Humanatees: Character Responsibility and the Efficacy of Serious Games

A thesis presented to the academic faculty in partial fulfillment of the requirement for the Degree
Masters of Science in Game Science and Design in the College of Arts, Media and Design

by

Tyler Corwin

Thesis Committee
Committee Chair: Dr. Casper Harteveld
Thesis Advisor: Dr. Casper Harteveld
Committee Member: Dr. Celia Pearce
Committee Member: Dr. Christoffer Holmgård Pedersen

Northeastern University
Boston, Massachusetts
April 2018
Table of Contents

List of Tables 5
List of Figures 6
Abstract 7
Introduction 8
Background 10
    Serious Games for Education 10
    Player-Character Relationships and Responsibility 12
Methodology 17
    The Game: Design and Development 17
        Experimental Design 17
        Design Overview 18
        Humanatees Synopsis 18
        Built in StudyCrafter 28
    The Lecture, the Pretest, and the Post-test 29
    The Character Relationship Survey 30
The Procedure 31
    Participants 31
    Data Collection & Analysis 32
Results 32
Discussion 41
Conclusion 45
    Future Studies 47
Acknowledgements

References

Appendix
List of Tables

Table 1: Average test scores. While participants from both conditions scored close to each other on the pretest, post-test scores were significantly different between conditions.

Table 2: The means, standard deviations, and interquartile ranges of the aggregate survey scores. Each table represents each condition, with the first one including both.

Table 3: The alpha scores of the aggregate scores, both based on the original surveys’ factors and those found from the factor analysis.

Table 4: The aggregate scores and how strongly they correlate to post-test scores.

Table 5: Time taken for participants to complete the Humanatees game and survey. Less time taken suggests that participants were more engaged.

Table 6: The Game Engagement scores. Participants in the control condition felt more engaged than those in the experimental condition, but not to a significant degree.

Table 7: The means and standard deviations for the individual survey items. Each table represents each condition, with the first one including both. The sixth item, “Q06,” was omitted from the test as it did not fit with the structure of the project.
List of Figures

Figure 1: Screenshot showing the screen participants had to click “I agree” to play.

Figure 2: Sample question from pretest. The correct answer is “illegal hunting.”

Figure 3: The character selection screen. Whichever character is clicked will appear throughout Humanatees.

Figure 4: Screenshots from the scene where the participant is introduced to their character. The image on the left is from the control condition, where the participant is told about their character by an omniscient narrator. The image on the right is from the experimental condition, with the character speaking directly to the participant.

Figure 5: A screenshot from the lecture. The lecturer is about to discuss the major threats to West Indian manatees.

Figure 6: The brief conversation after the lecture. On the left, the omniscient narrator asks participants in the control group for their thoughts on the lecture. On the right, Rasa does the same for participants in the experimental group.

Figure 7: Examples of wrong and right answers being given in the post-test.

Figure 8: The first three items from the Character Relationship survey. Participants click on one of the numbers between 1 and 7 for each item depending on how they feel it applies to themselves. Lower numbers represent disagreement with the statement, and higher numbers represent agreement.

Figure 9: An item from the Game Engagement Questionnaire. Participants click either yes or no depending on if the statement in the item applied to how they felt during the game.

Figure 10: The text from the debriefing (read in the order of top image, left column, right column).

Figure 11: Aggregate scores correlations. Deeper blue or orange squares indicate stronger positive and negative correlations respectively, with the numbers in each square showing the actual Spearman correlation score between two factors.

Figure 12: The Character Relationship survey item correlations. Deeper blue dots indicate stronger correlations while orange dots indicate negative correlations.
Abstract

Serious games serve a purpose besides entertainment. This purpose can be for training, advertising, or education. Researchers and designers have considered how serious games more effectively fulfill their purpose. In this research, the role of player-character relationships is considered. Depending on the game, players may be addressed directly or made to take on the role of a character. Games can use various factors to get players attached to their character, including the feeling of responsibility. This can make the game experience more meaningful for players. The research presented in this paper sought to investigate whether feeling responsible for one’s in-game character improved the efficacy of a learning game’s outcomes. An educational game called Humanatees was designed to incite the protégé effect to see if players learned more from the game when they felt responsible for their character. The results suggest that, while participants in the experimental condition were more successful on the tests than those who used a blank slate of a character, the protégé effect was not the catalyst for this effect. Potential future research is discussed, such as the use of longer games, other teaching methods, other subjects, and other types of serious games.

Keywords: Character Attachment, Manatees, Serious games, Educational games, Game-based learning, protégé effect
Introduction

Games intended to be educational can be designed to make the game user more engaged, increasing the likelihood of the game’s success. Games, particularly video games, can be thought of as one of two main types: either recreational or serious games. The term ‘serious’ connotes a purpose to the endeavor beyond simple enjoyment and recreation. Serious storytelling is “storytelling with a purpose beyond entertainment” (Lugmayr et al., 2016). A sub-genre of this is serious games, which Susi, Johannesson, and Backlund (2007) defined as “games used for purposes other than mere entertainment… serious games usually refer to games used for training, advertising, simulation, or education” (pp. 1, 3). Serious games have been used in numerous fields, such as military, government, education, the arts, corporate training, and health. The form that a serious game will take, its design, will heavily depend on the purpose it is trying to achieve.

Depending on the purpose of the serious game, some developers will have the game address players directly while some will put players into different kinds of roles and relationships with the game character or avatar. While game characters and avatars are not the same, research that studies one can apply to the other. Avatars in computing are the representation of a user in a digital space, such as an internet forum or, in particular, a video game (Banks, 2015). According to Banks (2015), players form different kinds of relationships with their avatars based on a number of factors, including self-differentiation, emotional intimacy, and agency (between
player actions and avatar actions). Avatars represent the users in an online space. Other users perceive the user as their avatar in that space, and that user perceives the others as their avatars (Pearce, 2006). A game character is a fictional, controllable being that the player uses to inhabit the game world (Lewis et al., 2008). Game characters have their own purposes depending on the game’s design. Some act as a pre-scripted story character with their own distinctive personalities and objectives. Others are vessels for players to project themselves onto to experience the game.

The feelings players have about their characters and the kinds of relationships players form with their characters (high character attachment) can influence how they experience the game. In their study of avatars, Bowman, Rogers, Sherrick, Woolley, & Chung (2013) found that players who reported on pleasurable or exciting video games tended to feel a greater sense of control over their online avatars. Meanwhile, players who “reported on meaningful video games... were players who tended to feel stronger senses of responsibility over their avatars” (Banks & Bowman, 2013, p. 2). When discussing meaningful experiences participants had during gameplay, Banks and Bowman were specifically referring to experiences that “fostered a sense of appreciation, introspection, and self-reflection of the game experience” (p. 2). Banks and Bowman’s research suggests that there is a connection between a player’s sense of responsibility for the avatar and the game’s meaningful aspects affecting the player more than it would otherwise. This can likely include a player’s responsibility for game characters as well. Harteveld (2011), when discussing a design philosophy called Triadic Game Design, stated that serious games involve worlds of Meaning, Reality, and Play. Here, meaning can be thought of as the understanding players gain from and about the game based on the game’s purpose and how it
achieves this purpose. The research cited indicates that encouraging players to feel responsible for their character in a serious game will help the game be more successful at achieving its purpose.

The purpose of this thesis was to study if a serious game’s efficacy and outcomes improved when players felt responsible for their in-game character. An experiment was performed where participants played a serious game with a purpose whose outcome could be measured. In this study, the serious game’s purpose was to be educational, so the game was considered more effective if the participants learned more from playing it. Because of the scope of the study, no follow up was performed to measure retention.

The serious game created for this study, Humanatees, taught players about the West Indian manatee. Some participants played the game with a blank-slate of a character while others were given a character that asked for their help, making them feel a sense of responsibility for that character. The hypothesis was that a serious game that fosters a strong sense of responsibility between a player and their character will be more successful at educating the player. The serious game would be considered more successful if the student’s performance on the test for immediate learning improved.

**Background**

*Serious Games for Education*
In his book *Triadic Game Design*, Harteveld (2011) discusses his reflections on serious game design after having worked in the field. The idea of Triadic Game Design, which is the result of these reflections, is made up of three parts: Reality, Meaning, and Play. Reality refers to the reality that the game is representing or portraying, and Play is the activity the player partakes in with the game. Meanwhile, Meaning is described as the meaning players get out of the game based on its purpose and design. Humans are “meaning creators,” as he puts it (p. 126). While players can gain meaning from any game, some games are designed with the intent of conveying meaning.

Serious games are games (typically digital games) with an objective besides entertainment (Susi et al., 2007; Wouters et al., 2013). Several different industries use serious games, including the military, government, corporations, education, and healthcare, for purposes such as performing simulations, advertising, teaching, and physical fitness (exergaming). They use games for their inherent ability to simulate situations in a safe environment. Art games also fit the definition of serious games, as they are intended to produce or provoke a reaction from the player. While serious games are not necessarily the same as game-based learning, many consider the two synonymous because most serious games intend to get some message across to the user, and this requires some form of learning.

Serious games with the purpose of education have come in several forms, such as e-learning (learning enhanced through interactive technology), edutainment, and game-based-learning (Susi et al., 2007). Edutainment games refer to those made to educate school children primarily in
science, math, and reading. It was thought that these games could hold their attention through
gameplay while also teaching them. Many early edutainment games failed to fulfill their
educational aims due to what van Eck, as cited by Susi et al., called “boring games and
drill-and-kill learning.” The games were not entertaining, and they relied on memorization rather
than an understanding of the material.

Modern attempts to educate through serious games focus more on “situational and
constructionist approaches” (Susi et al., 2007). These games teach students to better understand
the content so that they can identify patterns within problems and develop solutions
independently, rather than teaching specific solutions for specific problems (e.g., learning the
mechanics of multiplication instead of memorizing multiplication tables). Serious games are
showing positive effects as supplemental educational tools. For example, education through
serious games have been shown to yield higher learning gains and improve retention, although
motivation to learn is not significantly greater than that of conventional learning methods
(Wouters et al., 2013).

**Player-Character Relationships and Responsibility**

An important facet of a game’s design is the player’s in-game character. There are two types of
characters in games: game characters and avatars. Video game characters are fictional,
controllable beings that the player uses to inhabit the game world (Lewis et al., 2008). Avatars
are on-screen representations of the player, typically in an online game or virtual world, which
are manipulated through an external device, such as a controller or touch screen (Gazzard, 2009).
It is thought that players embody the avatar in the game space; however, Banks (2015) argues that the relationship between players and avatars can be mutual. Banks’ research suggested that players can have relationships with their avatars based on three factors, each one on a spectrum from high and low: self-differentiation (the extent to which the player and avatar are experienced as distinct social agents), emotional intimacy (the degree of sentimentality players felt for their avatar), and agency (between the player and the avatar being perceived as “in-charge”). From these factors, Banks (2015) listed four different types of relationships players and avatars can have: avatar-as-object (the avatar is how the player interacts with the game and nothing more), avatar-as-Me (the avatar is Me in the game), avatar-as-symbiote (the avatar is a new persona), and avatar-as-other (the avatar is its own being). While avatars are different from game characters, research that involves one can often be applied to the other.

Lewis et al. (2008) refer to another phenomenon they called “character attachment.” Game characters are different from avatars in that avatars represent the player in online spaces. The player sees other players in the form of their avatars, and other players perceive the player as their avatar (Pearce, 2006). Characters, on the other hand, allow the game player to immerse themselves into the game world by becoming attached to the character that is defined and scripted by the designers. Character attachment is the “internalization and psychological merging of a player’s and a character’s mind” (Lewis et al., 2008, p. 515). The mechanisms that result in character attachment were defined as “an individual’s feelings of (a) friendship and (b) identification with a video game character when an individual (c) is willing to suspend disbelief, (d) feels responsible for the game character, and (e) feels in control of the game character’s
Banks and Bowman (2013) found that the latter four mechanisms of character attachment may interact with the four avatar archetypes Banks detailed. They found this based on the interviews Banks had conducted with World of Warcraft users. The comparison of playful to serious games is natural when both have avatars and characters that players use in the game space. In the interactions between player and avatar/character, each of the mechanisms would be more or less present in certain relationships; for example, avatar-as-other relationships have players showing a stronger sense of responsibility for the avatar. They care about what the avatar wants. It is possible that these players are also likely to show a stronger appreciation for the game they are playing.

Responsibility has already been shown to positively affect one’s learning. Chase, Chin, Oppezzo, and Schwartz (2009) had 5th and 8th grade students work with software to learn about biology. Some students believed they were using the software to teach themselves while others believed that they were teaching Teachable Agents (TA). Students in the latter condition spent more time on learning activities (e.g., reading) and learned more (p. 342). The students made more of an effort to learn for their TAs than they did for themselves. The researchers called this the protégé effect, and they believed that it was caused by the TAs invoking a sense of responsibility that motivated the students to learn. While the students felt responsible for distinct digital agents, invoking responsibility for characters could motivate serious-game players to learn as well.

Research done by Bowman, Rogers, Sherrick, Woolley, & Chung (2013) suggests that players who feel a sense of responsibility or care for their avatar found their game experiences more
meaningful. They had players report on both their most enjoyable video game experience and
their most meaningful video game experience. Players who reported on enjoyable games tended
to feel a great sense of control over their avatars. Players who reported on “meaningful video
games – games that fostered a sense of appreciation, introspection, and self-reflection of the
game experience – were players who tended to feel strong senses of responsibility over their
avatars” (Banks & Bowman, 2013, p. 2). Because there is evidence of this connection between
players feeling a sense of care for their avatar and their improved appreciation for the game’s
experience, it is possible a player’s feelings of responsibility for their character may make
serious games more effective at fulfilling their purpose. Methods for effecting the player’s
relationship with their character have been studied before.

Studies suggest that players typically identify more strongly with characters that they pick for
themselves. This might involve players using or making characters that look like themselves, but
that does not apply to every type of player. Yang, Huesmann, and Bushman (2014) ran a study to
see if the effects of violent video games on player aggression are moderated by the genders of
the players and the genders of their characters. Men that used male characters showed an even
stronger effect, suggested that when there are matching genders, players more strongly identified
with their characters. This would suggest that players should ideally be given characters that
match their gender in order to strengthen the player-character relationship, but some players may
prefer dissimilar characters. Soutter and Hitchens (2016), discussed two factors that influence
character identification, a concept similar to character attachment. These factors were 1) the
similarities between the character and player and 2) life-satisfaction and the game’s
requirements. The former can refer to demographic information such as race or gender, as well as personality and past experiences. For the latter, “players identify with similar characters only if their life satisfaction is high and the game requires or allows for a similar character; however, if one’s life satisfaction is low or the game requires a dissimilar character, players will identify with a dissimilar character” (p. 1033). The more dissatisfied players are with their own circumstances, or the more the game requires a dissimilar character, the more likely players are to adopt the identity of a character dissimilar from themselves. This suggests that, in order to ensure players will form a stronger relationship with their game-character, they should be allowed to choose their character for themselves.

As Harteveld (2011) said, serious games must be designed according to their specific purpose. Characters and the relationships players form with said characters are one design element that must be considered. Research suggests that players who feel responsible for their character are more engaged in the game and therefore more likely to learn. This study intended to investigate whether an increased sense of responsibility for the character caused players to perform better in and learn more from the serious game.

In order to invoke a sense of responsibility in the participant, the protégé effect was used as a game mechanic. Suttie et al. (2012) and Arnab et al. (2014) sought to map game mechanics to learning mechanics in order to aid the development of serious games. Both sets of researchers found that the protégé effect, when used as a game mechanic in serious games, taught and encouraged players to take responsibility. According to Arnab et al. (2014), “The protégé effect
is not a learning goal but is the [Serious Game Mechanic] through which the goal can be achieved” (p. 401). The protégé effect is a tool through which responsibility is taught, and this responsibility can improve the efficacy of a learning game’s outcomes. Thus, Humanatees used the protégé effect to incite responsibility in users, and it did this by giving users a game character who needed their help.

**Methodology**

**The Game: Design and Development**

*Humanatees* is a game where the player chooses a character and attends a lecture about manatees. After a short break, the player is then quizzed on what they learned in the lecture. The player then fills out surveys regarding their relationship with their character and their engagement with the game.

**Experimental Design**

The main manipulation in this study was whether participants had a teachable agent of a character to incite the protégé effect. Participants were assigned to either the experimental condition or the control condition. In the experimental condition, participants were introduced to a character who asked them for help in learning about manatees. This was meant to create a feeling of responsibility for the character in participants. In the control condition, participants were told that they would need to learn about manatees using a blank slate character to interact with the
game world. The learning outcomes that would result from these two conditions were measured using the results of a test that asked participants questions about manatees.

Design Overview

*Humanatees* is a narrative-based game. The player clicks to read through lines of text. This text comes from an omniscient narrator as well as the player character and the teacher. For the latter two, the box containing the text will be connected to the speaker to indicate that they are talking. Occasionally, the player will be presented with two or more options for how they want to respond to a question. The player must click on one of the options, at which point different follow-up dialogue will be shown depending on what option the player chose. The player will also be prompted to give an open-ended answer to a question at least once when they play. Here, a text field will appear for them to type their answer into. In these cases, their answer is not recorded but were used to engage them in the narrative.

**Humanatees Synopsis**

The game starts with a briefing that explains what the player will do in the game and that there are no potential risks that may come from playing (e.g., the player can put in their name, but it is not recorded). To continue, the player must click “I agree” to show that they agree to the game’s conditions. The briefing read:

“In this experiment, you will be learning about and taking a test on manatees, specifically West Indian manatees. You will also take some surveys related to your engagement with the experiment. Only data
related to your test scores and survey responses will be recorded. As part of the experiment, you may have
the option of inputting your first name during this game. If you choose to do so, this information will not
be saved or recorded. If these terms sound agreeable to you… And if you are 18 years of age or older…
Please click “I agree.” This will be used as your way to show that you consent to these conditions.”

Figure 1: Screenshot showing the screen participants had to click “I agree” to play.

After the player accepts to the briefing, they are made to answer multiple-choice questions about
West Indian manatees. These questions are based on the information players will learn later
during the in-game lecture, and the questions act as a pre-test to measure how much players
already knew about West Indian manatees before the game started. Each question has three
choices and the option for the player to say “I don’t know.” This was included to allow players
who do not know the material to skip each question. If they were made to guess, then there was
the chance that they would get the question right, which would give the wrong impression of
their previous knowledge. Of course, some participants might have guessed regardless, but the option was available for those who were less confident in their guesses. The pretest can be found in Appendix 1.

![Sample question from pretest](image)

*Figure 2: Sample question from pretest. The correct answer is “illegal hunting.”*

In the next scene, the player is asked to choose a character from a set of eight: four men, four women, each with different skin-tones. The chosen character will appear in the following scenes. Based on the findings from Yang et al. (2014) and Soutter and Hitchens (2016), players should be able to choose the character that is most similar or dissimilar to themselves in order form a stronger relationship with that character depending on their life-satisfaction. Once the player has chosen their character, the game continues to the next scene.
The player is then introduced to their character and randomly placed into one of the two conditions. In the control condition, the player is shown their chosen character and told by an omniscient narrator that the player must use the character to attend a lecture. The narrator also shows the player that the character will say whatever choice the player makes, which will be used when the player takes the post-test.

In the experimental condition, the player is introduced to their chosen character, who introduces himself/herself as “Rasa” (based on the term Tabula Rasa, the philosophical idea that the human mind starts as a “blank slate”). Rasa asks the player for their name, to which the player has the option to give it to Rasa through an open response. Rasa tells the player that he/she has been attending a class to learn about manatees, but he/she has trouble paying attention due to the

*Figure 3*: The character selection screen. Whichever character is clicked will appear throughout Humanatees.
lecturer’s dry way of speaking. Rasa asks the player to attend the lecture to help him/her learn its content. This exchange is meant to incite the protégé effect within the player, as the player learns that Rasa needs their help with the class. The player is encouraged to learn for Rasa’s sake, giving them a sense of responsibility.

Figure 4: Screenshots from the scene where the participant is introduced to their character. The image on the left is from the control condition, where the participant is told about their character by an omniscient narrator. The image on the right is from the experimental condition, with the character speaking directly to the participant.

The next scene has the player read through a lecture being given by a teacher. The lecture starts by explaining some of the causes behind the endangerment of the West Indian manatee before listing some trivia facts about the species. West Indian manatees were primarily chosen as the subject of the lecture to lower the likelihood of participants knowing the information before having played the game. This information was obtained from a number of different sources (Gaspard III et al. (2013), Bauer (2010), Quintana-Rizzo et al. (2008), Reynolds (1981), and Runge et al. (2017)). This scene is the same for participants in both conditions.
Following the lecture, the player has a brief conversation with either the omniscient narrator for those in the control condition or Rasa in the experimental condition. This scene contains similar text and choices between the two conditions, and it was designed to take up the same amount of time between participants in both conditions. This was to keep the participants’ time with their character the same across conditions, as well as to keep participants in one condition from forgetting more of the lecture than those in the other condition. This section is primarily meant to give time between the lecture and the post-test to prevent a ceiling effect on the test scores. For players in the experimental condition, it also gives them an opportunity to get to know Rasa and feel more responsible for him/her.
The brief conversation after the lecture. On the left, the omniscient narrator asks participants in the control group for their thoughts on the lecture. On the right, Rasa does the same for participants in the experimental group.

The player is then presented with the post-test (Appendix 2). The teacher asks the player each question one at a time, and the player must choose their answer from one of four. When they choose an answer, the player character will say the question out-loud. This gives the player the impression that they are answering the questions as the character in the control group. In the experimental group, the player is instead giving Rasa the answer to each question to help them learn. If the player chooses the wrong answer, the teacher will say what the answer was, otherwise she will say that the player’s answer was correct. This helps players learn the material regardless of whether they remembered the lecture. When the test is completed, the player is given their scores from the pre-test and post-test.
After the post-test, the player is given a survey to evaluate their relationship with their player character. The survey was adapted from the character attachment survey by Lewis et al. (2008), with the survey item “I daydream about my character” was removed, as the player’s time with their character was so short. Additionally, seven items from Banks and Bowman’s (2015) Player-Avatar Interaction scale were added. The player responded to these items on a 7-point likert scale. The survey can be found in Appendix 3.

*Figure 7: Examples of wrong and right answers being given in the post-test.*
Figure 8: The first three items from the Character Relationship survey. Participants click on one of the numbers between 1 and 7 for each item depending on how they feel it applies to themselves. Lower numbers represent disagreement with the statement, and higher numbers represent agreement.

The player then takes the Game Engagement Questionnaire (Brockmyer et al., 2009). Each questionnaire item states how the player feels while playing the game, and the player responds either “yes” or “no” depending on if they did feel that way while playing Humanatees. This survey was included to check the level of engagement the player had while playing the game. If they get a low engagement score, that suggests that they were not seriously attending to the game or not engaged by its contents.
Figure 9: An item from the Game Engagement Questionnaire. Participants click either yes or no depending on if the statement in the item applied to how they felt during the game.

The end of the game is the debriefing. The player is told that the true purpose of Humanatees was to see if feeling responsible for the game character improves the efficacy of a serious game. The two conditions and the purpose of each test and survey are explained. The game gives the player the researcher’s contact information in case they have any questions, and they are asked to share the game with others while keeping the true purpose of the game a secret.
Our lab makes serious games, and we're always trying to find out how to improve their effectiveness using various techniques.

The purpose of this study was to understand how the feeling of responsibility for a character influenced the efficacy of a serious game at fulfilling its purpose.

In this case, that purpose was education.

There were two conditions. The control condition had the player use a blank slate of a character to answer the questions from the quiz.

In the experimental condition, the player interacted with and was made to feel responsible for their character.

The hypothesis was that players in the experimental condition, who felt responsible for their avatar, would learn more from the manatee lecture.

The first test was to measure how much about manatees you already knew. The second test was to measure how much you had learned.

That test with the 7 point likert scale asking about the character was to measure your friendship with the character, your suspension of disbelief, your feeling of control over the character's actions, your feeling of responsibility, your emotional investment, and how autonomous the character felt to you.

The last question was to measure how engaged you were during the experiment.

Well, I hope you enjoyed this experiment.

If you can, please share this with people you know. It will help ensure that the study has enough data to use.

... but...

Please do NOT tell anyone what the study is really about.

If you have any questions, you can contact me via email: corwin.t@husky.neu.edu

Thank you for your time.

**Figure 10**: The text from the debriefing (read in the order of top image, left column, right column).

**Built in StudyCrafter**

*Humanatees* was created in *StudyCrafter* (2018), a free program developed by the Northeastern Game Studio at Northeastern University, led by Prof. Casper Harteveld. *StudyCrafter* allows users to create research projects and share them online. These projects are designed to specifically research human behavior. Within these projects, users create scenes that situate the player into specific scenarios and contexts, including those that would be difficult to observe in real life. Users also designate the variables within the project, including independent and
dependent variables, and can observe the results once the player has finished playing the project. *Humanatees* was a project within *StudyCrafter* platform. A link to the *StudyCrafter* website was provided to participants in order for them to play *Humanatees*.

**The Lecture, the Pretest, and the Post-test**

In order to measure the effectiveness of the study condition, in terms of how well the participants learned from using the serious game, two tests evaluated how much participants learned about manatees by comparing what they knew before and after the lecture, similar to research done by Papastergiou (2008). Participants are presented with the pretest after they have agreed to the in-game briefing. This pretest asks participants one multiple-choice question at a time. The pretest and post-test each use six questions so as not to overwhelm participants. The pretest is meant to measure how much the participants already knew about West Indian manatees. Because of this, an “I don’t know” option was added to each of the multiple choice answers, making for a total of four choices for each question. This “I don’t know” option allows participants who do not know the answer to each question to move on without guessing the answer, which would have potentially given an unreliable assessment of their pre-existing knowledge about West Indian manatees. The questions from the pretest are shown in Appendix 1.

The post-test is given after the participant has read the lecture. After discussing the lecture with either the character (in the experimental group) or with the omniscient narrator as the character (in the control group), the participant is taken to a scene where the lecturer asks them each post-test question. There are six multiple-choice questions total, and each question has four
answers to choose from. If the participant answers right, the lecturer tells them they are correct, and if they answer wrong, the lecturer tells them what the correct answer was. Unlike the pretest, there is no “I don’t know” option. The questions from this post-test are shown in Appendix 2.

The questions for the pretest and post-test were created using the information given in the in-game lecture. This lecture given through a series of text boxes that the participant must read through. West Indian manatees were chosen as the subject of the lecture to lower the chance of participants knowing the information in the lecture beforehand. Additionally, it was believed that participants would be more enticed to play a game about manatees compared to one about some other sea life. The content of the lecture is based on research from a variety of sources (Gaspard III et al. (2013), Bauer (2010), Quintana-Rizzo et al. (2008), Reynolds (1981), and Runge et al. (2017)).

**The Character Relationship Survey**

The Character Relationship survey used in *Humanatees* was based on two existing surveys. The first 17 items are from the character attachment survey created by Lewis et al. (2008). These items measure factors in character attachment, including identification/friendship, suspension of disbelief, control, and responsibility. This survey fit the needs of this research, though the sixth item, “I daydream about my character,” was excluded from this experiment, as participants would not spend enough time with their character for this question to fit into the assessment. More recent research was also used to supplement this survey. Items from Banks and Bowman’s (2015) Player-Avatar Interaction scale were added. This scale measures factors related to
emotional investment, anthropomorphic autonomy, suspension of disbelief, and control. Several of the items overlap with those from the character attachment survey, but this scale was intended to measure how players interacted with their online avatars rather than characters, making it less relevant to this study. Thus, only items regarding emotional investment and anthropomorphic autonomy were adapted from this scale for use in the survey for Humanatees. The items from this new survey, as well as the factors that each item measures, can be seen in Appendix 3.

The Procedure

Participants

Participants were recruited from the Northeastern Campus and through various online outlets. Northeastern residents were invited personally to participate, though there was no follow-up to see if they had participated. Online, Humanatees was posted on the social media sites Tumblr, Twitter, and Facebook. An ad for the game was also posted on Northeastern University’s MyNEU events page.

The game was advertised online as letting participants help progress the science of game design while also learning about manatees. The ad also stated that only users age 18 or older could participate and that the game would take 15 minutes. A link to the game and contact information for the researcher were provided in this ad. Participants were also encouraged to share the game with those they knew. Through these recruitment methods, 113 users participated, with 92 of those completing the game and submitting their results. Two participants were removed from the
collected data, as they were outliers based on boxplot analyses of the post-test and investment scores.

**Data Collection & Analysis**

Once the data had been collected, it was analyzed in Rstudio. Because the survey used in *Humanatees* was created from two pre-existing surveys, the scale reliability was analyzed to test if the factors in the survey are predictive of one another. Additionally, to find if responsibility or any of the other factors in Character Attachment or player-avatar interactions affected participants’ learning, correlation and Mann Whitney U tests were performed to find the relationship between the factors and the test scores. Lastly, linear regression analyses were performed to find if any of the factors from the survey acted as mediating variables between the two conditions in the experiment and the post-test scores.

**Results**

Data from 90 participants was analyzed. The average pretest and post-test scores for participants in either condition are shown in Table 1. On average, pretest scores were higher for participants in the experimental condition \((M = 0.98, SD = 1.03)\) than they were for the control condition \((M = 0.95, SD = 1.02)\). Similarly, participants in the experimental condition had higher average post-test scores \((M = 5.18, SD = 0.83)\) than those in the control condition \((M = 4.76, SD = 0.91)\). The difference between the average pretest scores for each group is 0.03. The difference between the average post-test scores was 0.48.
Table 1: Average test scores. While participants from both conditions scored close to each other on the pretest, post-test scores were significantly different between conditions.

<table>
<thead>
<tr>
<th></th>
<th>Both (Pretest)</th>
<th>Both (Post-test)</th>
<th>Experimental (Pretest)</th>
<th>Experimental (Post-test)</th>
<th>Control (Pretest)</th>
<th>Control (Post-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.97</td>
<td>4.97</td>
<td>0.98</td>
<td>5.18</td>
<td>0.95</td>
<td>4.76</td>
</tr>
<tr>
<td>Std Dev</td>
<td>1.02</td>
<td>0.89</td>
<td>1.03</td>
<td>0.83</td>
<td>1.02</td>
<td>0.91</td>
</tr>
</tbody>
</table>

The most common pretest score was 0 out of 6 correct answers, and the most common post-test score was 5 out of 6. This is true for both conditions. With the Individual test questions, both conditions commonly got all of the pretest questions wrong. In the post-test, the experimental condition commonly got all 6 post-test questions right, while the control condition commonly got the first question, “The scientific name for the Florida manatee subspecies is...”, wrong and the other five right. That said, the experimental condition also struggled with the first question more than the others five, with just over half getting it right ($M = 0.6, SD = 0.5$). Because the difference between average pretest scores in each condition was so low, and because the most common pretest score was zero, only the post-test score will be discussed when analyzing the rest of the results. Based on the results of a Mann-Whitney U test, the post-test scores were significantly different between conditions, with participants in the experimental condition scoring higher, $U = 1287, p = 0.02, r = 0.25$. Therefore, participants who had Rasa as their character performed better on the post-test.
For the Character Relationship survey, Table 2 shows the means and standard deviations of the aggregate survey scores. The aggregate scores were normalized in order to more clearly assess them. The means and standard deviation for the individual survey items can be found in Table 7 in the Appendix.

Table 2: The means, standard deviations, and interquartile ranges of the aggregate survey scores. Each table represents each condition, with the first one including both.

<table>
<thead>
<tr>
<th></th>
<th>Friendship</th>
<th>Suspension of Disbelief</th>
<th>Control</th>
<th>Responsibility</th>
<th>Emotional Investment</th>
<th>Anthropomorphic Autonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.45</td>
<td>0.49</td>
<td>0.64</td>
<td>0.52</td>
<td>0.37</td>
<td>0.45</td>
</tr>
<tr>
<td><strong>St Dev</strong></td>
<td>0.23</td>
<td>0.25</td>
<td>0.24</td>
<td>0.22</td>
<td>0.2</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>IQR</strong></td>
<td>0.34</td>
<td>0.36</td>
<td>0.3</td>
<td>0.32</td>
<td>0.3</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>(both)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.47</td>
<td>0.56</td>
<td>0.57</td>
<td>0.57</td>
<td>0.38</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>St Dev</strong></td>
<td>0.2</td>
<td>0.25</td>
<td>0.26</td>
<td>0.23</td>
<td>0.21</td>
<td>0.24</td>
</tr>
<tr>
<td><strong>IQR</strong></td>
<td>0.19</td>
<td>0.36</td>
<td>0.4</td>
<td>0.33</td>
<td>0.3</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>(control)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.43</td>
<td>0.42</td>
<td>0.71</td>
<td>0.46</td>
<td>0.36</td>
<td>0.34</td>
</tr>
<tr>
<td><strong>St Dev</strong></td>
<td>0.26</td>
<td>0.24</td>
<td>0.19</td>
<td>0.18</td>
<td>0.2</td>
<td>0.24</td>
</tr>
<tr>
<td><strong>IQR</strong></td>
<td>0.38</td>
<td>0.36</td>
<td>0.2</td>
<td>0.28</td>
<td>0.3</td>
<td>0.33</td>
</tr>
</tbody>
</table>

A factor analysis was also performed on the individual items. The factors formed from analyzing the items from the Player-Avatar Interaction scale were consistent with the factors the original survey measured. The same was true when analyzing the items from the Character Attachment survey except for the items “I could see myself being attracted to my character,” “I get frustrated when my character does not perform the way I want them to,” and “I make decisions with my character’s best interests in mind” (items 5, 14, and 17 respectively), which failed to load properly. When the items, as they were presented in Humanatees, were analyzed together, the former items 5 and 14 failed to load while item 17 loaded in the factor related to emotional investment rather than responsibility. Additionally, item 3, “I consider my character a friend of
mine,” loaded in the emotional investment factor rather than the identification/friendship factor.

It is possible these items loaded on emotional investment because players who were more emotionally invested in their character made decisions with their character in mind and considered them a friend. Besides these former four items, all other items loaded on the factors they were originally designed for.

As the survey used in *Humanatees* was composed of two other surveys, the reliability of the scale and its different factors were tested using Cronbach’s alpha. The alpha score for the individual items showed high reliability ($\alpha = 0.81$). Additionally, the reliability of the individual factors were found to be generally high (Table 3). The difference in reliability between the factors from the original surveys and the factors found through the factor analysis was small overall except for responsibility ($\alpha = 0.76$) and factor 5 ($\alpha = 0.86$). Factor 5 contained the items that comprised the responsibility factor from the Character Attachment survey except item 17, “I make decisions with my character’s best interests in mind,” which had moved to Factor 1.

<table>
<thead>
<tr>
<th></th>
<th>Emotional Investment/ Factor 1</th>
<th>Suspension of Disbelief/ Factor 2</th>
<th>Control/ Factor 3</th>
<th>Autonomy/ Factor 4</th>
<th>Responsibility/ Factor 5</th>
<th>Friendship/ Factor 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alpha for original factors</strong></td>
<td>$\alpha = 0.85$</td>
<td>$\alpha = 0.85$</td>
<td>$\alpha = 0.77$</td>
<td>$\alpha = 0.80$</td>
<td>$\alpha = 0.76$</td>
<td>$\alpha = 0.77$</td>
</tr>
<tr>
<td><strong>Alpha for analysis factors</strong></td>
<td>$\alpha = 0.86$</td>
<td>$\alpha = 0.85$</td>
<td>$\alpha = 0.77$</td>
<td>$\alpha = 0.80$</td>
<td>$\alpha = 0.86$</td>
<td>$\alpha = 0.76$</td>
</tr>
</tbody>
</table>

Table 3: *The alpha scores of the aggregate scores, both based on the original surveys' factors and those found from the factor analysis.*
Because the reliability of the responsibility factor was improved by the removal of item 17, and because that item loaded on the emotional investment factor during the factor analysis, item 17 was moved to the latter factor. Additionally, because identification/friendship showed a negligible drop in reliability when items 3 and 5, “I consider my character a friend of mine” and “I could see myself being attracted to my character” respectively, were omitted, the former was moved to emotional investment, and the latter was dropped completely, having failed to load during the factor analysis. Finally, item 14, “I get frustrated when my character does not perform the way I want them to,” failed to load during the factor analysis, and its removal does not impact the control factor’s reliability, so it was also dropped. The factors from the factor analysis were used for the rest of the data analysis.

Responses to the survey showed differences in how the two conditions affected them. The feelings of friendship or identification toward the character were not significantly different between conditions, $U = 1017, p = 0.97, r < 0.01$. Participants in the experimental condition suspended their disbelief significantly more than those in the control condition, $U = 1324, p = 0.01, r = 0.27$. Participants in the control condition felt significantly more control over their character than those in the experimental condition, $U = 679, p < 0.01, r = 0.29$. Participants in the experimental condition felt significantly more responsible for their character than those in the control condition, $U = 1310, p = 0.02, r = 0.26$. Participants’ emotional investment for their character was not significantly different between conditions, $U = 1108, p = 0.44, r = 0.08$. 
Participants in the experimental condition felt that the character had significantly more autonomy than those in the control condition, $U = 1461, p < 0.00, r = 0.38$.

Spearman correlations were performed on the aggregate scores (Figure 11). The correlations between the individual survey item scores shown in Figure 12 in the Appendix. These correlations suggest that, in addition to the individual items correlating strongly with items that measure the same factor, items related to identification/friendship and emotional investment with the character were also strongly correlated ($r = 0.54, p < 0.00$). Additionally, identification/friendship and control over the character were correlated ($r = 0.43, p < 0.00$). Identification/friendship and responsibility for the character were correlated ($r = 0.21, p < 0.05$). Responsibility and emotional investment were strongly correlated ($r = 0.46, p < 0.00$). Emotional investment and control were correlated ($r = 0.34, p < 0.01$). Lastly, emotional investment and anthropomorphic autonomy were correlated ($r = 0.33, p < 0.01$).
The correlations between the post-test scores and the aggregate scores were also analyzed (Table 4). Control had a significant negative correlation with post-test scores, suggesting that the feeling of control over the character hurt post-test scores, although the correlation was weak ($r = -0.22$, $p = 0.04$). None of the other factors had a significant correlation with post-test scores, suggesting that these factors had little effect on the post-test scores.

Figure 11: Aggregate scores correlations. Deeper blue or orange squares indicate stronger positive and negative correlations respectively, with the numbers in each square showing the actual Spearman correlation score between two factors.
Table 4: *The aggregate scores and how strongly they correlate to post-test scores.*

<table>
<thead>
<tr>
<th></th>
<th>Friendship</th>
<th>Suspension of Disbelief</th>
<th>Control</th>
<th>Responsibility</th>
<th>Emotional Investment</th>
<th>Autonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation with Post-test scores</td>
<td>$r = -0.16$</td>
<td>$r = 0.02$</td>
<td>$r = -0.22$</td>
<td>$r = -0.05$</td>
<td>$r = -0.02$</td>
<td>$r = 0.12$</td>
</tr>
<tr>
<td>P-values of the above correlations</td>
<td>$p = 0.14$</td>
<td>$p = 0.64$</td>
<td>$p = 0.04$</td>
<td>$p = 0.62$</td>
<td>$p = 0.87$</td>
<td>$p = 0.28$</td>
</tr>
</tbody>
</table>

A linear regression analysis was also performed to see if any of the factors from the survey had a mediating effect on the relationship between the two conditions and the post-test scores. The condition affected responsibility ($p = 0.02$). When responsibility was introduced as a mediator between condition and post-test score, the effect of responsibility was not significant ($p = 0.26$) while the effect condition had on post-test score remained significant ($p = 0.01$). While the experimental condition showed improved learning outcomes and higher feelings of responsibility, responsibility itself did not seem important for the learning outcome.

Control had a non-significant mediating effect on relationship between condition and post-test score. Control is significantly affected by condition ($p < 0.01$). When control was introduced as a mediator, the relationship between condition and post-test score is no longer significant ($p = 0.13$), but the relationship between control and post-test score is not significant ($p = 0.09$).

Similarly, anthropomorphic autonomy was significantly affected by condition ($p < 0.00$). When autonomy was introduced as a mediator, the relationship between condition and post-test score
was no longer significant \( p > 0.05 \), but the relationship between autonomy and post-test score was not significant \( p = 0.77 \).

The amount of time participants took to complete the project was recorded along with the Game Engagement Questionnaire scores to see if participants in either condition were more engaged with the game (Table 5 and Table 6). There was a significant negative correlation between time \( (M = 667.15, SD = 225.18) \) and engagement \( (M = 7.34, SD = 4.39) \), \( r = -0.21, p = 0.04 \). This suggests that participants who took less time to complete the project were more engaged. The time participants in the experimental condition took \( (M = 692.98, SD = 209.59) \) was not significantly different from those in the control condition \( (M = 641.12, SD = 239.32) \), \( U = 1216, p = 0.1, r = 0.17 \). Participants in the experimental condition \( (M = 6.69, SD = 4.28) \) were also not significantly more engaged in the game than those in the control condition \( (M = 8.0, SD = 4.45) \), \( U = 836, p = 0.15, r = 0.15 \).

Table 5: Time taken for participants to complete the Humanatees game and survey. Less time taken suggests that participants were more engaged.

<table>
<thead>
<tr>
<th>(Time Taken)</th>
<th>Both</th>
<th>Experimental</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>667.15</td>
<td>692.98</td>
<td>641.12</td>
</tr>
<tr>
<td>Std Dev</td>
<td>225.18</td>
<td>209.59</td>
<td>239.32</td>
</tr>
</tbody>
</table>

Table 6: The Game Engagement scores. Participants in the control condition felt more engaged than those in the experimental condition, but not to a significant degree.

<table>
<thead>
<tr>
<th>(Engagement)</th>
<th>Both</th>
<th>Experimental</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>7.34</td>
<td>6.69</td>
<td>8.0</td>
</tr>
<tr>
<td>Std Dev</td>
<td>4.39</td>
<td>4.28</td>
<td>4.45</td>
</tr>
</tbody>
</table>
Discussion

The hypothesis was partially supported by the results. The hypothesis was that players would learn more from a serious game that fosters a strong sense of responsibility between a player and their character. While both the post-test scores and the feeling of responsibility for the character were both significantly higher in the experimental condition, post-test scores and responsibility were not correlated. Additionally, responsibility was not found to be a mediating variable between the conditions and post-test scores. This suggests that, while participants in the experimental condition showed improved learning outcomes, responsibility was not the cause.

Humanatees was successful at teaching participants about manatees. The average pretest scores for each condition were lower than 1, with participants most commonly getting none of the answers right. Additionally, the difference between the average pretest score for each condition was incredibly low (difference of 0.03). The post-test scores, on the other hand, were higher than the pretest scores for both the experimental ($M = 5.18, SD = 0.83$) and control conditions ($M = 4.76, SD = 0.91$), with participants in the experimental condition scoring higher than the control condition (difference of 0.48, $p = 0.01$). The question participants most commonly got wrong, especially in the control condition, was the first question in the post-test, “The scientific name for the Florida manatee subspecies is...”. This was likely because the scientific name was difficult to remember and easily confused with the fake names given in the wrong answers.

Overall though, these results suggest that participants in both conditions came into the
experiment with little to no knowledge of manatees, and they both learned from the in-game lecture, but those in the experimental condition scored higher on the post-test. Something about the experimental condition had an effect on participants’ learning.

While most items from the Character Relationship survey were reliable and validated, several items were moved or removed in the factor analysis. Items 5 and 14, “I could see myself being attracted to my character” and “I get frustrated when my character does not perform the way I want them to,” failed to load during the analysis. It could be that these two items were not as relevant to Humanatees players, as players did not spend enough time with their character to feel attracted to the character nor were there opportunities for the character to behave in a way the player did not want them to. Items 3 and 17, “I consider my character a friend of mine” and “I make decisions with my character’s best interests in mind,” were moved to the factor related to emotional investment. It is possible that the former item was more closely related to the emotions of participants than identification/friendship. Similarly, while items 15 and 16 related to understanding what the character wanted or needed, participants who agreed with item 17 may have felt more emotionally invested in their character rather than feeling responsible for them. It is also possible that the sample size from the experiment was too small for a factor analysis, with only 90 responses. Regardless, the majority of factors from the original surveys were replicated in the factor analysis and were shown to be reliable.

Some factors from the Character Attachment survey and the Player-Avatar Interaction scale were affected by the conditions. The participants’ reports of their suspension of disbelief, their
feelings of responsibility for the character, and the character’s anthropomorphic autonomy were significantly higher when participants interacted with a character who was distinct from themselves. Suspension of disbelief was stronger in the experimental condition perhaps because social feedback (in this case talking/speech) actively drew participants in and was more involving than having a blank slate character. Responsibility was intentioned to be stronger in the experimental condition through the protégé effect, and these results suggest that it succeeded in that regard. Lastly, because Rasa was a character with dialogue in the experimental condition, that seemed to cause participants to think of Rasa as an autonomous agent.

The feeling of control over the character was higher for those in the control condition. As the player was repeatedly told and shown that they control the character’s actions, this makes sense. The feelings of identification/friendship and emotional investment for the character, however, were not different between conditions. These latter two factors were also strongly correlated with one another. Both involve the player growing to care about or identify with their character. Perhaps while the game created a sense of responsibility in participants for the character, participants did not care for the character itself.

What caused the difference in test scores between the conditions is unclear. Responsibility was not found to be a mediating variable between the conditions and post-test scores. The control condition, the condition that was given a blank slate character, felt more control over their character, as stated above. While none of the other factors from the survey were significantly correlated with the results of the post-test, there was a weak negative correlation between feeling
control over the character and the post-test scores ($r = -0.22, p = 0.04$), suggesting that participants who felt more control over their character performed worse on the post-test. The lecture that taught participants about West Indian manatees never required the players to control their character, so it is possible that participants who were expecting to control their character during the lecture were less attentive to the lecture, as they were waiting for a chance to control the character directly. The results of the linear regression analysis, however, suggested that control was not a mediating variable between the two conditions and the post-test scores ($p = 0.09$). It could be that something related to the feeling of control inhibited players’ ability to learn; however, it is unclear what that is.

A potential explanation for the difference in post-test scores between the two conditions is what motivated participants to engage with the game and the lecture. Most, if not all, of the participants played *Humanatees* because they wanted to learn about manatees. Once participants started playing however, the difference between the conditions may have affected what motivated participants to continue with the game into the lecture. Participants in the experimental condition were motivated to help Rasa, as a relationship had formed between them and Rasa. Those in the control condition were motivated to control their character for the sense of freedom that player-control provides. The lecture only catered to the motivations of participants in the experimental condition, as participants had no control over the lecture other than progressing to the next line of text. This may have caused the participants in the control condition to score lower on the post-test, as their motivations were not being met.
Participants who completed *Humanatees* more quickly reported higher engagement on the Game Engagement Questionnaire. This suggests that participants either played through the game more quickly because they were more engaged and wanted to play more of it quickly, or they were more engaged by the game when they saw the contents faster. Either way, neither time nor engagement were significantly different between conditions, nor was the reported engagement high to begin with (7.34 on average, with 19 being the maximum). Considering many participants likely had experience playing more complex games, it is possible that *Humanatees* was simply not as engaging as other games participants were used to. It is possible that participants were not engaged enough with the game for character attachment to be effective.

Other aspects of the project’s design may have affected the results. Due to the limited scope of the project, retention was not evaluated. It is possible that the two conditions would have differed in how long participants retained the information. Additionally, the surveys adapted for *Humanatees* were meant for players who had spent a significant amount of time with their character or avatar. Participants only had around five to ten minutes to spend with their character in *Humanatees*. Lastly, this project did not study how a long-term relationship between the player and the character (over the course of several play sessions) may have affected learning. Future studies could study the effects of long-term character attachment.

**Conclusion**
Serious games have purposes besides entertainment. They are made up of playful activities, informed by reality, and given meaning by the player. Players ascribe more meaning to games that have made them feel responsible for their character or online avatar. Because of this, the research question asked if serious games that make players feel responsible for their character would be more effective at fulfilling their purpose. The hypothesis was that an educational game that gave players a character they were responsible for would be more effective at teaching them and result in a better learning outcome. To test this, *Humanatees*, a game that taught players about West Indian manatees, was created and posted online for people to play. This game put players into one of two conditions. One condition had players using a blank slate of a character that they were told they control. The second, experimental condition used the protégé effect to make players feel responsible for an autonomous character by having that character ask for help from the player. Players were tested before and after an in-game lecture to measure their learning. They also filled out a survey to evaluate their feelings of identification/friendship, control, responsibility, and emotional investment in their character, their sense of the character’s anthropomorphic autonomy, and their suspension of disbelief and engagement with the game.

The results from the game partially supported the hypothesis. Participants in the experimental condition felt more responsible for their character and scored higher on the post-test, which showed that participants with the character eliciting the protégé effect had improved learning outcomes over those with the blank slate character in the control condition. Responsibility was not correlated with post-test scores however, nor did responsibility act as a mediating variable. Participants in the control condition felt significantly more control over the character than those
in the experimental condition, and the feeling of control was negatively correlated with post-test scores. This suggested that participants who felt more control over their character performed worse on the test. The negative correlation between control and post-test scores was weak however, and control did not act as a mediating variable between the conditions and post-test scores. The experimental condition, where participants had a character who talked to them and asked for their help, scored higher on the post-test than those in the control condition. It is possible that something related to the feeling of control inhibited learning in the control condition, but it is unknown what that is. The difference in player motivation between the conditions is one potential explanation to be explored in future research.

**Future studies**

Future studies can build on the results of this one. For example, a study could alter the current design of *Humanatees* in order to potentially change the outcome. The lecture was delivered through several paragraphs of text. The lecture could instead be gamified. The game could require players to protect manatees from hazards that threaten them, such as boat motors. If the player controlled their character to play these segments, it may also change the effect control or autonomy has on the learning outcome, as the player would have more control than the character. The ability to perform well or to change the outcome of the game would also give players in the control condition more motivation to play and learn.
One thing that should be examined is what aspects related to players controlling their character affect their ability to learn. The study with *Humanatees* suggested that players will not learn more from controlling their character, but from helping them; however, helping the character did not cause the improvement. Will periods of no player input cause a disconnect between the player and the game? Will controlling the character be enough to significantly hinder learning outcomes? This could be tested using various teaching methods, such as lectures or playful teaching activities that players must reflect on after the fact.

Another study could use a longer educational game to teach players. A study that requires participants to play and replay the game over longer periods of time could examine how more time spent with the character influences the learning outcome. A longer game could also test learning retention. More complex characters and gameplay may affect these results as well.

Other player demographics could be explored as well. This study required participants to be of 18 years or older. Because of this, the protégé effect and the factors related to character attachment were not explored with teen or child participants. Younger players might benefit from feeling empathy for their character.

Different lecture subjects might influence the way participants engage and learn with the material. *Humanatees* sought to teach players about manatees, and it was advertised as such, meaning that most participants likely wanted to learn about manatees. It is worth investigating
how character attachment and the protégé effect could influence learning when participants are less interested in the material.

Lastly, serious games with purposes besides education should be examined. Art games such as *The Graveyard* (2008) by Tale of Tales could be studied in place of an educational game like *Humanatees*. In the case of *The Graveyard*, participants would either play the game normally or be given backstory for the old woman they control in the game in order to foster their sense of responsibility for the woman. This would allow the study to aid developers of future art games by providing insight into characters as an element of design.

**Acknowledgements**

I would like to thank my committee members for helping me shape this thesis. Celia Pearce, thank you for the insightful feedback and words of encouragement. Casper Harteveld, thank you for your guidance and for giving me the opportunity to work with *StudyCrafter*. Christoffer Holmgård Pedersen, thank you for being available when I needed your help and advice.

I would also like to thank Sam Snodgrass and the other members of the *StudyCrafter* team for helping me make *Humanatees* the best it could be. I want to thank Alex Liesse, who was diligent in helping me recruit enough participants for this thesis; you went above and beyond, my friend. Lastly, I want to thank my parents, Bill and Kathy Corwin, for always supporting me.
References


The researchers sought to map game mechanics to learning mechanics to create a list of serious game mechanics. The protégé effect is an example of a learning mechanic.


Four different types of player-avatar relationships are suggested. The Avatar-as-Other relationship has been shown to be connected to players feeling responsible for their avatar.


Banks and Bowman analyze player-avatar archetypes through the lense of character attachment. They also mention that players to feel responsible for their avatar are ones who report on meaningful game experiences.


Banks and Bowman propose a metric for measuring avatar relationships. This metric can be used to measure how players felt about their avatar in the StudyCrafter scenario.


Corwin 50
The researchers asked players to report on their most exciting and meaningful game moments. Exciting moments happened while players felt in control while meaningful moments made players feel a sense of care for their game avatar.


The authors proposed a Game Engagement Questionnaire to measure a player’s engagement when playing a game. This questionnaire can be used to measure if players are significantly more engaged with a serious game when they feel responsible for their avatar.


The researchers use the term “the protégé effect” to describe students learning better when they try to teach someone else. The participants in this study will similarly be learning either for themselves or for their avatars.


Harteveld introduces Triadic Game Design - Reality, Meaning, and Play - as a way to think about the design of serious games. Games must be designed well to get across meaning to a player.


A metric for measuring character attachment is proposed here. This CA scale “has a significant relationship with self-esteem, addiction, game enjoyment, and time spent playing games.”

The authors give an overview of serious storytelling. Serious storytelling is storytelling with a purpose other than entertainment.


The researchers evaluated the effectiveness of game-based learning for computer science education. In their study, they used a pretest and a post-test to evaluate the extent of the students’ learning.


The relationship between flow and avatar/character identification is explored here. Flow and character identification are strongly and positively correlated. No relationship between character customization and character identification were found, which contradicts other research here. Perhaps it moreso depends on how similar the character is to the player, and customization is just related to that.

This software is a platform where users can create and participant in social experiments. A scenario for this thesis was be made using StudyCrafter.


The writers give an overview of the definitions and work done using serious games. They explain the different uses serious games have, such as education, and list how different fields use serious games.


The researchers mapped certain game mechanics to pedagogical aspects to better understand what game mechanics can influence learning for serious games. The protégé effect is one example of a game mechanic and is related to responsibility.


In this game, players control an old woman as she moves through a graveyard. If an Art Game alternative to this research thesis is ever done, this game could be used.


The authors analyzed how serious games have had beneficial effects on players. Evidence shows education through serious games yields better learning, but motivation to learn is not significantly greater than that of conventional learning methods.


Examines the gender of the in-game character as a moderator variable when examining violent video games and player aggression.
Appendix

Pretest 1: Which of these is NOT a risk of quasi-extinction for the West Indian Manatee
1. water-crafts such as boats
2. red tide mortality
3. illegal hunting (o)
4. I don't know

Pretest 2: What have West Indian Manatees been found to eat?
1. plants
2. plants and small fish (o)
3. fish eggs
4. I don't know

Pretest 3: West Indian manatee have been the focus of conservation efforts since they were listed under the...
1. Endangered Species Act of 1973 (o)
2. Endangered Species Conservation Act of 1969
3. Endangered Species Preservation Act of 1966
4. I don't know

Pretest 4: Manatees use _ to communicate.
1. echolocation
2. pheromones (o)
3. body language
4. I don't know

Pretest 5: Manatees can live up to 60 years, but the average age of the carcasses of non-calves recovered in Florida was...
1. 14
2. 13
3. 12 (o)
4. I don't know

Pretest 6: In terms of their vision, manatees...
1. They can't see color
2. They see red and green, but struggle with blue.
3. They see blue and green, but struggle with red (o)
4. I don't know

Appendix 1: Each question from the pretest. Each correct answer is marked with a (o).
Post-test 1: The scientific name for the Florida manatee subspecies is...
1. Trichechus manatus linnaeus
2. Trichechus manatus inunguis
3. Trichechus manatus latirostris (o)
4. Trichechus manatus manatus

Post-test 2: How is red tide, a harmful algal bloom, a threat to the West Indian manatee?
1. It makes it difficult to see underwater
2. It's corrosive to their skin.
3. It releases toxins into the air (o)
4. It makes it difficult to move underwater

Post-test 3: Manatee seek warm-water habitats. If the water is lower than _ is unsuitable.
1. 23°C
2. 20°C (o)
3. 18°C
4. 15°C

Post-test 4: How long, on average, can manatees hold their breaths
1. 24 minutes
2. 22 minutes
3. 20 minutes (o)
4. none of the above

Post-test 5: Vibrissae is the term for what part of the manatee?
1. Cuts from boat propellers
2. The flippers
3. The backbone
4. The face/body hairs (o)

Post-test 6: Manatees have been recorded eating the feces of other manatee. We think they do this for what reason?
1. Their low metabolic rate
2. Nutrition
3. To assert dominance
4. To obtain information (o)

Appendix 2: Each question from the post-test. Each correct answer is marked with a (o).

Factor 1: Identification/Friendship
Q01: I sometimes forget my own feelings and take on those of my character.
Q02: I enjoy pretending my character is a real person.
Q03: I consider my character a friend of mine.
Q04: I enjoy pretending I am my character.
Q05: I could see myself being attracted to my character.
Q06: I daydream about my character.

**Factor 2: Suspension of Disbelief**
Q07: I direct my attention to possible errors or contradictions in the video game.*
Q08: It is important for me to check whether inconsistencies are present in the video game.*
Q09: I concentrate on whether there are any inconsistencies within the video game.*
Q10: I think about whether the action or the video game presentation was plausible.*

**Factor 3: Control**
Q11: I enjoy controlling my character.
Q12: I control my character.
Q13: My character does what I want him/her to do.
Q14: I get frustrated when my character does not perform the way I want him/her to.

**Factor 4: Responsibility**
Q15: I know what my character wants.
Q16: I know what my character needs.
Q17: I make decisions with my character’s best interests in mind

**Factor 5: Emotional investment**
Q18: This character is very special to me.
Q19: I would be heartbroken if I lost this character.
Q20: I appreciate this character.
Q21: I love this character.

**Factor 6: Anthropomorphic Autonomy**
Q22: This character has its own thoughts and ideas.
Q23: This character is autonomous and acts on its own.
Q24: This character exists independently from me.

Appendix 3: The items from the Character Relationship survey used in Humanatees. Questions 1 through 17 were adapted from the Character Attachment survey (Lewis et al., 2008). Questions 18 through 24 were adapted from the Player-Avatar Interaction scale (Banks & Bowman, 2015). Items marked with a * are reverse coded. Question 6, “I daydream about my character,” was removed from the survey because the player’s time with their character was so short.

Table 7: The means and standard deviations for the individual survey items. Each table represents each condition, with the first one including both. The sixth item, “Q06,” was omitted from the test as it did not fit with the structure of the project.
Figure 12: The Character Relationship survey item correlations. Deeper blue dots indicate stronger correlations while orange dots indicate negative correlations.