Evaluating an Educational Game for History

Kalpa Anjur, advised by Dr. Erik Harpstead

May 11, 2020

1 Abstract

We investigate the effect of interactivity in historical video games on the player’s ability to learn. Interactivity refers to the player’s ability to click on dialogue choices, move around and interact with objects, and talk to non-player characters in the game. We design a game for historical education centered around the Homestead Steel Strike during the Industrial Revolution while incorporating transformational game design principles [1]. We will then randomly present participants with either an interactive (movement, choices, and puzzles) or a non-interactive (dialogue only) version of this game and measure their learning through the use of qualitative pre-test and post-test surveys. These surveys will then be evaluated using a preconceived rubric.

2 Introduction

As education has become more and more valuable in society, the need for more engaging mediums to convey information to students has risen. As a result, educational video games have emerged as a reasonable alternative to traditional educational materials like textbooks or videos. However, the most effective method of designing such games has been an open research question for years. In order to engage students, developers have tested different genres of games (role playing games,
platformers, etc.), different control mechanisms, and even different visuals and sound effects within games. However, one of the most debated features taken into account when designing educational video games has been the level of agency or interactivity that the player has within the game. *Interactivity* refers to the amount of control that a player has over their in-game character, whether it be moving the character around, choosing who the character speaks to, and even being able to do things within the game like solve puzzles or make choices during dialogue scenes.

In this paper we examine the effect of interactivity in educational video games on the player’s ability to learn by designing and testing a game. Our game covers the subject of history, and it is centered around the Homestead Steel Strike, which occurred during the Industrial Revolution in late 1800’s America. We also incorporate transformational game design principles as presented by Culyba [2]. We plan to randomly present student participants with either an interactive (movement, choices, and puzzles) or a non-interactive (dialogue only) version of this game and measure their learning through the use of qualitative pre-test and post-test surveys.

Effective methods for game design have been analyzed and developed across various studies. Some of the most widely-used principles belong to the transformational framework, developed by Culyba [2]. The framework includes principles that guide transformational game design, aimed at changing the player’s perception to the world via a variety of different focuses. Furthermore, Abeele et al. [1] developed a player experience inventory survey in order to measure the level of engagement and interest a player feels during a game. This serves as another useful tool in evaluating a game’s design.

The importance of interactivity, or agency, has also been investigated in different settings in past studies. For example, a study run by Snow et al. [5] showed that students who succeeded the most in a biology reading-based game were those who had the choice of what they wanted to do within the game. They had more interest in the subject matter and a continued motivation to pursue the
Another interesting consequence of agency was presented by Nguyen et al. [4], who used an educational math game to show that agency did not seem to make a difference in learning due to the presence of indirect control features. Indirect control refers to subtly directing the player’s attention to a specific aspect of the game without explicit instruction, and could prove important in conjunction with interactivity. This concept was further explored by Harpstead et al. [3] using the same math game. The study analyzed how making the player think they were in indirect control while not allowing any further agency did not negatively affect learning. This illustrated that the illusion of agency could still be preserved while allowing for specific educational topics to be conveyed.

This paper in particular focuses on the subject of historical video games because modern education puts an emphasis on science, technology, engineering, and mathematics (STEM). This has led to a lack of interest in essential subjects like social sciences and history. As a result, many students are left with the opportunity to learn new things about history from educational video games that they likely didn’t know before, which is ideal for testing students like those at Carnegie Mellon University, a STEM-focused university. Furthermore, history is an educational subject notorious for lengthy textbooks, slow videos, and numerous events to memorize. Thus, proposing an alternate medium for providing historical education could prove an effective method of overcoming the shortcomings of other mediums, prompting a stronger interest in historical education. Lastly, interactivity in particular is one of the differentiating factors between video games and just videos. Evaluating the different levels of interactivity provides a method of testing how different the educational video game is from traditional educational mediums, beyond just the initial mask of being a ”game” and not a ”textbook”. This in turn provides more clear insight on if an educational video game is effective merely because it is labelled as being a ”game”, or if the design of its internal content is also a significant factor.
3 Iterating on the Historical Game’s Design

In order to create the most effective final version of our historical video game, we iterated on different versions. Initially, we created a rough sketch of a game and controls that incorporated principles of transformational game design. These were then transitioned to a paper prototype using Microsoft PowerPoint, which also contained placeholder graphics such as clip art, along with dialogue and dialogue choices. Lastly, after testing this prototype, the final version was created using the open-source game engine called Unity, version 2019.2.9f1. This included upgraded graphics as well as music and sound effects from royalty-free asset stores. The Unity version also allows the player to move back and forth and interact with object using keyboard keys.

3.1 Initial Design

The initial game design was a design document listing characters, learning goals for players, important scenes that would occur, possible puzzles, and intended mechanics for the game. During this stage of design, after much discussion it was decided that the two main characters would be from rich and poor backgrounds, and that the game would illustrate their lives in parallel in order to more clearly contrast their lifestyles. Furthermore, it was decided that the game would incorporate the principles of transformational game design as described by Culyba [2] in its script, mechanics, and puzzle design.

In particular, we focused on the principles of ”Knowledge”, ”Experience”, ”Belief”, and ”Society”. Knowledge refers to the player directly learning a new piece of information, which is key in our study to measuring how effective the video game is at teaching information to students. ”Experience” refers to the player’s experiences within the game; this was key to shaping their perception of their new knowledge. For example, if the player had to read a lot of dialogue text within the game, it might feel more like a ”book” to them, whereas if a player had to work on puzzles during
the game, they may see the dialogue text as more of a "conversation" they are having before they move on to the next puzzle. This particular characteristic also involves differences in interactivity, which we plan to test. Lastly, "Belief" and "Society" are tied together, and they refer to a player’s perspective on their surroundings and how they see society as a whole. Our game was intended to shape players’ perspectives on American society during the Industrial Revolution time period; although the many new inventions provided convenience for the rich who could afford them, the divide between the rich and poor only increased, making life harder for a majority of the population who was significantly poorer. In order to make sure the players were able to see this perspective, we had to account for the ways in which we conveyed this information, and resulted in the "parallelism" between the rich and poor characters. Players would be able to see what life is like on both sides of the class divide.

3.2 Paper Prototype

After the initial design was fleshed out, a paper prototype (see Figure 1) of the game was designed and implemented using Microsoft PowerPoint. This included placeholder clip art for characters and backgrounds, as well as click-through dialogue and puzzles. There were captions under the art to make it clear to players that this was not the final version. There was no music and there was no movement controls for players.

The prototype was tested right after development by around twenty students in the Carnegie Mellon University community to evaluate the effectiveness of the scripts, puzzles, and story at conveying information. Testers took a verbal survey after playing the prototype to gauge their interest level throughout the game, their understanding of the information presented to them, and their engagement with the characters and puzzles. This testing yielded various pieces of important feedback, including improvement on character speech patterns for historical accuracy, increasing
difficulty of some puzzles, and cutting down on some dialogue. This feedback was recorded for the next iteration of design.

### 3.3 Final Version: Unity Game

The feedback from the paper prototype was taken into account and applied to building a final version of the game using Unity (see Figure 2). The final version included improved art assets, music, and sound effects. As the research team did not contain any artists or musicians, we found assets from royalty-free websites such as opengameart.org that we believed to fit the game well. For example, the music during the "rich" character’s scenes is calm, classical music while the music during the "poor" characters scenes is more gloomy and dark. This adds to the mood of these different scenes while accentuating the differences between classes, as intended by the transformational design principles [2]. The two main characters also have art assets created by the same artist, illustrating their parallelism. Furthermore, the Unity version allows for the player to move left and right using the left and right arrow keys, and to choose which characters and objects
they interact with using the Enter key. However, these controls and the puzzles are disabled for
the second version of the game that contains less interactivity. Instead, the second version only
presents dialogue text that the player must click through, while still including the improved art
and sound assets.

4 Evaluating the Design via Qualitative Testing

After the final version of the game was developed, a testing method was designed in order to
evaluate the different effects of the two different levels of interactivity within the game. The entire
process is to be conducted remotely, involving two surveys and a version of the game itself.

4.1 Motivations Behind Testing Structure

It was decided that testing would be evaluated qualitatively using a pre-test (see Appendix A) and
post-test survey (see Appendix B) in order to gauge the players’ learning of key concepts presented
in the pre-test survey. In particular, a rubric (see Appendix C) was developed to evaluate the
results in each survey and see if the player has any differences in score before and after playing the game. In particular, we decided to score the pre-test results and post-test results separately with the same rubric in order to prevent any sort of bias when scoring results (i.e. purposely looking for simple differences in pre- and post-test results rather than looking for learning improvements). The rubric was developed according to what the game was intended to convey; correct results correspond to the initial learning goals.

Furthermore, the post-test survey includes the exact same questions as in the pre-test survey, but also involves a Player Experience Inventory survey as designed by a separate study [1] to measure how the player felt while playing the actual game. This inventory was included to provide an additional buffer when evaluating the effectiveness of the game’s design in addition to its interactivity. In other words, in case the levels of interactivity of the game yield similar survey results, we can look back at the experience inventory to gauge how engaging the game was; in the case where players did not feel engaged or immersed, we can conclude that the game’s design may have been at fault, which resulted in the levels of interactivity not having a significant difference. However, as we have completed the initial prototype testing of the game’s design, this will ideally not be the case.

4.2 Testing Setup

Participants will be contacted via email with links to a Google Form for the pre-test and post-test survey, and a Google Drive link of a Unity game. First, the participant will take the pre-test survey. This will measure their initial knowledge of the subject matter (e.g. Industrial Revolution, class disparity, etc). Once they have submitted the survey, the player will play a randomly assigned version of the game. The tester will alternate interactive and non-interactive versions of the game between participants to allow for an equal-sized subject pool per condition while also preserving
randomness as to which participant gets what version. The game contains a warning at the begin-
ing to ensure the player understands its contents. Once the participant has played the game
to completion, they will complete the post-test survey. Once completed, the tester will verify that
both survey results have been received and then the test will be over.

5 Surprises and Lessons Learned

As the experiment was being designed, we faced a transition from in-person classes to remote classes.
As a result, the initial experimental design submitted to the Institutional Review Board (IRB) for
approval had to be revised to account for the remote setting, and then resent. Furthermore, due
to the large volume of revision requests that the IRB received, we still have not heard back from
the IRB with an approval. Due to this setback, testing during the semester will be delayed until
approval is granted. This turned the initial plans for research upside-down, as we unfortunately
were not able to get results to present by the end of the semester in this paper. However, we plan
to continue the experiment over the summer and finish execution of the initial plans. In the future,
we will try to account for both remote and in-person testing situations so that we do not have to
resubmit to the IRB later on.

Outside of this situation, one other surprise resulted from the initial testing of the paper proto-
type. Although we initially believed that players would focus more on the story and mechanics of
the game, we found that the dialogue between characters was important to players. In particular,
speech patterns seemed to make a big difference in immersing players in the game. While characters
from the ”upper class” within the game had eloquent and regal speech patterns, those from the
”lower class” in the paper prototype spoke similarly to modern teenagers. This seemed to prompt
players to expect more of a ”class gap” between the upper and lower classes, and they wanted to
see ”lower class” characters speak in an even less educated manner, with more contractions and a
strong dialect. This was an important oversight that was solved in the final design of the Unity game and was interesting to us, while also further proving the importance of initial prototype testing.

6 Conclusions and Future Work

After receiving approval from the IRB, we will begin remote testing of the game on Carnegie Mellon University students. We will then analyze the pre-test and post-test survey results to evaluate the effect of the different levels of interactivity on the participants' learning. We hypothesize that the students who play the interactive version of the game will have the higher score increase from pre-test to post-test as compared to the non-interactive game players. This is largely based on the results from the study by Snow et al. [5], indicating the positive effect of interactivity on learning. Students who can participate directly in the game will feel like they have an effect on the game’s story and be more inclined to read and retain the dialogue exchanges in detail, as opposed to those who only read the dialogue. Furthermore, as the game is dialogue-heavy, there are fewer opportunities for indirect control features in the visual design within the game; as a result, the conclusions presented by Nguyen et al. [4] should not be applicable to our results. Instead, the player experience inventory [1] will provide a clearer indication of the effectiveness of the game’s design in conjunction to the interactivity levels.

This experiment could be further modified and built upon to test effectiveness across different subjects other than just history. For example, interactivity within math, science, and English games could also be tested. Furthermore, one could measure levels of interactivity within the game in even more detail, rather than just two levels (existent or non-existent). For example, the number of puzzles, choices, and characters could be varied in three or more different versions of the game to see if too much interactivity overrides the initial goal of educating the player. One could also
evaluate the player’s learning ability during the game by having in-game quizzes directly related to learning material, rather than having an external survey to measure these results.

In conclusion, as video games have become a key medium for education, there are many different ways to design, develop, and test their effectiveness. What remains important is making sure students are able to obtain and retain key information while staying interested in learning. Once the experiment is conducted, the results will be able to shape our understanding of what the best way is to do so.

7 Acknowledgements

I would like to thank my advisor, Dr. Erik Harpstead, for his guidance during our weekly meetings throughout the semester. I would also like to thank the Carnegie Mellon University Game Creation Society for providing resources and support during game development, as well as providing playtesters for the paper prototype. Lastly, I would like to thank Professor Bogdan Vasilescu and Professor Jonathan Aldrich for their guidance during the research process throughout courses 15-300 and 15-400.
A Appendix A: Pre-Test Survey

What was the economic class system like in America during the late 1800s/early 1900s?

What were some similarities between the richer classes and the poorer classes in this system?

What consequences were presented by developments in technology during the late 1800s/early 1900s with regards to living conditions for the rich and the poor?

Were these consequences widely known by everyone (factory owners, workers, their families, etc)?

How did knowing (or not knowing) about these consequences affect the attitudes of the rich vs the poor?
A Appendix B: Post-Test Survey

What was the economic class system like in America during the late 1800s/early 1900s?

What were some similarities between the richer classes and the poorer classes in this system?

What consequences were presented by developments in technology during the late 1800s/early 1900s with regards to living conditions for the rich and the poor?

Were these consequences widely known by everyone (factory owners, workers, their families, etc)?

How did knowing (or not knowing) about these consequences affect the attitudes of the rich vs the poor?

Player Experience Inventory Please rate how much you agree with each question about your experiences playing the game to the best of your ability. Please use the following scale:

1 – Strongly Disagree
2 – Disagree
3 – Somewhat Disagree
4 – Neutral
5 – Somewhat Agree
6 – Agree
7 – Strongly Agree
Playing the game was meaningful to me.

I felt I was good at playing this game.

I was fully focused on the game.

I felt like I had choices regarding how I wanted to play this game.

I wanted to find out how the game progressed.

It was easy to know how to perform actions in the game.

The challenges in the game were at the right level of difficulty for me.

The game gave clear feedback on my progress towards the goals.

I liked the look and feel of the game.

I grasped the overall goal of the game.
A Appendix C: General Assessment Rubric

0 - blank answer, or shows lack of understanding (or misunderstanding) of question

1 - makes an attempt or educated guess to answer question, but is completely incorrect

2 - some details are correct but overall idea presented by answer is incorrect

3 - general idea presented by answer is correct, with no specific details

4 - correct general idea with some correct specific details

5 - contains very specific details and correct information about the time period. Indicates a deep understanding of the material
References


