically require students to illustrate their knowledge by conducting research themselves. However, students do not usually get to set up and run an experiment until the final years of their undergraduate program, if at all. The following quote highlights other issues in the education of research methods:

> "too often research methods instructors teach the way they were taught. This approach to teaching preserves the normative paradigm in which the teacher is all knowing and the student is an empty vessel waiting to be filled. It goes without saying that such pedagogical approaches relegate the student to a non-participant status…Other strategies are needed that empower students to participate in their own learning process" [11:120].

Therefore, students are still not fully engaged even when opportunities for "active learning" are presented to them in their undergraduate careers.. Based on this, we believe that a more effective education about research involves one that is experiential and authentic. In other words, one that allows students to create their own experiments or to replicate past experiments [33][34]. Replicating previous findings has been suggested as a method to better understand research methods and concepts [9]; Collaborative Replications and Education Project, 2013). This entails a constructionist approach to education. Papert [28] defined constructionism (similar to constructivist theories) as a view of learning as a reconstruction rather than as a transmission of knowledge. Also, he added (similar to experiential learning theories) that learning is most effective when the learner creates a meaningful product. This educational approach is consistent with the revised Bloom's Taxonomy [2], where "create" is the highest level of cognitive processing. In order to create, a student must understand and evaluate the material. Well-known examples of constructionist approaches are programming environments such as *Scratch* [31] that enable novices to make their own interactive stories, games, and animations. Through this tool, students learn programming by creating.

We argue that *Mad Science* can be used as a tool for teaching research. Although the type of student involvement might different across different types of classes (e.g. in some courses, *Mad Science* may be used mainly for illustrative purposes while in research methods courses it may be used as a platform for students to conduct research themselves), this tool is useful across courses and not just suited for courses teaching research methods solely. This idea is especially relevant to the social sciences, as students often learn about theories and concepts that may be relevant or applicable to their everyday lives outside of the classroom. Thus far, it does not appear that such a game, emphasizing a constructionist approach has been created before. Research has primarily focused on making players, in particular K-12 students, participate in the basic natural sciences, such as biology, chemistry, and physics [38].

Although *Mad Science* addresses educational gaps in current education, our ultimate goal is to deploy it for informal education. We aim to accomplish participatory crowdsourced research: enabling citizens to perform science. Currently, citizen science refers to individuals who voluntarily participate in a scientific process—constructed and framed by professional researchers—by contributing their time, effort, and/or resources without having necessarily a formal science background [29]. Existing forms on a large-scale involve using idle computing time, tapping into human computational power, or registering observations [26,30]. Two well-known game examples concern tagging pictures and solving protein puzzles: the *ESP Game* [1] and *Foldit* [6], respectively. Both games have attracted thousands of users, yet neither one lets their users conduct science nor educates them about the science.

With *Mad Science*, we plan to bridge the gap between the scientific community and the public at large. The tedious tasks of clicking and tagging that are made into an engaging activity in current crowdsourced citizen science games is reminiscent of the scientific hierarchy defined by William Whewell [5]. In his Great Tide Experiment of 1835 he described citizen science participants as "subordinate laborers" who are capable of collecting information, while professional scientists have the ability to make meaning of the information. Through *Mad Science* we want to transform the way science is performed, enabling a more diverse and global audience to take on the role of scientists, who are able to introduce scenarios or hypotheses that professional scientists might not normally consider, while also offering large amounts of crowdsourced data to scientific researchers. In short, we hope to democratize science. What we aim to create is a culture of curiosity, where everyone can be involved with contributing to science in a manner that does not reflect the scientific hierarchy that is present in current crowdsourced citizen science games.

Game Design

Mad Science is a cross-platform game. It will be playable on mobile devices and browsers, making it portable and accessible in most locations. In essence, playing the game is focused around dialogue choices. Players are immersed into situations where they need to make a decision on how to respond. A typical situation involves responding to a conversation with a Non-Player Character (NPC), a character controlled by no human player. In responding, players are provided a closed or open response option. With a closed option, players select their preferred response among a set of predefined answers; with an open option, players type in their response. This latter option should be used with care because NPCs are scripted and will likely not respond in a meaningful way.

It can further be considered a “massively single-player game.” Although players play alone, other players populate the content in the game and players can also socialize with each other. The majority of experiments will be user-created. Players will be able to access all user-created content. New objects are added to libraries and experiments can be copied and changed once a player participated in those experiments. Players will not inhabit the same virtual space, but are able to contact players after participating in their experiments and by visiting players’ “personal spaces” (including their houses, yards, and laboratories). Madworld is a dynamic environment. Every time players visit Madworld, its appearance will change. Except for the player’s personal space and crucial landmarks, the world will be populated with new personal spaces and experiments. In this manner, players will remain curious about what this game will have to offer and will be exposed to randomly sampled experiments.

Although players can access experiments through various shortcuts, the basic manner to participate in experiments is by exploring Madworld. Players will then run into an experiment and can choose to participate in it or not after being informed (in the style of typical informed consent forms) regarding what the experiment is about. At all times players can exit an experiment, which is in accordance to standard research protocols. When players finish an experiment, they will get to see the immediate results for the data collected so far. They are further able to copy the experiment and access the data/results of this experiment at all times.

For creating experiments players will get access to a set of tools, ranging from a tool to create objects (Figure 1) and a tool to change the characters in the experiment (Figure 2) to a tool to script the dialogue and events in the experiment. In *Mad Science*, players will have access to backgrounds (e.g., an office, hospital, or classroom) in which to place their experiment. Players also have access to pre-constructed furniture and objects that can be placed in the environment that can be changed between players. Players will be able to use a simple storyboard style (e.g., Figure 3) for creating their experiments, which will require no programming skills on the part of players. By providing all of the necessary tools and art assets to create fast and easy-to-play scenarios, we hope that players from varying backgrounds and with different interests will be able to create scenarios that test the hypotheses they generate. Players will be required to create and provide informed consent and debriefing forms. This feature will allow researchers who are interested in using the platform to follow the protocols established by their institutions for conducting research. It will also give those who are part of the general population a chance to understand the ethical rigor that is applied to human behavioral research. The platform will have policies in place to flag the use of inappropriate language and for inappropriate experiments.

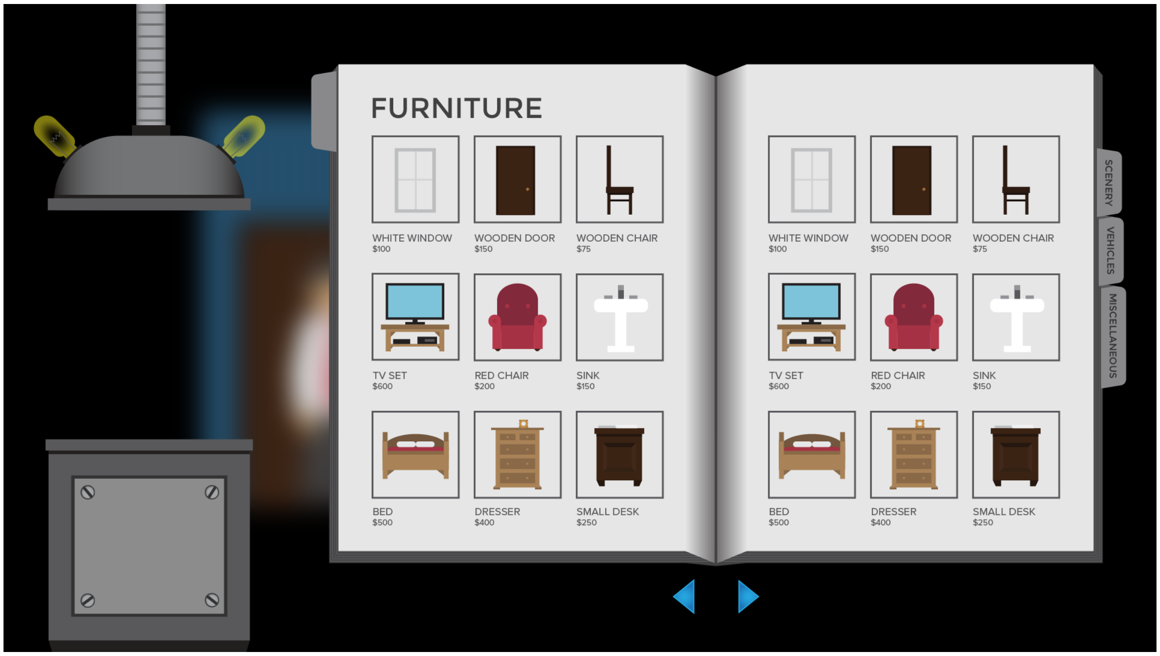

Figure 1. An illustration of the object creator tool in *Mad Science*. Players will be able to create objects that can be placed in scenarios to create different settings.

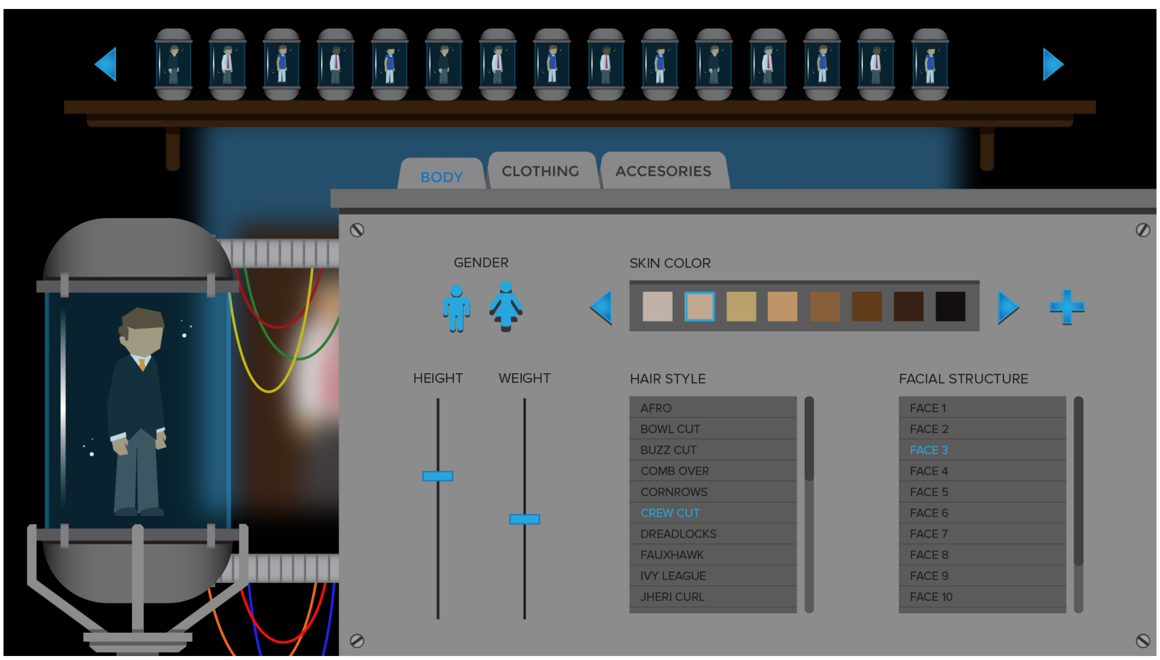

Figure 2. An illustration of the character creator tool in *Mad Science*. Players will be able to create the characters that can be act out the scenarios in their experiments.

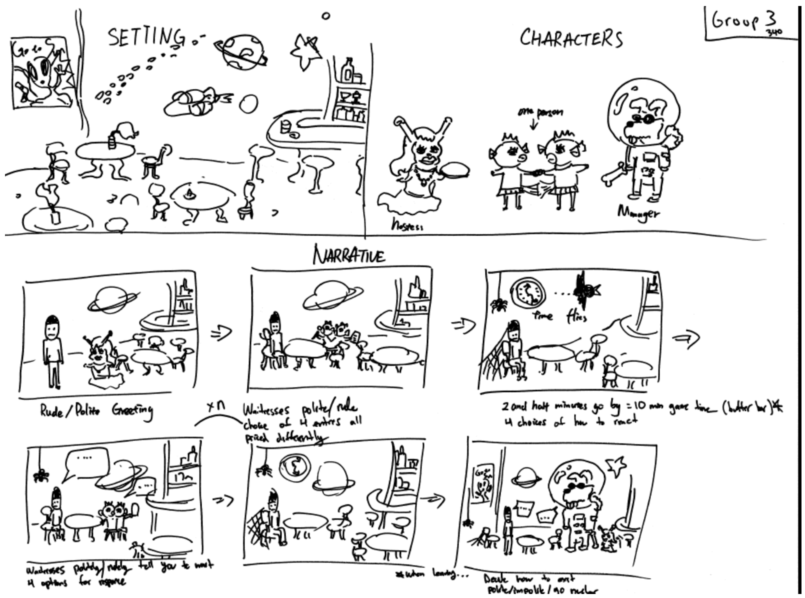

Figure 3. Example of a visual narrative created by students. In this scenario, the students were interested in testing the effect of server politeness/rudeness on customer decisions.

Players will not have immediate access to all of the tools and will get more access to the features of each tool as they gain more experience. Herein lies the motivation for players to explore Madworld and participate in experiments. This progression model is typical for most games and allows us to provide the necessary scaffolding for players to learn how to play the game and, more importantly, how to successfully create playable scenarios that can test their hypotheses. Of course, by participating in experiments players will have other gains as well, such as being able to copy the experiment and gain access to the data. After players have unlocked all features, they can participate in "themed challenges" where these expert players are encouraged to build experiments that meet a particular theme (e.g., society and gender roles, or financial conflicts of interest).

In terms of data collection and visualization, players can record player choices but also reaction times and other important behavioral data, and will have various options for visualizing the results. Players can perform advanced statistical manipulations by exporting the data and using the statistical software of their choice. Players can further invite people to participate through social media such as Twitter and Facebook. Participants with an invite do not need to be a registered player to participate in an experiment. However, they will not be able to access anything else unless they register. Teachers interested in having students play specific scenarios can create and invite the student to their experiment via a number of options (social media, email, course management websites, etc.). This means that for class demonstrations, students will not need to register for the game in order to participate.

Mad Science is a game that will be aimed at general audiences and will be a tool that allows us not only to crowd-source results regarding how people respond to different manipulations in our own studies, but also to crowd-source results about the types of studies players create and the results of their studies. Creating a social environment for experimentation will allow the general population to test how different manipulations affect behavior, but it will also be a free and easy-to-use tool that traditional researchers can use to create their own experiments and to collect data from a large number of participants that they would not have had access to in the past.

Implementation of Mad Science in the Classroom

The previous section described the concept and design of *Mad Science*. This is a work-in-progress and it will take much iteration before the final game is engaging and ready to be released for a general audience. In the meantime, we are playtesting prototypes and collecting evidence by implementing the game in various classrooms. The implementation that we describe here is in the course COMM 1131: Sex, Relationships and Communication, an elective course offered by the Department of Communication Studies. The course focuses on communication as it occurs in sexual and romantic relationships, specifically on the role of verbal and nonverbal communication in these relationships. The course can be marked as a content-focused course where research is used to inform students about what is known. Students in this class do not perform any empirical research. We integrated a prototype of *Mad Science* as part of an activity spanning two, 60-minute classes. Our objectives with this implementation were two-folded. First, our aim was to replicate well-known studies from the decision sciences using students who had not been receiving lectures around the tested content. In this prototype we included experiments on the framing effect, decoy effect, and emotional contagion. This replication is important from an educational as well as a scientific perspective. From an educational perspective the tool will be more effective if the results from existing studies provide the same kind of results in *Mad Science*. Otherwise students will not understand the relevance of the work. From a scientific perspective, replicability is crucial for validating the platform for pursuing research.

Second, our aim was to get a sense of what experiments students would want to make themselves and how they would go about doing this. One of our major design challenges is to make sure that novices can create experiments and since the students in this class are not taught research methods, this class was a good test case to see what students would do after having been exposed to the platform. This class activity informed how we would continue to design the digital tools creating experiments. To that end, this activity did not make use of our digital tools. Testing a design in this way is recommended for game design and referred to as "paper prototyping" [10]. In this exercise, we provided scaffolding to contextualize it around the content of the course.

We implemented *Mad Science* as follows. In the first class, we started by explaining the context and purpose of the two classes and then letting the students play the game in the same room, which involved completing the three experiments. Students were not told what concepts the scenarios tested before playing the scenarios. The exact procedure and content of these experiments are described in detail in the next section. This did not proceed without interruption. Although we tested the game multiple times, the game crashed because our website could not handle over one hundred requests at the same time. Fortunately, after a while almost everyone was able to access the game. Nevertheless, unexpected challenges can be part of introducing new technology and those looking to implement *Mad Science*, or any other technology, should acknowledge this possibility and have a plan for how to handle issues that may arise in the classroom [see also 17].

After playing the game, the students were debriefed about the three experiments with classic examples from the literature and specifically about how those concepts were being tested in the experiments the students had just experienced. At the end of class students were instructed about a homework assignment that served as preparation for the next day's class, where the students were asked to create an experiment themselves. The homework assignment stated the following:

> Your assignment is to think about and document an experiment that you would be interested in creating using *Mad Science*. Determine (1): what you want to test or explore; and (2) how you might create an environment in which this could occur.

To help the students, we posed the following five questions that students would need to think of:

1. In what setting will your experiment take place (e.g., in a classroom, on a date, at a family dinner)? Describe this setting in as much detail as possible. What does the setting look like? What are the features of the physical environment you want to create?

2. What do you want to find out (e.g., how people make decisions about something, how disclosure is related to liking, how reciprocity is related to closeness)? Explain why this interests you and why you would like to learn more about it.

3. What will you vary for the players in your scenario (e.g., the framing of a decision (positive or negative), level of disclosure (superficial or deep), level of reciprocity (at the same level as your partner or at a lower level than your partner)? Describe how this variation will change how the scenario is played.

4. What do you think will be affected by the things you vary for players (e.g., the decision that is made, how much liking is reported for a partner, how much closeness is reported by a partner)?

5. What do you expect to happen? Why do you think this will happen? Explain in detail. You might think about what you have learned in class or personal experiences you have had to help you address this question.

We further provided an example of how this assignment would need to be filled out based on the decoy effect scenario they had previously played. In return for a small amount of extra credit, students were asked to bring a hard copy of their homework assignment to class the following day. In this class, students were divided randomly over four classrooms and placed into groups of three. Their assignment was to come up with an experiment and describe it using a visual narrative on a large sheet of paper (18" x 24"). A facilitator and student teaching assistant were present to assist through this process. The homework was intended to prime a discussion among the students. Although students were given the option to pick one of the experiments from their homework assignments as a group, most groups chose to design a new experiment. As with the homework assignment, here too we provided an example of what we expected students to deliver. In the upper left corner we wanted them to sketch the setting and in the upper right the characters that would be part of their experiment. The rest of the visual narrative had to describe visually and with text how the experiment would proceed from start to finish. This process is similar to how they would do this in the game, using the in-game tools.

Creating experiments using these visual narratives went surprisingly well. Most students finished this class activity between 30-40 minutes. One of the student groups recreated the prisoner's dilemma experiment in this manner, which was interesting because when asked, the group stated they were unfamiliar with the prisoner's dilemma. Most of the created experiments revolved around dates, which is not surprising considering the topic of the class and the target audience. Figure 3 shows one of the visual narratives created by students who created an experiment with aliens and looking at how server behavior influences customer decisions.

The implementation provided us with confidence that novices in a non-research methods course can create experiments. The students in this study were at a point in their college curriculum before the need for taking research methods courses and thus would have limited exposure to research methodology. We also noticed that the students enjoyed playing the three experiments, and that our implementations seemed memorable as many students referred to the experiments on the second day of class. Future research should provide us with more conclusive answers on the efficacy of the tool for education. However, the current work does already provide empirical insights on the replicability of the experiments and in the remainder we will discuss the experiments and the results.

EXPERIMENT

Throughout this paper we have argued that *Mad Science* is a valuable tool for researching and teaching decision-making concepts. As a first step toward validating *Mad Science* as a platform for accomplishing these goals, we selected three concepts that have been shown to affect decisions and programmed scenarios meant to replicate previous findings from the literature and to provide creative digital versions to illustrate the concepts to our participants in *Mad Science*. The three scenarios were the framing effect, the decoy effect, and emotional contagion.

The framing effect is a well-supported decision bias in the literature [41]; [27].[21][20] The framing effect describes the tendency for decisions to be affected by whether outcomes are framed positively (gains) or negatively (losses). More specifically, when outcomes are framed as potential gains, decision makers tend to be risk averse, choosing certain options; however, when outcomes are framed as potential losses, decision makers are risk seeking, preferring more probabilistic gambles to certain losses. We predicted that participants would show similar risk-averse decisions when outcomes were presented as gains and risk-seeking decisions when outcomes were presented as losses in the *Mad Science* scenario.

The decoy effect is another decision bias that has been supported by a number of studies [20];[7]. The decoy effect describes the tendency of preferences between two options (e.g., two different cars) to be affected by a third, asymmetrically dominated option. Asymmetrical dominance occurs when the third option is better than one alternative but is clearly worse than the other alternative. For instance, would you rather have (1) a car that gets 40 miles per gallon of gas (MPG) for $25,000 or (2) a car that gets 30 MPG for $20,000? In this instance, it is not clear which option is better and you would expect that buyers would have no clear preference. Now, consider a third (3) car that gets 35MPG and costs $30,000. Option 3 costs more than Option 1, but it gets less MPG; however, Option 3 costs more than Option 2 and gets more MPG. Option 1 asymmetrically dominates Option 3. Although no decision maker would be expected to select Option 3, it makes Option 1 look like a better deal and would increase preference for Option 1 as opposed to a situation where Option 3 was never offered. We created a decoy effect scenario in *Mad Science*, predicting that we would replicate this finding using in-game (fictitious) currency and real time delays.

Finally, because of controversies surrounding the emotional contagion studies using Facebook [25], we wanted to create an illustrative scenario for students to participate in and later discuss. Emotional contagion is the effect of others' emotions or emotionally charged content on the emotions of individuals who come into contact with them [19]. For example, in the Facebook study, researchers controlled the content presented to a sample of users, showing largely negative content and using negative language or largely positive content and positive language. They found that participants who saw the negative content tended to use more negative words in their posts, while participants who saw the positive content tended to use more positive words in their posts. We decided to use the emotional contagion scenario because of the coverage of and controversy surrounding the Facebook study. Being able to quickly create scenarios in *Mad Science* offers the benefit of illustrating and discussing research that is currently relevant for students. A successful replication would find that player responses were affected by whether the dialogue in *Mad Science* included either positive or negative words.

METHODS

Participants

Participants were 100 undergraduate students who completed the experiment as a part of a class activity. Only 95 students completed all three scenarios. Of the participants who completed scenarios, 70 were female and 25 were male. In response to how often they played digital games, 36 said "Never", 37 said "A couple of times per year", 11 said "A couple of times per month", 8 said "A couple of times per week", and 3 said "Every day". Thus, we had a diverse sample population in terms of their previous digital game experience.

Materials and Design

Three scenarios were programmed for *Mad Science* using XML code for scripting and the Unity Game Engine (http://unity3d.com). We have since created an experiment creator tool in *Mad Science* that will allow players to create experiments without programming skills. The scenarios were loaded to a website that participants could access once the link was distributed. In *Mad Science*, participants controlled their avatar by pointing and clicking and were required to make selections, using their mouse, by clicking on text in the game. Participants clicked on arrows on a dialogue box to scroll between all options for responses (i.e., choosing an answer or choosing the dialogue with which their character would respond to other characters in the game). By requiring them to scroll through options, we were able to record whether participants viewed each option before making their selection. All response options were randomized for each participant when they started each scenario to avoid any possible order effects. Whenever we state that players were instructed or advised in the game, the instruction came from a NPC in the game.

Framing Effect. As shown in Figure 4, participants controlled an avatar in a game show setting in the Framing Effect scenario. Participants were asked to complete five trivia questions, one from each of the following categories: mathematical logic, science, animals, geography, and movies. For each question answered correctly, participants earned $100 of fictitious money. To ensure all participants ended the game show with some money in their bank, all participants started with $100. Participants were told at the beginning of the scenario, through character dialogue, that for every $100 they had in their bank account, an equal amount was being placed aside for a potential gift that could be given to a classmate at the end of the game (there was not an actual classmate receiving the gift). Participants were told that they would have to decide how the gift would be given.

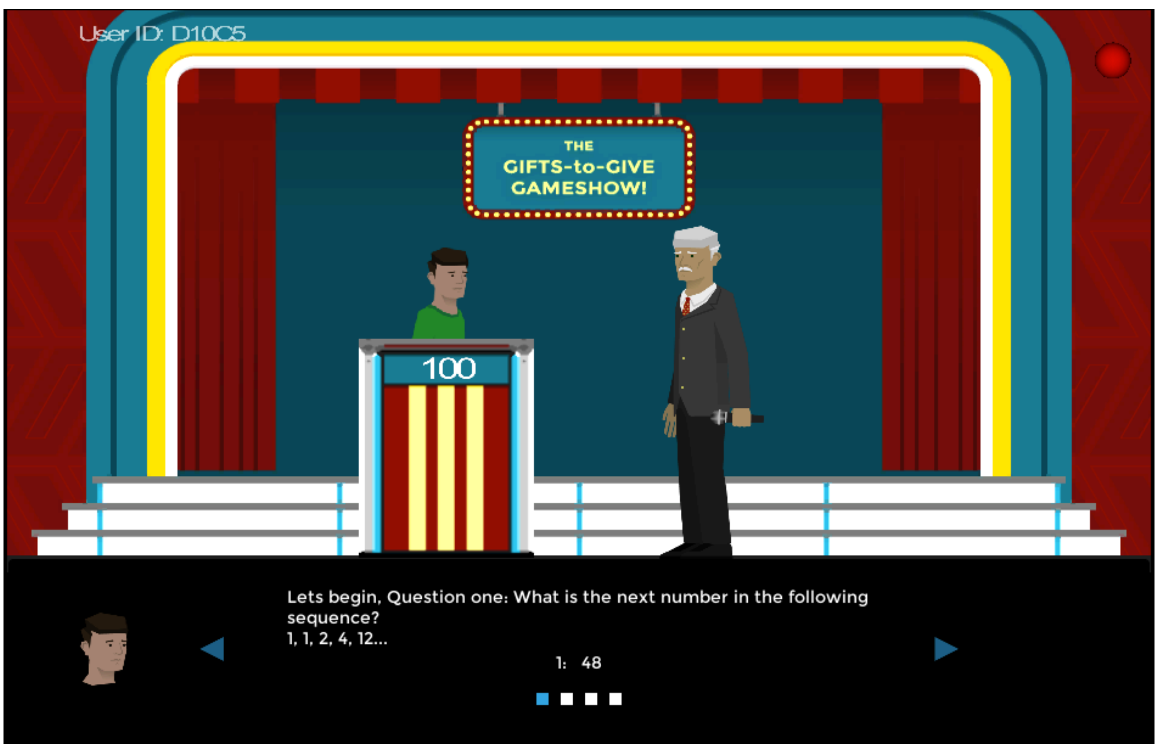

Figure 4. A screenshot from the Framing Effect scenario. The blue arrows were used to scroll through options and players clicked on the text to make a selection. The face in the dialogue box reflects which character is speaking or making a choice.

Participants were randomly assigned to one of two versions of the Framing Effect scenario. While the expected value (probability*amount) that could be given to their classmate was the same in each version, the information was provided as either a gain or a loss. In the gain frame version, participants were given the following two options for giving the gift to their fictitious classmate:

1. Of the $500 (actual amount in the bank account was shown) your classmate receives $100 (the amount in the bank was multiplied by 0.2).

2. 20% chance your classmate receives $500 (actual amount in the bank account was shown) and 80% chance they receive $0 of the $500 (actual amount in the bank account was shown).

In the loss frame version, participants were given the following two options for giving the gift to their fictitious classmate:

1. Of the $500(actual amount in the bank account was shown), your classmate loses $400 (the amount in the bank was multiplied by 0.8).

2. 80% chance your classmate loses $500 (actual amount in the bank account was shown) and 20% chance they lose $0 of the $500 (actual amount in the bank account was shown).

Research has shown that participants tend to prefer the risky option (Option 2) when the choice is presented as a potential loss and prefer the certain option (Option 1) when the choice is presented as a potential gain [41].

Decoy Effect. Figure 5 shows a screenshot from the Decoy Effect scenario. The same scene was used in the Emotional Contagion scenario. In the Decoy Effect scenario, participants were advised that they would receive $100,000 when they arrived at Madworld (the island where they would be able to conduct their own research in the future). Currently, players are unable to access Mad World thus the $100,000 is completely fictitious currency, as it cannot be saved and spent in the game at a later time. Players entered a boathouse and were required to interact with a NPC working at a sales desk. Players were given three options for boats that could be used to get them to Madworld. Each boat had an associated cost (in-game currency) and an amount of time to build (real time delays). Players were required to wait for the entire time, associated with their choice, to elapse before progressing to the final scenario. The first two options were the same for every participant. The third option randomly varied between participants. The third option varied in order to change which of the other two options asymmetrically dominated it. The following are the options presented to players:

1. Will cost $25,000.00 and will take 30 seconds to build.
2. Will cost $50,000.00 and will take 20 seconds to build.

3a. Will cost $30,000.00 and will take 31 seconds to build.

3b. Will cost $55,000.00 and will take 21 seconds to build.

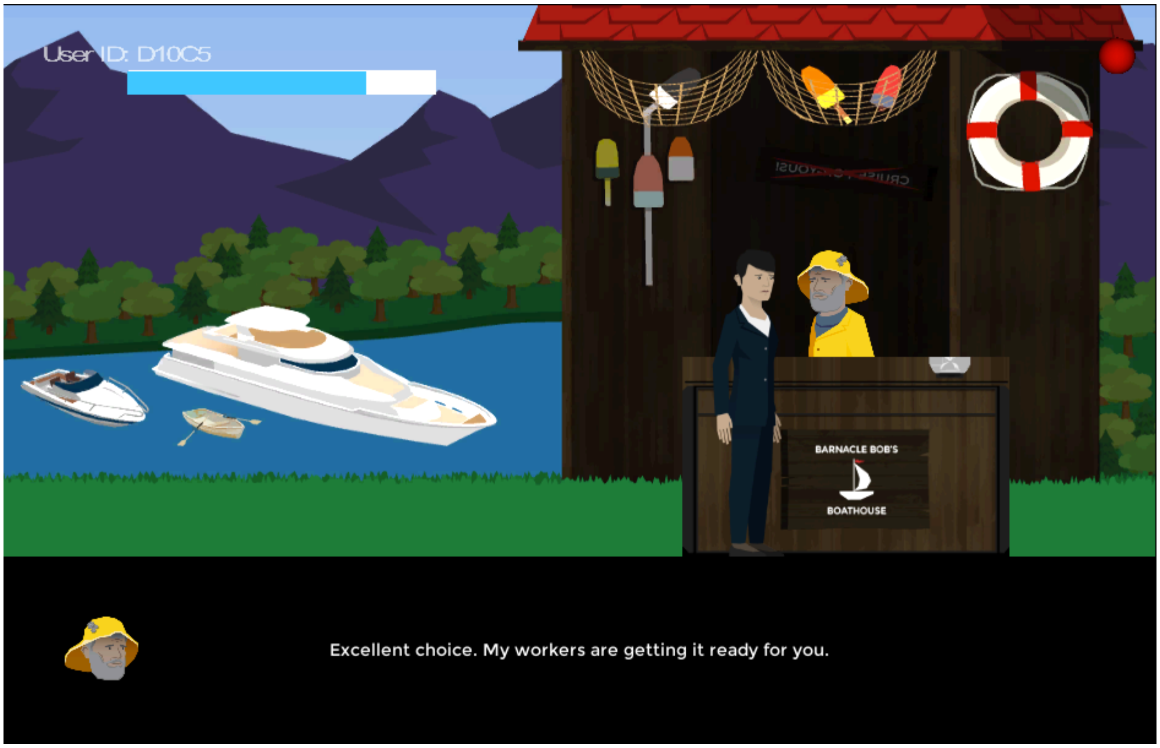

Figure 5. A screenshot from the Decoy Effect scenario. The blue rectangle in the upper left portion of the screen is the timer that visually depicts the remaining delay as time lapses.

Emotional Contagion. In the Emotional Contagion scenario, participants revisited the boathouse, this time controlling a different avatar. There were two NPCs in this scenario. The first NPC was a customer who had reportedly just interacted with the second NPC, a salesperson working at the counter. Once the scenario began, the customer walked toward the player's avatar, stopping immediately in front of the avatar to begin dialogue. The customer describes their interaction with the salesperson (attempting to return a previously purchased, defective boat). There were two versions of this scenario, one of which used all negative words (e.g., furious, frustrating, jerk) to describe the interaction and the other using all positive words (e.g., helpful, kindness, awesome) to describe the interaction. Negative and positive words were selected from the 13,915 English lemmas analyzed by [42], using higher ratings in positive and negative affect, respectively.

After the conversation with the customer, the player was required to interact with the salesperson. The salesperson asked how they were doing/how they could help them and later told the player that the item that they were attempting to purchase was out of stock and would not be in until the end of the week. After both prompts, the player was provided dialogue response options that ranged from positive to negative. For responding to the salesperson asking how they could help the player, the following options were provided for the player to choose between:

1. I feel great. Can I get a boat, please?
2. Ok, I guess. I want a boat.

3. I am irritated. Give me a boat.
4. I have been better. I need a boat.

To respond to the salesperson stating that the desired product was currently out of stock, the following response options were provided for the player to select from:

1. I will try to find it somewhere else first.
2. No thanks. I refuse to come back here again.
3. Thank you for your help. I will come back at the end of the week.

Replicating the previous findings would occur if players selected negatively worded responses when the customer used negative words to describe their interaction and positively worded responses when the customer used positive words to describe their interaction.

Procedure

Participants were instructed to bring their laptops to the classroom the day that of the activity, which was a common request for this particular course. Prior to coming to class, participants were told that they would be playing several scenarios in a game but were not told about the nature of the scenarios. Participants were shown a short video describing *Mad Science* then read an Informed Consent form, both accessible on the website and required to be accessed before beginning the scenarios. Participants completed three scenarios in the game via the *Mad Science* website provided in class. All participants completed the same three scenarios in the same order (1. Framing Effect, 2. Decoy Effect, 3. Emotional Contagion).

RESULTS

All independent and dependent variables were nominal and randomly sampled between participants. Thus, all analyses used chi-square tests of independence to evaluate the relation between variables. In all studies, there were no significant differences in the decisions as a result of gender or previous game experience thus the data was combined across these variables.

The Framing Effect. Participants were randomly assigned to either the gain or loss framing scenario. We predicted that the framing scenario would affect risk-seeking or risk-averse decisions. A chi-square test of independence was conducted to test the relation between frame and the option selected for giving a gift to a classmate. The relationship between frame and choice was statistically significant, X^2 (1, N = 100) = 17.52, $p < .01$. As shown in Figure 6, participants were more likely to select the certain option (Option 1) when the outcomes were framed as gains and more likely to select the risky option (Option 2) when the outcomes were framed as losses. These results are consistent with the previous literature thus our replication was successful.

The Decoy Effect. Of the 100 participants, only 99 completed this scenario due to bandwidth issues. To evaluate the relationship between which option dominated the third, inferior option (asymmetrical dominance) and choice preference between the other two options, a chi-square test of independence was run. The relation between asymmetrical dominance and choice preference was not statistically significant, X^2 (1, N = 99) = 2.16,

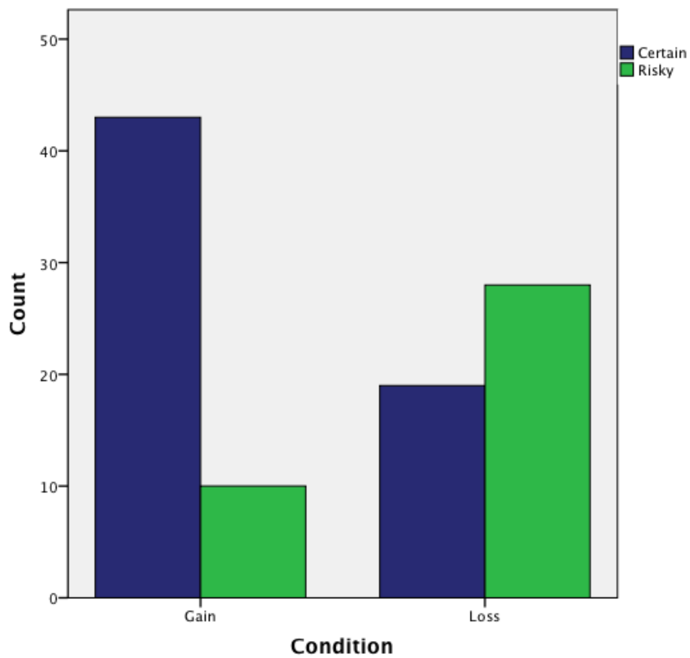

Figure 6. This graph illustrates the results of the Framing Effect. When outcomes were framed as gains, participants were more likely to select the certain option. When outcomes were framed as losses, participants were more likely to select the risky option.

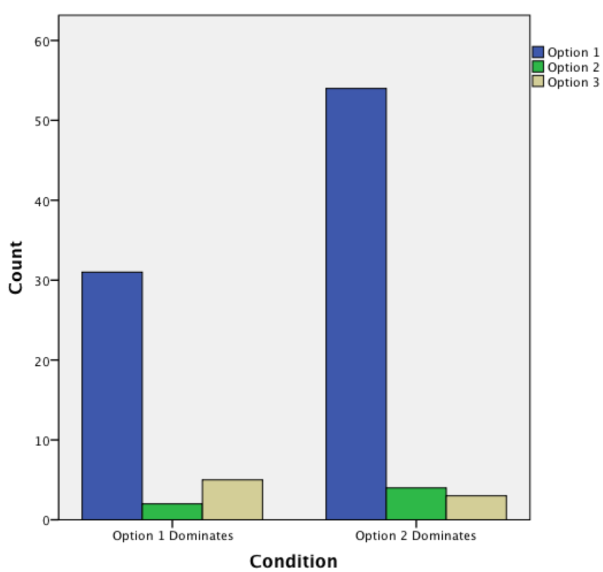

Figure 7. This graph illustrates the results of the Decoy Effect. No matter whether Option 1 or Option 2 dominated clearly inferior Option 3, most participants selected Option 1, which was the least expensive option (fictitious in-game currency).

$p > .05$. As illustrated in Figure 7, participants tended to pick Option 1 in both conditions. As a reminder, Option 1 was always the least expensive (in-game currency) option.

Emotional Contagion. Of the 100 participants, only 95 completed this scenario due to bandwidth issues. To test whether the emotional charge (positive or negative) was related to the emotional charge of dialogue responses, we ran a chi-square test of independence on both the response to how they could be helped and finding out that the item they wished to purchase was not in stock. The relationships between positively and negatively charged dialogue and the emotional charge of the responses were not statistically significant for either response (both p's $> .05$). In response to how they could be assisted, 90% of participants in each condition selected Option 1 ("I feel great. Can I get a boat, please?"). In response to the product not being in stock, 56% and 62% of participants chose Option 1 ("I will try to find it somewhere else first."), in the positive and negative conditions; 36% and 38% chose Option 3 ("Thank you for your help. I will come back at the end of the week.") in the positive and negative conditions, respectively. Almost no participants selected to respond with the negative, rude options.

DISCUSSION

Of the three concepts we tested, only the framing effect was replicated. Participants were risk-seeking in the face of loss and risk-averse in the face of gains. Although the in-game currency being won on the game show and the classmate were known to be fictitious, participants made decisions consistent with what would be predicted from past research. Participants were required to earn the money by putting in effort to answer a series of questions. It is possible that the effort invested may have made the money feel more real; however, additional studies would need to compare giving the participant money with earned money. The results of the Decoy Effect also suggest that participant decisions were under the control of the fictitious money stimulus.

In the Decoy Effect scenario, the majority of participants selected the least expensive option (measured as in-game currency) despite the longer delay times associated with that option. Although our manipulation did not replicate the decoy effect, which we argue is due to such small differences in the amount of time required to build, our results do suggest that by telling players that money in the game is important, players are sensitive to the money manipulation in the game environment. This is an interesting finding because the money has no real-world consequence but the time delay does have a real world consequence. Also, it is important to highlight that the fictitious game currency appeared to motivate participants to be interested in the outcomes of their in-game decisions. We suggest that future iterations of this scenario may include increasing the time delays or changing the setting and item being purchased to determine whether we are able to replicate previous findings.

In the Emotional Contagion scenario, participant responses were not influenced by the emotional charge of the preceding dialogue. It was, however, interesting to observe that in a virtual environment, participants still tended to prefer positive, polite responses regardless of the previous context. This may also reflect students' knowledge of the course's content, which emphasized positive, pro-social communication in relationships. This finding suggests that social norms likely continue to influence virtual behavior. It will be necessary to think through possible modifications to the present scenario that would allow us to successfully replicate previous emotional contagion research results.

Although, the results of the experiments did not replicate all of the previous research, our findings are positive for future implementation of digital games in the classroom. Additionally, observing null results opens the opportunity to discuss in the classroom what might have led to the null results and what students believe would change the results. This is an important step that researchers often have to take and has been shown to lead to an increased understanding of concepts by students [43].

For the purposes of the present study, we did not test participants' understanding of the framing effect, the decoy effect, and emotional contagion before the in-class activity. We chose not to administer a pre-test because we did not want participants to know which concepts were being tested before they played the scenarios. Going forward, researchers should validate their measure to ensure that the learning outcomes are achieved through the use of *Mad Science*. We have begun implementing mixed testing methods for our current and future prototype testing sessions.

Although we argue that *Mad Science* is a platform that researchers can use for creating experiments and instructors can use to teach concepts and research methods, there are limitations to its use to meet these goals. When individuals engage with virtual environments, there are risks that the behaviors exhibited in those environments may not truly represent the behaviors that would be exhibited in real world environments. Blascovich and Bailenson [3] argue that, for the most part, individuals do behave similarly in both virtual and real environments, but that researchers must be cautious because situations may differ in ways not previously considered. This means that we must be weary in how we interpret our results from digital games. However, through crowdsourcing research studies in a game environment, *Mad Science* provides a platform where even these questions can be tested out at a level that would not be possible at the individual level. Furthermore, learning how to replicate previous findings in digital games, researchers and students can learn what considerations might need to be made in using *Mad Science* for creating new experiments.

Future directions for *Mad Science* should include crowdsourcing for increasing the library of playable scenarios. In order to make the platform more useful in the classroom, it will be important to increase the number of replicated studies that can be used to illustrate a larger number of concepts from the decision sciences. Additionally, feedback from users must be used and integrated to ensure that the platform is user-friendly and is used to meet the goals of educating users about psychological concepts and research methods. The designers of *Mad Science* will need to work with other researchers and educators to make a meaningful platform and to help increase the use of the platform in different settings.

CONCLUSION

Our results suggest that *Mad Science* can be an excellent tool for many purposes including replicating existing research findings in the literature. Through the process of replication in the classroom, students can gain a deeper understanding of how research is conducted in the decision sciences and see how decision biases are manifested in different settings. Additionally, because of the ability to easily modify variables and settings, *Mad Science* can be used to allow students to test their own assumptions about how modifications in classic experiments or new experiments may change how decision biases are affected.

REFERENCES

1. Von Ahn, L. and Dabbish, L. Labeling images with a computer game. *ACM CHI 2004 Conference on Human Factors in Computing Systems*, ACM Press (2004), 319–326.

2. Anderson, L.W., Krathwohl, D.R., Airasian, P.W., et al. A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition. *White Plains, NY: Longman*, (2001).

3. Blascovich, J. and Bailenson, J. *Infinite reality: Avatars, eternal life, new worlds, and the dawn of the virtual revolution.* William Morrow, New York, 2011.

4. Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., and Boyle, J.M. A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education 59*, 2 (2012), 661–686.

5. Cooper, C. Pearls across the Zooniverse: When crowdsourcing becomes citizen science. *Scientific American*, 2013. http://blogs.scientificamerican.com/guest-blog/2013/02/28/pearls-across-the-zooniverse-when-crowdsourcing-becomes-citizen-science/.

6. Cooper, S., Khatib, F., Treuille, A., et al. Predicting protein structures with a multiplayer online game. *Nature 446*, (2010), 756–760.

7. Doyle, J.R., O'Connor, D.J., Reynolds, G.M., and Bottomley, P.A. The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases. *Psychology and Marketing 16*, 3 (1999), 225–243.

8. Federation of American Scientists. *Harnessing the power of video games for learning.* FAS, Washington, D.C., 2006.

9. Frank, M.C. and Saxe, R. Teaching replication. *Perspectives on Psychological Science 7*, 6 (2012), 600–604.

10. Fullerton, T., Swain, C., and Hoffman, S.S. *Game design workshop: A playcentric approach to creating innovative games.* Morgan Kaufmann Publishers, Burlington, MA, 2008.

11. Garner, M., Wagner, C., and Kawulich, B. *Teaching research methods in the social sciences.* Ashgate Publishing, Ltd., 2009.

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

13. Goedert, K., Darling, J., Shore, W., Anderson, D., and Ceynar, M. Learning about the means to the end: What US introductory psychology students report about experimental participation. *Psychology Learning & Teaching 6*, 2 (2007), 91.

14. Granic, I., Lobel, A., and Engels, R.C. The benefits of playing video games. (2013).

15. Green, C.S. and Bavelier, D. Action video game modifies visual selective attention. *Nature 423*, 6939 (2003), 534–537.

16. Harteveld, C. *Triadic game design: Balancing reality, meaning and play.* Springer, London, UK, 2011.

17. Harteveld, C. Making sense of virtual risks: A quasi-experimental investigation into game-based training. 2012.

18. Harteveld, C. and Sutherland, S.C. Finding the game in decision–making: A preliminary investigation. (2014).

19. Hatfield, E. *Emotional contagion.* Cambridge University Press ; Editions de la Maison des sciences de l'homme, Cambridge [England] ; New York : Paris, 1994.

20. Huber, J., Payne, J.W., and Puto, C. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research 9*, 1 (1982), 90–98.

21. Kahneman, D. and Tversky, A. Prospect Theory: An analysis of decision under risk. *Econometrica 47*, 2 (1979), 263.

22. Kato, P.M., Cole, S.W., Bradlyn, A.S., and Pollock, B.H. A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. *Pediatrics 122*, 2 (2008), 305–317.

23. Kiili, K. Digital game-based learning: Towards an experiential gaming model. *The Internet and Higher Education 8*, 1 (2005), 13–24.

24. Kolb, D.A. *Experiential Learning: Experience as the source of learning and development.* Prentice Hall, Upper Saddle River, NJ, 1984.

25. Kramer, A.D.I., Guillory, J.E., and Hancock, J.T. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences 111*, 24 (2014), 8788–8790.

26. Law, E. and Ahn, L. von. Human computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning 5*, 3 (2011), 1–121.

27. Levin, I.P., Schneider, S.L., and Gaeth, G.J. All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes 76*, 2 (1998), 149–188.

28. Papert, S. *Mindstorms: Children, computers, and powerful ideas.* Basic Books, New York, NY, 1980.

29. Prestopnik, N.R. and Crowston, K. Gaming for (citizen) science: Exploring motivation and data quality in the context of crowdsourced science through the design and evaluation of a social-computational system. *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on*, IEEE (2011), 28–33.

30. Quinn, A.J. and Bederson, B.B. Human computation: A survey and taxonomy of a growing field. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2011), 1403–1412.

31. Resnick, M., Maloney, J., Monroy-Hernández, A., et al. Scratch: Programming for all. *Communications of the ACM 52*, 11 (2009), 60–67.

32. Rosell, M.C., Beck, D., Luther, K., Goedert, K., Shore, W., and Anderson, D. The pedagogical value of experimental participation paired with course content. *Teaching of Psychology 32*, 2 (2005), 95–99.

33. Russell, S.H., Hancock, M.P., and McCullough, J. The pipeline: Benefits of undergraduate research experiences. *Science 316*, 5824 (2007), 548–549.

34. Seymour, E., Hunter, A.-B., Laursen, S.L., and DeAntoni, T. Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education 88*, 4 (2004), 493–534.

35. Shaffer, D.W. and Gee, J.P. The Right Kind of GATE: Computer games and the future of assessment. *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research. Charlotte, NC: Information Age Publishing*, (2012).

36. Shute, V.J., Ventura, M., Bauer, M., and Zapata-Rivera, D. Melding the power of serious games and embedded assessment to monitor and foster learning. *Serious games: Mechanisms and effects*, (2009), 295–321.

37. Squire, K.D. Replaying History: Learning world history through playing Civilization III. 2004.

38. Squire, K.D. *Video games and learning: Teaching and participatory culture in the digital age.* Teachers College Press, New York, NY, 2011.

39. Steinkuehler, C. Games for impact: Global challenges, local initiatives. 2013.

40. Tversky, A. and Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science 185*, 4157 (1974), 1124–1131.

41. Tversky, A. and Kahneman, D. The framing of decisions and the psychology of choice. *Science 211*, 4481 (1981), 453–458.

42. Warriner, A.B., Kuperman, V., and Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior research methods 45*, 4 (2013), 1191–1207.

43. White, B.Y. and Frederiksen, J.R. Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction 16*, 1 (1998), 3–118.

44. Young, M.E., Cole, J.J., and Sutherland, S.C. Rich stimulus sampling for between-subjects designs improves model selection. *Behavior Research Methods 44*, 1 (2012), 176–188.

45. Young, M.E., Sutherland, S.C., Cole, J.J., and Nguyen, N. Waiting to decide helps in the face of probabilistic uncertainty but not delay uncertainty. *Learning & Behavior 39*, 2 (2011), 115–124.

46. Young, M.E., Webb, T.L., Sutherland, S.C., and Jacobs, E.A. Magnitude effects for experienced rewards at short delays in the escalating interest task. *Psychonomic Bulletin & Review 20*, 2 (2013), 302–309.

47. Young, M.F., Slota, S., Cutter, A.B., et al. Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research 82*, 1 (2012), 61–89.