Evaluation Plan for an Educational Game

Josie Wittwer

School of Education (CECH), University of Cincinnati

IDT 8130: Mater's Project in IDT

Dr. Gi Woong Choi and Dr. Janet Zydney

October 7, 2020

Abstract

An effective educational game can enhance the learning process, but not any educational game can become a valuable educational tool. A game based on practical learning theories and evaluating its performance needs to be created to assure its value for users. Usability testing is a methodology that can ensure an end product is suitable for learners. *Life: To Save or Spend* is the artifact that I will be testing for this evaluation plan. *Life: To Save or Spend* is an educational game designed for young adults new to personal finance. It is a 'choose your own adventure' style game that depicts real-world spending situations. By playing the game, users can understand what it is like to make certain financial decisions. This evaluation plan seeks to assess the design and technology of *Life: To Save or Spend* by usability testing. The usability testing will be performed by experts and non-experts in the field of educational technology and personal finance. The usability testing results will be analyzed and used to revise the game and turn it into a valuable learning tool.

Keywords: Evaluation plan, educational game, design and technology, usability testing

Evaluation Plan for an Educational Game

The instructional design process is iterative, constant changes and updates are happening to eLearning artifacts. New technology is being rapidly introduced, and we are learning new things about our learners. It is essential to make eLearning tools with this iterative process in mind because it can make the evaluation method more manageable. Developing and refining artifacts will help ensure the eLearning product will be a useful educational tool for all users. During my fall 2019 semester in the Instructional Design and Technology (IDT) master's program at the University of Cincinnati, I created an educational finance game titled, *Life: To Save or Spend*. This game was inspired by my interest in budgeting and making sound financial decisions.

The game is styled as a choose-your-own-adventure decision-making journey. Learners play as a character wanting to save up for an end of year goal. They take their character on a monthly trip making individual decisions that impact their budget. *Life: To Save or Spend* was a collaborative group project submitted as a final prototype for an Educational Game Design course. I will discuss *Life: To Save or Spend* in more detail later on in this paper. This paper will discuss my evaluation plan in revisiting *Life: To Save or Spend* by using the knowledge and skills gained in my graduate studies. For this plan, I will discuss the evaluation methodology that will best suit my artifact. I will also list the evaluation instruments, sampling methods, analysis procedures, data collected, and create a timeline that will be feasible given the time to complete this project. At the end of this process, I will present the data I have collected, enhance my artifact, and display my findings in my digital portfolio located at joswittwer.com.

Background

Educational Game Design

The role of games in learning can contribute to the growth and sustainability of education and teaching (Mercer, Kythreotis, Robinson, Stolte, George, and Haywood, 2017). Educational games can be used to enhance the project-based instructional strategy of learning. *Life: To Save or Spend* was created with this in mind. Finance and budgeting can be a difficult and challenging topic to grasp. Finance is not just about the addition and subtraction of numbers; there is an element of decision-making and learning about how to decipher the best choice when spending money. Delivering learners with an educational tool like a game can enhance the learning process and provide them with a new way of looking at a difficult subject.

I have a background in multimedia design and have created a variety of items using technology. I was excited to learn about educational game design during my fall 2019 semester. In this course, we learned the fundamentals and frameworks of game design. Throughout the semester, we read literature about the theories and best practices of game design and tested and played a variety of different educational games. For our final project, we were tasked with creating an educational game prototype. We were placed in groups to work together and come up with an idea for an educational game. The first iteration of the game was to be designed in Twine. Twine is an open-source software program where users can develop non-linear interactive story-based games.

Creation of *Life: To Save or Spend*

I have always had an interest in personal finance. The idea for *Life: To Save or Spend* came about as a group effort after weighing what sort of themes and concepts would work well in a Twine based game. The game would function as a choose-your-own-adventure series

where characters are tasked with realistic scenarios about which money spending decisions to make. In Twine, we created one character trying to save money for an end of year goal, specifically a vacation. Each month she would run into different scenarios about spending or saving money. To reach her end of year goal, she would need to decide to spend the least amount of money. After completing the initial draft in Twine, we received feedback from our peers and then began the next phase of the design process.

We ran into several problems during this phase, such as which program to use, how to export the game, and divide the work equally. We started creating and developing the game in Unity, but the software's learning curve was not feasible to complete what needed to be completed by the time the final project was due. We then decided to design the game in Adobe Animate. Adobe Animate is an animation program that lets users add multimedia and interactive features to create a video game. We added other elements to the game during this stage, such as graphic design images, music, and sound effects. These elements create a distinct atmosphere that gives life to the gaming experience (Kalmpourtzis, 2018).
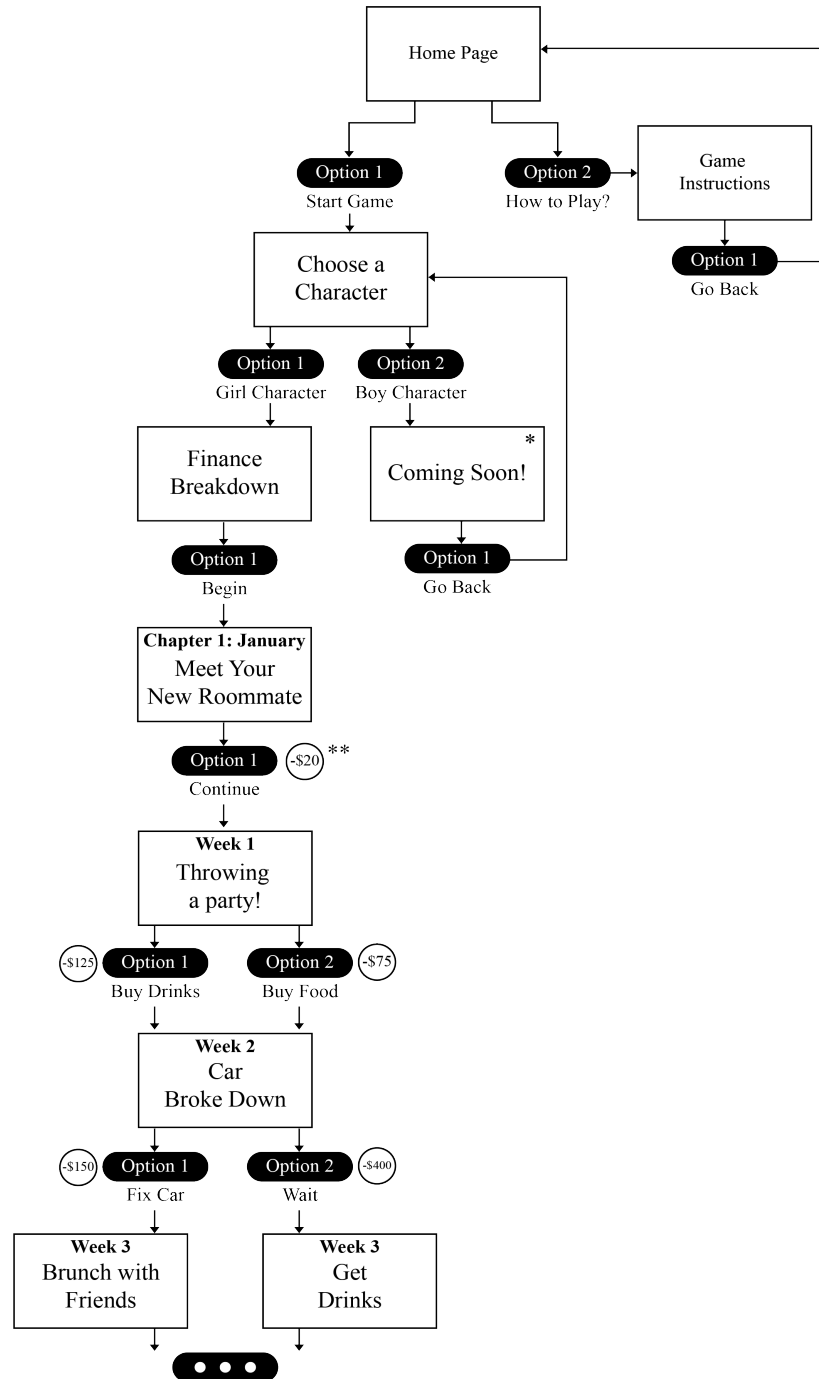
We narrowed the target audience to young adults between the ages of 18-26. The initial audience was high school students, but we felt an older generation would have more experience with money spending and job exposure. We created characters and scenarios to best suit this audience and designed two different characters, but only one character had a completed storyline for the finished prototype.

Once a player has selected a character, they will start navigating their way through various financial decisions. The character will have a job with a monthly income. Before the decision-making scenarios, there will be a finance breakdown to reveal how much can be spent for them that month. Figure 1 shows an image of what this looks like.

Figure 1

Finance Breakdown Screenshot
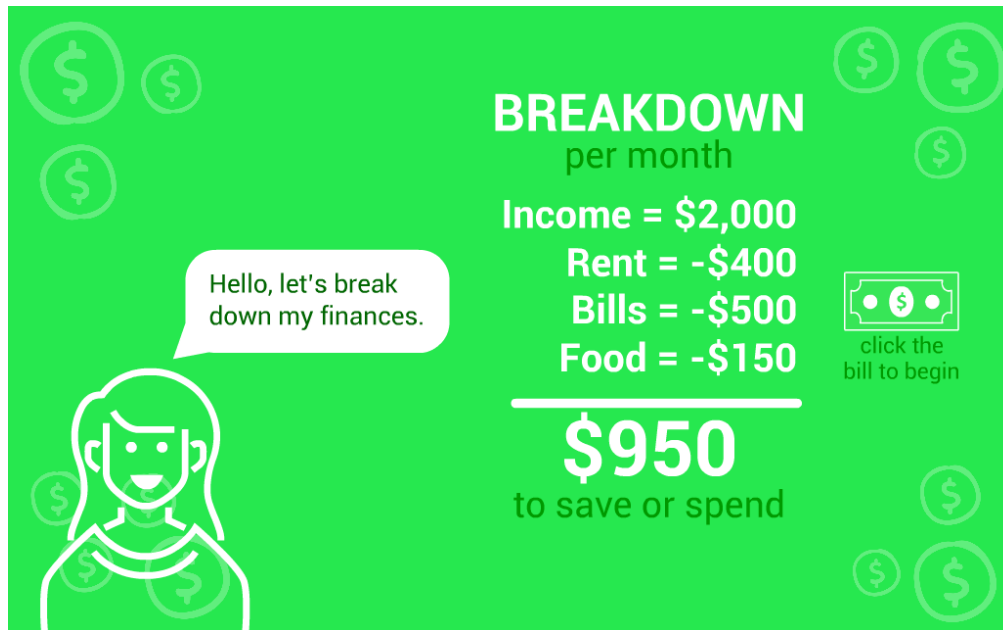
Months are divided into chapters, and each week they will have a decision to make. For instance, in one scenario, the character's roommate is throwing a party and needs help buying food or drinks. The user has to decide if the character should buy food or drinks without knowing the exact cost, similar to how it would be in real life. The reason behind this is because sometimes we may budget to spend a certain amount on something and end up paying more. At the end of the journey, they will have saved a set amount of money given their gameplay decisions. My game prototype only has the month of January completed. Compiling all the information and data for January took the rest of the semester, and that is why it is the only month complete. Figuring out how to Figure 2 displays a flow map of the game and character scenarios.

Figure 2

Flow Map of Life: To Save or Spend

```
                              Home Page
                                 |
              ┌──────────────────┴────────┐
           Option 1                    Option 2        Game
          Start Game                  How to Play?   Instructions
              |                                          |
          Choose a                                   Option 1
          Character                                   Go Back
              |
       ┌──────┴──────┐
    Option 1      Option 2
  Girl Character  Boy Character
       |              |
    Finance        Coming Soon!  *
    Breakdown          |
       |            Option 1
    Option 1        Go Back
     Begin
       |
  Chapter 1: January
    Meet Your
  New Roommate
       |
    Option 1  (-$20) **
    Continue
       |
     Week 1
    Throwing
    a party!
       |
  ┌────┴────┐
(-$125) Option 1   Option 2 (-$75)
   Buy Drinks      Buy Food
       |
     Week 2
      Car
   Broke Down
       |
  ┌────┴────┐
(-$150) Option 1   Option 2 (-$400)
   Fix Car         Wait
       |              |
   Week 3         Week 3
  Brunch with      Get
   Friends        Drinks
       |              |
       └──── ● ● ● ───┘
```

*This character was under construction when the final assignment was submitted.
** These deductions were hidden from players.
*** The game has a total of 4 weeks. This flow chart is used as an example of the scenarios.

(Note: This flow chart was created in Adobe Illustrator)

Since turning in the final prototype, I transferred the game content using Articulate

Storyline. The original software used, Adobe Animate, needs flash support to play. Flash will no

longer be supported in Google Chrome by year-end. Since this will be an artifact displayed in my portfolio, I had to change the software so it will be supported long-term. I want to have the game in a format supported using popular online browsers like Google Chrome, Safari, and Microsoft Edge. These popular online browsers are ending their support of flash as well.

Why an Educational Game?

Educational games can enhance problem-solving skills, "learning is not just about acquiring knowledge but also about being able to properly use it when necessary. Information or skills we acquire and develop become useful when problematic situations appear in our lives that need to be solved" (Kalmpourtzis, 2018, p. 80). According to Kalmpourtzis (2018), an educational game in a learning context can present an efficient way to solve difficult problems. *Life: To Save or Spend* can offer learners with exciting and engaging experiences while learning about personal finance that a traditional classroom setting may not be able to provide on its own. The use of this game in the classroom can help fill the gaps in a traditional classroom.

Evaluation Methodology

Designing the right way to evaluate an eLearning artifact can be challenging. There are several different assessments and methodologies to choose from, but the method should align with the assessed artifact. To begin an evaluation, you must look at what you are evaluating and what you want the end result to look like. "To do evaluation well, you should start by defining what you are trying to evaluate" (Dirksen, 2012, p. 375). For my educational game *Life: To Save or Spend*, I would like to evaluate the game's design and function to ensure that learners can use the game as it is designed. The evaluation methodology I will be using is usability testing through user playtesting and individual questionnaires. Since *Life: To Save or Spend* is still in the creation phase, I want to get insight into how users interact with the artifact early in production to amend the artifact quickly. The testing conditions will be a mix of online remote testing using Zoom and in-person observation testing.

In *Usability Testing Essentials* by Carol Barnum (2020), usability testing is "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (p. 11). Usability testing focuses on the user operating the artifact and observes their experience using the artifact. According to Barnum (2020), usability testing concentrates on the following factors: effective, efficient, engaging, error-tolerant, and easy to learn.

Usability testing aligns with my artifact because it allows me to evaluate the game's ease of use for users. In *Educational Game Design Fundamentals* by George Kalmpourtzis (2018), usability tests are generated to observe their ease of use. They also seek to identify if the game is understandable, error-free, and intuitive. Given that I recently converted the game into a new program, a usability test aligns with my artifact because others have not tested the function and features. A usability test will help me observe my game's behavior amongst potential users. The feedback I will receive will allow me to make sure the game is functional and playable to display in my online portfolio. Since this an educational game, one way to conduct a usability test is to playtest the game.

Playtesting is when a participant plays the game while being observed by the game designer (Kalmpourtzis, 2018). Playtesting is a robust evaluation tool because it can provide real-time feedback and reveal unexpected circumstances. Playtesting is a way to gain further insight into the player's experience with the game. Playtesting will add to the iterative design process. A continuing process of evaluating, playtesting, and revising will make for a powerful comprehensive artifact. Questionnaires are also efficient tools for assessing an artifact because they can provide insightful feedback (Kalmpourtzis, 2018). When developed using a program like Microsoft Forms, they can be sent out and collected more efficiently. The data collected can provide me with quantitative data that can be analyzed quickly using statistical formulas with Microsoft Excel. My playtesting and questionnaire proposal will be discussed in more detail later on in this paper.

Evaluation Instruments

"The purpose of usability testing is to find problems and make recommendations to improve the utility of a product during its design and development" (Lee, 1999, p. 34). I will be

assessing and evaluating usability using five different components of usability. These components were taken from the dimensions of usability testing section from Lee's *Usability Testing for Developing Effective Interactive Multimedia Software: Concepts, Dimensions, and Procedures* (1999) article: (1) learnability; (2) performance effectiveness; (3) flexibility; (4) error tolerance and system integrity; and (5) user satisfaction.

Learnability refers to how the users will absorb and understand the tasks it takes to play the game. Problems related to learnability may include not being able to navigate the game quickly and easily. The navigation needs to be memorable and straightforward; that way, users can quickly transfer the knowledge to succeed in playing the game. Performance effectiveness refers to the ability to implement tasks over time, usually measured in speed and accuracy (Lee, 1999). With high-performance effectiveness, users will be able to play the game more quickly over time.

Flexibility refers to the "variations in task-completion strategies" (Lee, 1999, p. 36). When playing an educational game, it is important that the user does not make many errors while playing. Error tolerance refers to the number of mistakes that appear in-game design while playing. System integrity is the prevention of system and data loss. Suppose the game malfunctions, the integrity of the artifact is lost and unusable. Finally, user satisfaction refers to the enjoyment of the user and their actual feelings of the game. User satisfaction should be high and acceptable so that the users feel engaged and motivated. Elements to measure user satisfaction include color, graphic images, sound, and interaction (Lee, 1999). These five dimensions of usability and their objectives are also listed in the table below.

Table 1

Dimensions of Usability Testing

Dimension	Goals and Objectives
Learnability	To evaluate the degree of user's ability to operate the game to a defined level of competence after time playing.
Performance Effectiveness	To measure the ease of using the system, either by speed of performance or error rate.
Flexibility	To evaluate the degree to which the game enables a user to achieve their goal.
Error Tolerance & System Integrity	To test error tolerance in using the game and system reliability for preventing corruption and input loss.
User Satisfaction	To measure the user's perceptions, feelings, and opinions of the game.

* This table has been adapted from *Usability Testing for Developing Effective Interactive Multimedia Software: Concepts, Dimensions, and Procedures* (Lee, 1999).

With these usability components in mind, I will need to obtain measurable data that aligns with the goals of usability testing. To gather appropriate feedback, I will be conducting several playtesting sessions and collecting a System Usability Scale (SUS) survey.

Playtesting

"Playtesting, or using play to guide game design, gives designers feedback about whether their game is meeting their goals and the player's expectations" (Choi, Forlizzi, Christel, Moeller, Bates, and Hammer, 2016). Playtesting will help me achieve results in the five categories stated above by observing the users' interaction with the game in real-time, providing objective and

subjective feedback that will be quantified and analyzed to make improved recommendations. Playtesting can take on many forms, from informal and qualitative to a more structured and quantitative method. The end goal of playtesting is to gain valuable feedback from users to increase the game's overall experience.

The playtesting sessions I will be conducting will be to obtain qualitative and observational feedback. These will be recorded in both remote and in-person sessions. Qualitative observation methods are useful because they can provide nonverbal input in facial expressions, eye-movement, and interaction cues. These components can attest to the users' learnability, satisfaction, and performance of the game and help decide what recommendations should be implemented to boost the game's appeal and educational value. The analysis procedure of my playtesting sessions will be discussed in detail later on in this paper.

To conduct my analysis procedure, I will create a script to help moderate and introduce the usability playtesting. Moderating usability testing can be challenging. A moderator must be patient and empathetic to build a strong relationship with the tester (Nodder, 2015). It is important to observe how participants behave with my artifact. This will be key to analyzing potential problems. Relationship building can make testers feel more comfortable, thus allowing them to give honest and constructive feedback. The moderator script will allow me to make the usability consistent for each tester. According to Nodder (2015), the critical thing to get across to participants is that the study is not a test of them, and they can stop at any time. When participants are relaxed and at ease, it will be easier to get a better insight into their interaction with the artifact. A detailed moderator script is presented in Appendix A.

After reading the introduction to the game from the moderator script, I will have users begin playtesting the artifact. I will ask questions of my users during and after their playtesting

session. A detailed list of questions will be outlined in Appendix B. These questions have been

adapted from Tracy Fullerton's *Game Design Workshop: A Playcentric Approach to Creating*

Innovative Games (2018). The questions are separated into three sections: (1) in-game

observations, (2) postgame questions, and (3) revision ideas. While participants are playtesting, I

will be observing their habits and take notes of how users are interacting with the game. The

observation notes will be qualitative findings. I will look for consistencies and inconsistencies

amongst the users. For instance, if players are confused or unable to get a particular button to

function correctly, I will know that the artifact, not the user, is not working correctly. I will also

be testing for functionality. According to Fullerton (2018), "if they can play your game from start

to finish without any input or assistance from you, your game is functional. If they can't, figure

out what was missing, and revise the game to make it functional" (p. 314).

System Usability Scale (SUS)

The other instrument in assessing my game's usability will be a System Usability Scale

(SUS). A SUS is a standardized questionnaire used to evaluate users' satisfaction with an artifact

like a game immediately after usability testing. A standardized questionnaire offers advantages

by supporting objectivity, replicability, quantification, economy, communication, and scientific

generalization (Sauro and Lewis, 2016). Given the time and scope of this project, creating a

custom questionnaire survey will not be sufficient. A SUS will quickly identify performance

results in efficiency, effectiveness, and overall ease of use (Sauro and Lewis, 2016). System

Usability Scale surveys should be given to participants directly after evaluation testing.

Participants will provide their immediate responses for each question.

Numerical data is convenient to synthesize because it provides statistical information that

can be easily explained. A SUS consists of a 10 question Likert scale survey. The results of the

survey will be converted into a score. The scores will then reveal whether or not the product is usable. A System Usability Scale will be a valuable instrument in usability testing because "it has been studied by numerous UX researchers and has been proven a valid questionnaire in studies with as few as five participants" (Barnum, 2020, p, 233). The analysis procedure and breakdown of the SUS scoring results and exact questions will be provided in detail later on in this paper. Next, I will discuss other instruments that will be used for my evaluation.

The planning process for usability testing requires determining what instruments need to be acquired before testing (Barnum, 2020). A key consideration in the evaluation and usability of an educational game is the technical equipment necessary to conduct the tests. Since I have changed my game's digital platform, I need to ensure that everything is operational and working before I begin testing. This is also a critical step in the book Game Design Workshop by Fullerton (2018); it is vital to playtest the prototype first to get a version that can be playtested by many.

In-Person Testing

For in-person testing, I will have my laptop for users to test on. My computer has the system and technical requirements for the game to work correctly. I have a 15" MacBook Pro with 16GB RAM; this will help reduce game lag and performance issues. I will also have a backup device, charger, wireless mouse and keyboard, headphones, and speakers. Since the global pandemic is still affecting communities, I will bring sterile wipes and masks to make sure CDC guidelines are being met. For in-person testing, I will bring a script along with questions I will read and ask before and during playtesting. The script and questions will be discussed later on in this paper. A detailed script is listed in Appendix A, while the questions will be detailed in Appendix B. The script will provide users with a detailed overview of the scope and purpose of

my usability testing. I will also use screen capture to record in-person playtesting to reference later on during data collection and observation notes.

Remote Testing

For participants that are unable to meet in-person, I will conduct remote playtesting. There are both benefits and hindrances to remote testing. Remote testing can allow participants to do the testing from the comfort of their own home. It can also open the possibility of recruiting participants from far away areas. However, remote testing requires a heavy reliance on technology. For remote playtesting, users must have their own personal laptop or computer. They will also need a connection to highspeed internet to prevent any lag or disconnect when performing the playtest. A microphone and speakers will be required to communicate between myself and the tester. A video conferencing system will be used to record and share remotely. Zoom will be used to conduct remote testing. Zoom is a free conferencing tool and is available on both Windows and Apple iOS devices. It has the capabilities of recording the playtesting for future feedback use.

Questionnaire Survey

Microsoft Forms will be used for the design and creation of the SUS questionnaire survey. Microsoft Forms is an online survey creation system. It can be easily emailed to participants. The results of the survey can also be easily collected and exported to Microsoft Excel for data analysis. Participants will need access to an email service provider and connection to an internet provider to send and receive surveys.

Sampling Methodology

I plan on recruiting a minimum of five participants and a maximum of seven participants. Of the participants, I will recruit one to two subject matter experts (SME) in finance. The game

is about personal finance, and scenarios deal with the price of products and merchandise. Recruiting and testing with these SMEs will allow me to get accurate feedback on financial items and choices. I have acquaintances that have economic and accounting backgrounds that will participate in conducting the usability playtesting. My SME contacts can do remote playtesting sessions. They are knowledgeable about using technology and are under the age of 30.

Since the game is geared towards young adults, I will recruit users in the age range of 18-30. This specific target audience has a couple of years of experience in the job market but may not attain sound financial decisions as competently as an older generation. Also, the game setting is geared more towards a younger generation. Other factors of potential participants are they have experience using and understanding technology. Some usability testing may be done remotely, so participants need to have their own computing devices that can connect to the remote video conferencing tool Zoom. "The ideal playtester is someone who represents your target audience" (Fullerton, 2018, p. 280). Given this information, recruiting users with the attributes stated above will help me collect relevant and useful data from playtesting. This sampling method is convenient and purposeful. Conducting playtesting with a random sample of participants may hinder the process of fixing issues with my artifact as they may reveal problems not relevant to the game's purpose. In *Game Design Workshop: A Playcentric Approach to Creating Innovative Games* (2018) by Tracy Fullerton, different types of playtesters are suggested for varying categories of prototyping. Figure 3 shows the best suitable playtesters, given the prototyping stage.

Figure 3

Types of Playtesters for each Prototyping Stage (Fullerton, 2018, p. 281)

My game prototype is in the refinement stage so participants will be selected within the target audience. Since it will take time to find the right people, I plan to recruit participants one to two weeks before starting the research phase of my evaluation (see Table 8).

The sample of users should exemplify the target populace in which the system will be represented in (Sauro and Lewis, 2016). By recruiting and sampling participants between the ages of 18 and 30, the assessment results will provide useful information and feedback that will help me create better recommendations to make my artifact a suitable learning tool.

Analysis Procedure

It It is essential to be able to analyze, measure, and synthesize the behavior and attitudes of the testing subjects as they interact with an educational artifact. My evaluation methodology will yield two results. One will be observational notes from playtesting sessions, and the other will be questionnaire results from the SUS survey. These two findings will be processed, interpreted, and used in conjunction to inform me how to make useful recommendations that I will provide in my evaluation report.

First, I will discuss how I will analyze and interpret the results of the System Usability Scale. Since a SUS is a standardized and widely used survey, there are procedures in place that help researchers make sense of the data. As mentioned before, a SUS is a 10 question Likert scale survey (a detailed SUS is presented in Table C1). To score the SUS I will:

First, sum the score contributions from each item. Each item's score contribution will range from 0 to 4. For items 1, 3, 5, 7, and 9, the score contribution is the scale position minus 1. For items 2, 4, 6, 8, and 10, the contribution is 5 minus the scale position. Multiply the sum of the scores by 2.5 to obtain the overall value of SUS. (Barnum, 2020, p. 233)

Table 4 represents a sample SUS result with the score contribution (raw score) and the final score.

Table 4

Sample SUS Results

Participant ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SUS Raw Score	SUS Final Score
1	3	3	2	5	3	2	1	1	5	1	22	55
2	4	1	5	1	5	1	5	1	5	1	39	97.5
3	3	2	2	2	2	1	2	3	5	1	25	62.5

Minimum	55
Maximum	98
Average	72
Median	63

Table 4 shows that three participants took the SUS and yielded a median score of 63. To make

sense of the final score, I will interpret the SUS scores into a grading scale, as presented

in *Quantifying the User Experience* (Sauro and Lewis, 2016). Table 5 shows how this grading

scale will work.

Table 5

Grading Scale Interpretation

Grade	SUS	Range	Adjective	Acceptable
A+	84.1-100	96-100	Best Imaginable	Acceptable
A	80.8-84.0	90-95	Excellent	Acceptable
A-	78.9-80.7	85-89		Acceptable
B+	77.2-78.8	80-84		Acceptable
B	74.1 – 77.1	70 – 79		Acceptable
B-	72.6 – 74.0	65 – 69		Acceptable
C+	71.1 – 72.5	60 – 64	Good	Acceptable
C	65.0 – 71.0	41 – 59		Marginal
C-	62.7 – 64.9	35 – 40		Marginal
D	51.7 – 62.6	15 – 34	OK	Marginal
F	25.1 – 51.6	2– 14	Poor	Not Acceptable
F	0-25	0-1.9	Worst Imaginable	Not Acceptable

Given the grading scale, a SUS score 84.1 to 100 is considered an A+, and a score of 0 to 25

would be an F. If my game scores 80 and above, I will consider it successful, and users perceive

the artifact user-friendly. The results and scale will help me better quantify and understand the results of my playtesting observations.

The results from playtesting may generate lots of data and subjective opinions. To obtain workable solutions from the collected data, I will identify, code, and prioritize usability issues. This procedure will help make sense of the data by allocating impact to determine each issue's severity, allowing me to decide which problems need to be fixed right away. One approach to organizing usability problems and issues is to plot and code the data similar to what is shown in Table 6 below.

Table 6

A Sample Representation of how to Organize and Code Data

ID	Game Play	Task Critically	Where	Description	Impact	P1	P2	P3	P4	P5	P6	Frequency	Severity
1	Selecting the start button	8	Home page	Struggled to start the game.	5	1	1	1	1		1	83%	33.3
2	Viewing buttons	2	Character page	Buttons were inconsistent.	5	1						17%	1.7
3	Income breakdown	5	Breakdown Page	The income breakdown did not make sense.	4	1					1	33%	6.7

*Table adapted from Rosemberg (2017)

The The first step is to collect and generate data by reviewing recorded playtesting sessions and making notes. These notes will be organized and structured using the techniques given by Carlos

Rosemberg in *Turning Usability Testing Data into Action without Going Insane* (2017). The issues will have an identification where the problem will be described along with where it happened and the task that was being conducted. From there, the issue will be coded and prioritized using a linear value from 1-5, with five being severe/critical and one being not harsh or critical. Severe and critical issues will depend on how it affects user experience. For example, if the user can not perform a function of the game due to an error, then that issue will be graded as severe. However, if the issue is a preference of opinion, then it will be rated less critical. Next, the impact score will be defined. The impact score will also be graded linearly from 1-5. The breakdown of the score will be as follows:

Task criticality: Impact on the game or the user if the task is not accomplished.

Issue frequency: Amount of times the issue occurred between participants.

Issue impact: How much it impacted the player trying to accomplish the task.

5 (prevents the user from performing a task)

4 (causes frustration or delay)

3 (major effect on performance)

2 (minor effect on performance)

1 (suggestion from the participant)

Scores adapted from Rosemberg (2017)

Next, the frequency between participants will be tallied; that way, severity can be calculated. The severity will be calculated by multiplying task critically, impact, and frequency. The final severity score will help me prioritize and logically provide solutions to issues. Table 7 will show the severity score range.

Table 7

Severity Score Range

	Severity Score Range
Less Severe	0-5
	0-15
	14-25
Most Severe	25-35

I will conclude my evaluation report by providing recommendations based on these procedures.

Timeline

Table 8 contains a timeline of deliverables and due dates for the completion of this semesters master's project. The semester started the week of Monday, August 24, 2020 and will end on Wednesday, December 2, 2020. This timeline will help keep me on track in making sure everything gets completed on time. Certain due dates and deliverables will be adjusted if needed.

Table 8

Timeline of Deliverables and Due Dates

Task	Deliverable	Due Date
Evaluation Plan	Initial Version	September 13, 2020
Evaluation Plan	Feedback from Peers	September 20, 2020
Evaluation Plan	Revised Version	September 27, 2020
Participant Recruitment	Send Email	September 29, 2020
Playtesting (In-Person)	Observation Notes	October 5, 2020

Task	Deliverable	Due Date
Playtesting (Remote)	Observation Notes	October 10, 2020
Questionnaire Survey (Completed)	Microsoft Forms	October 10, 2020
Data Collection	Survey Results	October 12, 2020
Analysis	No Deliverable	October 13, 2020
Artifact Revision	Update Game Based Analysis Results	October 15, 2020
Evaluation Report	Initial Version	October 18, 2020
Evaluation report	Feedback from Peers	October 25, 2020
Evaluation report	Revision	November 1, 2020
Portfolio Revision 1	Website Content	November 8, 2020
Portfolio Feedback	Website Complete	November 11 , 2020
Presentation & Defense	Peer feedback & revision	November 24 , 2020

Conclusion

By conducting playtesting sessions using a moderator script and in-play questions, pinpointing errors in usability and functionality will be easier to recognize. Upon completion of this evaluation plan, I will begin the testing and evaluation report. This detailed plan will help me collect the necessary data using my chosen methodology. Usability testing by playtesting will help me identify errors in the game design prototyping production process. In order to have an educational game, it needs to be playable for all learners.

References

Assistant Secretary for Public Affairs. (2013, September 6). *System usability scale (SUS).*

Usability.gov. https://www.usability.gov/how-to-and-tools/methods/system-usability-

scale.html

Choi, J. O., Forlizzi, J., Christel, M., Moeller, R., Bates, M., & Hammer, J. (2016).

Playtesting with a purpose. *Proceedings of the 2016 Annual Symposium on Computer-

Human Interaction in Play*. https://doi.org/10.1145/2967934.2968103

Barnum, C. M. (2020). *Usability testing essentials: Ready, set-- test.* Morgan

Kaufmann Publishers.

Dirksen, J. (2012). *Design for how people learn.* [Kindle Version] (pp. 1-424).

Berkeley, CA: New Riders.

Fullerton, T. (2018). *Game design workshop: A Playcentric approach to creating innovative

games* (4th ed.). CRC Press.

Kalmpourtzis, G. (2018). *Educational game design fundamentals: A journey to creating

intrinsically motivating learning experiences* [Kindle Version]. CRC Press.

Lee, S. H. (1999). Usability testing for developing effective interactive multimedia software:

Concepts, dimensions, and procedures. *Educational Technology & Society, 2(*2), 34-49.

Mercer, T. G., Kythreotis, A. P., Robinson, Z. P., Stolte, T., George, S. M., & Haywood, S. K.

(2017). The use of educational game design and play in higher education to influence

sustainable behaviour. *International Journal of Sustainability in Higher Education, 18(*3),

359-384. https://doi.org/10.1108/ijshe-03-2015-0064

Nodder, C. (2015, November 2). UX foundations: Usability testing [Linkedin.com online

course]. https://www.linkedin.com/learning/ux-foundations-usability-

testing/welcome?u=2133849

Rosemberg, C. (2017). *Turning usability testing data into action without going insane*. Toptal

Design Blog. https://www.toptal.com/designers/usability/turning-usability-testing-data-

into-action

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user*

research. Morgan Kaufmann.

Appendix A

Moderator Script

This script This script will be used to introduce playtesting participants to my game. It will be used to build a relationship and make sure the user feels comfortable and informative data is being collected.

Hello, and thank you for participating in this usability test. My name is Josie, and I will be conducting the session with you today.

This playtesting session will take approximately 30 minutes to complete.

Today you are going to be testing an educational game called *Life: To Save or Spend*.

This game was designed as part of a Game Design Course in my graduate program at the University of Cincinnati. For my final capstone project, I am conducting an evaluation plan to determine the game's usability. I want to find out if the game works for users like you; that way, I can display it in my online portfolio.

This isn't a test of you. There are no right or wrong answers. You are helping me test my artifact. Please be honest – you won't hurt my feelings.

Today you will be working with a game prototype. Some of the features are still in development.

If you need to take a break at any point, just let me know. To help me take better notes, I would like you to talk through your experience testing the game. Tell me your thought process through each step. This will help me take better notes and gain more insight into the usability of the game.

For this session, you will play the game as if you were playing any other game.

Let's begin.

[Actual playtesting begins]

[Ask questions from Appendix B: In-game Questions]

[Playtesting ends]

[Ask questions from Appendix B: Post Playtesting Questions]

Next, I would like you to answer a questionnaire survey using Microsoft Forms. This is a System Usability Scale (SUS). This survey will help me obtain quantitative data about the functionality of the game. This questionnaire is anonymous.

[Send link to participate]

Thank you for volunteering your time in helping me complete the testing. Your feedback has been very helpful.

Do you have any questions or comments for me?

Thanks again for participating in this study.

Using this script will help me keep each playtesting session consistent amongst individual participants.

Appendix B

Playtesting Observations, Comments and Questions

The following questions have been adapted from the book *Game Design Workshop: A Playcentric Approach to Creating Innovative Games* by Tracy Fullerton (p. 301). These questions will give me a better insight into the functionality and user experience of the game. I will provide the question in the column on the left, and then explain the reason for asking the question in the column on the right.

Table C1

Playtesting Questions for Users

Playtesting In-Game Questions	The following questions will be asked during the in-game playtesting.
Does the navigation seem complicated?	This question can be asked to understand whether or not the user can navigate properly.
Why did you make that decision?	This is important to gain insight into a user's decision-making process.
Why are you pausing there?	If there is a long pause, I will know that there is an error with the game's flow.
Post Playtesting Questions	The following questions will be asked post playtesting session.
General Questions	
What was your first impression?	First impressions can determine whether or not users will enjoy the game or not.

How did that impression change as you played?	This will help me understand if the game got easier or harder throughout the playtesting.
Was there anything you found frustrating?	Frustrating aspects can reveal inconsistencies with usability.
Did the game glitch at any point?	Glitching may mean the game exported improperly, or it is not functioning properly.
Were there particular aspects that you found satisfying?	These questions will help me understand which aspects are better developed than others.
What was the most exciting thing about the game?	
Did the game feel too long, too short, or just about right?	
Formal Elements	
Describe the objective of the game?	This will help me determine if the theme and objectives were consistent throughout.
Was the objective clear at all times?	
What types of choices did you make during the game?	Did the participant decide to do something because that was the only option?
What was the most important decision you made?	This will give me better insight into their specific gaming preference.
How would you describe the conflict?	
Do you prefer to play alone or with human opponents?	

What elements do you think could be improved?	Participants will shed light on what elements need to be improved.
Dramatic elements	
Was the game's premise exciting?	Dramatic element questions will help me gain further insight into the theme and storyline of the game.
Did the story enhance or detract from the game?	
As you played, did the story evolve with the game?	
Is this game appropriate for the target audience?	
Did you feel a sense of dramatic climax as the game progressed?	
How would you make the story and game work better as a whole?	
Procedures, interface, and controls	
Were the directions easy to understand?	The interface and controls are what makes the game playable and user friendly.
How did the controls feel? Did they make sense?	
Could you find the information you needed on the home screen?	
Was there anything about the interface you would change?	
Did anything feel clunky or awkward?	
Are there any controls or interface features you would like to see added?	
End of session	

Overall, how would you describe this game's appeal?	
What elements of the game attracted you?	End of session questions will help reveal
What was missing from the game?	any other issues that may not have been
If you could change just one thing, what would it be?	asked before. These questions will also
Who do you think is the target audience for this game?	help determine any inconsistency or
Does this game enhance the learning process?	problems with the playability aspect.

Appendix C

System Usability Survey

      The following survey will be given to participants upon their usability playtesting session. The SUS will be uploaded to a Microsoft Form and sent via email for each participant to take and submit.

Table C1

Survey Questionnaire

Rate each question between 1-5 with 1 being "strongly disagree" and 5 being "strongly agree."					
I think that I would like to use this game frequently.	1	2	3	4	5
I found the game unnecessarily complex.					
I think that I would need the support of a technical person to be able to use this game.					
I found the various functions in this game were well integrated.					
I thought there was too much inconsistency in this game.					
I would imagine that most people would learn to use this game very quickly.					
I found the game very cumbersome to use.					
I felt very confident using the game.					
I needed to learn a lot of things before I could get going with this game.					