

Experience, Experiment, Evaluate: A Framework for Assessing Experiential Games

Nicholas Lytle¹, Mark Floryan², David Amin²

¹North Carolina State University, Department of Computer Science
nalytle@ncsu.edu

²University of Virginia, Department of Computer Science
{mfloryan, dma3fq}@virginia.edu

Abstract

*The design of effective educational games has proven to be challenging for many years, leading to sparse and somewhat inconsistent insight into the principles governing such systems. While attempts at constructing frameworks for educational games certainly exist, their nature is often quite general (limiting the practical utility) or noticeably specific (limiting the scope of projects to which that framework might be applied). We present a design framework for a broad, but well-defined genre known as experiential games. We have named our framework the Experience, Experiment, Evaluate (EEE) framework and believe it to be an adequate lens to analyze such games. This article presents the EEE framework in detail and provides example analyses of three games (a U.S. Civil War history game, a medical diagnosis game, and *The Legend of Zelda: Ocarina of Time™*). We present results from one game that predicted pedagogical success through learning gains by adhering to the framework, and are able to use the framework to explain an increased effectiveness in a design change of the other game. In particular, the medical diagnosis game, *Rashi*, is shown to elicit higher quantity and quality of student responses when features were added that more tightly bound the game to our framework. Additionally, we provide evidence that activities within our U.S. Civil War game, 'A Nation Divided', are more successful in providing learning gains to students when those activities more carefully apply the ideas within our framework. We do not present any empirical results regarding *Ocarina of Time™*, but include it as an anecdotal example of how commercial games have applied these principles successfully to teach mechanics of games to players, and argue that this is, in many ways, an exercise in pedagogy. We conclude by offering suggestions for strategically incorporating elements of our framework in the development and design of future systems and express the applicability of our framework to both experiential and non-experiential games.*

Keywords: *Experiential Games, Serious Games, Game Frameworks;*

1. Introduction to Problem Area

There have been multiple attempts at developing a standard framework for the design and evaluation of educational games, but none has become a clear favorite. This is, in part, because of the difficulty of developing a framework that applies broadly enough to address the elements an educational game should have without becoming so broad as to be impractical for making real comparisons. For example, Amory proposes an extension to the Game Object Model [1] that considers the components



of an educational game as a series of interfaces and objects similar to the paradigms of OO programming. This framework has 6 core categories for the 57 unique concepts that are spread across 12 spaces. The core categories include the importance of designing engaging and accurate narratives, the importance of social collaboration, and the role of gender in gaming. Additionally, the framework considers the importance of well-defined challenges in games, both complex and relevant to the player so as not to seem trivial, the importance of progressively more difficult challenges, and experiences that are consistent with their real world source material. Our proposed framework is in many ways similar to this system, but some aspects of this model were considered extraneous for our purposes. Specifically, the focus on social collaboration as an educational tool was less of a concern for our case studies, which focused more greatly on the player exploring the games. The experimentation process we define is much more individual in scope, allowing learners to independently discover the mechanics of the experience, collaboration is not relevant to this particular approach. As such, EEE seeks to focus on core factors relevant to experiential games, such as the importance of building first hand experiences and the factors that make for an effective puzzle or challenge.

Additionally, Winn developed a framework [2] analogous to the popular MDA framework [3] which naturally makes it easy to apply to assessment of educational games, as the familiar core aesthetics remain and are extended with educationally relevant concepts. These broadly fall into three new categories: learning, storytelling, and user interaction with the game. This framework, as MDA before it, helps with bridging the gap between a designer's perspective of the game and a player's. For example, the designer develops the content with pedagogy in mind, and must constantly consider how this content will translate into learning for the player. While it is valuable to consider these factors and others, such as the importance of a user interface on user engagement, some aspects of the framework were problematic for evaluating the games we've examined, specifically the focus on narrative. We argue that while this type of framework is useful for discussing and breaking down games, they aren't as helpful at explaining the success of particularly well designed games. Other frameworks [4][5][6][7], while very useful as general strategies for design, don't recommend explicit design principles, rendering their application somewhat difficult. Akkerman et al. evaluate an educational historical game [8] in a way that could potentially be applied to some other experiential games, but would be tricky to generalize to experiential games overall. As an alternative to systems such as these, our framework aims to be broad enough to apply to any experiential game yet not so general as to lose specific meaning, allowing for an in-depth discussion of the important elements of the game and making comparisons between such games easier in the future.

Activity Theory-based Model of Serious Games (ATMSG) [9] was developed to create a means of identifying relationships and categorizing elements of the instructional, learning, and gameplay layers of serious games. This allows for a convenient means of organizing and understanding the serious game as a hierarchical system allowing analysis of not only these relations within the game, but within the environment that the game is used. ATMSG is useful for identifying game components and relations, and provides a useful generalized process for understanding a serious game. Our framework enhances this framework by 1) narrowing the scope of use to a specific, but still broad class of games (experiential games), and 2) providing specific design principles for optimizing the utility of an experiential game. In this sense, our framework attempts to provide a means of optimally organizing, presenting, and mapping game components in order to achieve the best learning outcomes possible. We thus view our framework as providing useful insights into the design of serious games, while still remaining congruent with ATMSG.

Recent developments include the creation of a model-based framework for designing educational games [10]. This framework has also been implemented into a web application that allows designers to use this as a more formal means of laying out their game. This framework affords developers of these systems the ability to more broadly see the relationships between the learning related elements and the game-specific structure. While not incompatible with our framework, our aim is to present a less-general framework that focuses more on the design elements one should consider in that specific subset of educational games that are experiential.

For the purposes of our research, we will define experiential games in a manner similar to Nicholson [11], as games that require learning by action, especially simulating real world events in some capacity. Our framework seeks to identify the core factors essential to this specific subset of educational games, by exploring the factors crucial to the development of two experiential educational games, and the mechanics of a game for entertainment that teaches core game concepts effectively by creating specific experiences for the player. In identifying these essential factors, we

seek not only to provide a system for comparison of experiential games, but also to recommend best practices for the development of experiential games in the future.

2. The Framework

An experiential game can be looked at from a design perspective as having to meet three distinct criteria. First, it must provide mechanics of interaction that the player can **experience** within the game space. Second, the game must construct scenarios that allow players to **experiment** with these mechanics in order to learn how to use these mechanics as well as how certain mechanics relate to others within the game world. Finally, in order to assess that the player has adequately learned a skill, the player must be presented with a scenario that allows them to **evaluate** their understanding of a given mechanic. This also serves the purpose, in the case of educational experiential games, for pedagogically determining if the student has learned the appropriate information or skill. This three-part lens can be applied on a macro level to assess the effectiveness of the system as a learning machine overall, or on a micro level through assessing individual mechanics or topics that are introduced in the game. These three categories were chosen through a data-driven analysis of “A Nation Divided” (Sections 3.1.1-3.1.3) analyzing successful and unsuccessful components of the initial tests (Modules 1-3, see below for more detail) to identify metrics for predictive success. These metrics were then decoupled from the specifics of the game to try to provide design decisions general for any experiential game.

Overall, the EEE framework provides a pedagogical lens through which game mechanics can be evaluated. The following sections provide more detail on each component of the framework.

2.1 Experience

Perhaps the most necessary aspect of experiential games is the experience that is being created. The means of interacting with the world are chosen and each maps explicitly to a learning goal or concept. In the case of educational games, the experience must be intentionally designed and explicitly understood in order to produce an effective as well as engaging experiential game.

First Hand/Interactive Mechanic - In order for a game to meet the definition of an experiential game, there must be a direct means for the player to interact with the world as an agent. In an educational game, these mechanics also must map to an appropriate Learning Mechanic in order for effective learning to occur [12]. We posit that the more direct the interaction, the better. For example, talking to a historical figure directly (e.g., Abraham Lincoln) is more useful and engaging than hearing about this character through a generic NPC.

Taxonomizing Information - An understanding of how each lesson, idea, or topic relates to one another is necessary for the creation of an appropriately designed system. Topics with clearly recognizable dependencies should be presented in an appropriate order to ensure users have sufficient background knowledge to tackle new concepts. More importantly, these definitions and relationships must be communicated to the user naturally through the interaction.

Encapsulating - Well encapsulated games create clear distinctions (either systematically through their mechanics or structurally through their presentation of concepts) as to what is and is not part of the topic. In doing so, it presents a clear organization to the player as well as reduces cognitive load. This is especially important in serious games, as the border between real content and superfluous information provided by the designers can often become blurred.

Experience <i>Craft the user experience by developing the educational content & mechanics</i>		
First Hand/Interactive Mechanic	Taxonomize Information	Encapsulating
The game must involve direct interaction with the environment and educational content in a meaningful way. Designers should craft experiences that allow users to interact as closely with the content as possible.	Information within the experience should be well organized by the developer and communicated to the player in a digestible manner.	Developers should make clear to players what topics are within the scope of the game and what topics are not. Players need to be aware of the boundaries of the interaction.

Experiment <i>Give players the means to experiment themselves and discover the knowledge</i>		
Relating Ideas	Clustering	Labeling
The game should be designed to enforce relationships between ideas in the domain. This can be done visually, through narrative, etc. Games offer a unique medium through which relationships between concepts can be well established.	Topics within the game that are related should be spatially or temporally clustered near one another. This reinforces relationships between topics in a meaningful way.	Items that belong to a particular group should be labeled as belonging in some manner. This allows players to understand the disparate objects within the game and discern their respective properties themselves.

Evaluate <i>Players should be provided the means to evaluate their progress easily and accurately</i>		
Clearly-Defined Goal	Motivating Goal	Difficult to Game the System
The player's goal should be clearly defined and communicated up front. In addition, every activity required should be well-crafted and relate meaningfully to this goal.	When possible, the goal should be as motivating as possible. This should be done, however, in an authentic manner. It is generally not useful to add "hip", but out-of-place mechanics to liven up the goal.	It should be difficult (nearly impossible) to get through the game without approaching the given tasks in the intended manner.

Figure 1. Summary of the EEE framework and its components.

2.2 Experiment

The second category of our framework is experimentation. Under the assumption that well-mapped and proper mechanics are chosen, it is necessary to properly introduce and teach the encompassing concepts in a natural and appropriate manner. This should include appropriate design of scenarios to properly test and teach players how these concepts are understood in isolation as well as how a particular concept relates with others already previously developed.

Relating Ideas - While engaging with a certain mechanic or concept, players must be presented with means of relating the ideas currently being tested and introduced with ones that are assumed to be well-formed and previously established. In doing so, relationships between these ideas can further develop the understanding of the idea currently tested, re-solidify the concepts taught previously, and provide a better universal understanding of the topic at large. Games provide an excellent medium for doing this both visually and through specific tasks and challenges.

Clustering - Either spatially within the game world or temporally through the sequence of play, ideas should be introduced in a way that allows for a natural grouping of the mechanics or topics. If

a player is supposed to learn how to use the shotgun, spatially placing the player in a close-quarters environment gives them room to explore the mechanic. If a player is supposed to be learning how to use a newly introduced flashlight, immediately cutting out the lights temporally links the mechanic with its use case. Likewise, if a student is meant to be learning about a historical battle, the concepts related to that battle should be present, and visually organized to lessen cognitive load.

Labeling - Signaling visually, or through some other means, what objects belong to the current topic is necessary for effective learning. Having items in the world clearly demonstrate their relation and interaction potential with the current mechanic allows players to better sense how these items fit together prior to any action being taken. For example, related items can share a unique color, or challenges that rely on a given mechanic can be consistently labeled with some word or symbol.

2.3 Evaluate

Throughout the entirety of the experience, a player must be presented with a means of assessing whether or not they have a correct understanding of whatever concept they are currently engaged in. In an educational game, it is also necessary to create effective tests of a user's ability in order to demonstrate learning gains. Thus, proper, goal-oriented means of evaluating the player's abilities with a mechanic or concept are necessary in experiential games.

Clearly-Defined Goal - In many of these systems, a clearly-defined goal is presented to the player. Whether this be a final boss at the end of a level or the end of a set of levels in one world, the game presents the player a clear marker past which they can reasonably assume that they have acquired the knowledge about the related topic or mechanic. By ensuring the player constantly knows what they are trying to achieve, frustration can be reduced, allowing for more of a focus on learning.

Motivating Goal - In addition to being clearly defined, the end goal for the user must be motivating enough for a player to attempt to try and achieve it. This can either be done thematically through the narrative created in the game or structurally through the mechanics of the game itself (e.g. a major reward is given to players who complete the goal). Well-designed goals will be motivating in multiple ways, to ensure players with a variety of means of motivation will be engaged.

Difficult to Game the System - Gaming the system is a well-studied area [17][18][19] aimed at preventing students in educational contexts to cheat the intended nature of the activity. In order to provide a proper metric for evaluating a player's ability in a certain topic or skill, the means of assessment should be designed in such a way to ensure that only those who acquired the necessary related skills are able to complete it. If a final boss of a dungeon is beatable through a means other than whatever mechanic or item the dungeon was testing you on, the final boss is not a good evaluation of the skill.

3. Analyzing 3 Games Under Our Framework

3.1. Game 1: A Nation Divided

"A Nation Divided" is a web-based game developed in Javascript using a rendering engine called PIXIJS [20]. The game is played from a top-down perspective as the player takes control of a nameless protagonist (see fig. 2, left) navigating the game world. The player interacts with other non-playable characters and entities and in many cases, the player is able to choose what response to give to these characters from a list of options.

The game currently contains four distinct modules (more detail below) focusing on distinct aspects of the American Civil War. In each module, the player is placed in a specific location and must complete a variety of tasks. This includes activities such as reading newspapers, communicating with townspeople, and engaging historical figures such as Harriet Tubman, Abraham Lincoln, etc. The quests for each module can be completed in any order, but all must be completed to progress to the ending segment. For many quests, the player must supply a non-playable character with the correct response to a question (or series of questions) to make progress. This helps ensure that the player understands a piece of information within the domain before



continuing. If the player is unable to supply the correct answer to that character, there are characters within the vicinity that give hints or help.



Figure 2. While playing, students walk around town interacting with historical figures and locals (left). Every module contains a unique end game scenario quizzing the student's knowledge (right).

After all of the tasks are completed, the final section of a given module is available to the students. These ending scenarios are unique to each module and have their own interaction scheme and win condition. For example, Module 3 (Mid-war) features an end-game segment that tasks players to control a Union blockade and shoot down ships that have Confederate identifiers (fig. 2, right). These ending segments usually tie back to the theme or plot that was developed in the previous section.

The game employs a cartoony aesthetic with colorful graphics and light, calming music. While there are certainly aspects of the time period that could have been explored or shown in a more detailed, mature light, the game was specifically designed for an elementary classroom setting and the design was catered around this intended young user. The dialogue, the primary vehicle for delivering information in the game, was not developed to be completely in-line with the speaking mannerisms of individuals in the 1800s. Instead, the primary goal was to make comprehensible and engaging dialogue that an average 5th grader could read and understand.

As stated earlier, the game contains four distinct modules, details of which follow:

1. *Underground Railroad*: Taking place on a southern cotton plantation in the 1850s, the player must aid Harriet Tubman in collecting supplies, information, and travelers to take on the Underground Railroad. Once ready, the player must complete a turn-based board game moving a group of runaway slaves from city to city navigating away from slavers trying to catch the party. The game is completed when either the slaves are caught (resulting in a game over) or the slaves are able to escape out of the country.

2. *Secession*: Set in an unspecified Virginian plantation in the year 1860, the player is tasked with delivering newspapers containing the news that Abraham Lincoln has been elected president. The player must present the correct newspaper to the non-playable characters (either a northern abolitionist newspaper or a southern newspaper) based off of the information and views that they relate to the player. On completion, the game shows each southern state seceding from the Union in order coupled with information about the beginning of the Civil War.

3. *Mid-war*: Set in 1862 on an unspecified Virginia plantation, the player must help General McClellan with a series of tasks for the Union. This includes relaying messages to Union naval officers, helping wounded soldiers remember which battles they took place in, and aiding Clara Barton and her nursing staff. Once completed, the player takes control of a Union blockade shooting down ships that have concepts associated with the Confederacy (e.g. "Virginia", "General Lee", "Stars and Bars", etc.) and allowing ships that have concepts associated with the Union to pass.

4. *Gettysburg*: Set in Gettysburg, Pennsylvania during the Consecration of the National Cemetery at Gettysburg (1863), the player is tasked by Abraham Lincoln to collect information about what happened during the battle and how people are reacting. The player must piece together a timeline of events from General Meade, ask for Frederick Douglass's opinion on the battle and war, and help reporters get the facts of the battle. Afterwards, the player aides Lincoln in constructing the Gettysburg Address by select the correct missing word in a segment of the address given its synonym.

Although similar to one another, each module is unique in exactly how material is presented and game mechanics are designed.

A preliminary version of the EEE framework developed in this paper was used as a development tool for the creation of the final module and the analysis of ‘A Nation Divided’. As such, many of the elements of our current framework are prominently demonstrated in the game’s 4 modules (especially the 1st and 4th). Below are examples of a few of our EEE framework elements demonstrated by this game.

	FRAMEWORK ELEMENTS								
	Experience			Experiment			Evaluate		
	Firsthand / Interactive Mechanic	Encapsulating	Taxonomizing Information	Relating Ideas	Clustering	Labeling	Clearly Defined Goal	Motivating Goal	Difficult to Game the System
Module 1	Yellow	Green	Yellow	Yellow	Green	Green	Green	Green	Yellow
Module 2	Yellow	Yellow	Red	Red	Yellow	Red	Yellow	Red	Red
Module 3	Green	Red	Yellow	Red	Yellow	Green	Red	Green	Red
Module 4	Green	Green	Yellow	Yellow	Green	Green	Green	Green	Green

Figure 3: Adherence of each module of ‘A Nation Divided’ to the EEE Framework. Modules 1-3 were used as preliminary data driving the development of the framework. The framework was then applied directly to the design of Module 4 for testing and validation.

First Hand/Interactive Mechanics: The mechanics and narrative always directly task players with an objective relevant to the experience a Civil War era civilian might have had, such as finding relevant information about the events of the war from newspapers. The direct interaction with historical figures in dialogue was another easy way to design first hand interaction for the player.

Taxonomizing Information: We set up our quest system to have a chain of dependencies, such that players would not be presented with information until they had achieved steps that presented them with important details for context. For example, talking to a couple characters who described details of the war prompted players to go talk to General McClellan, where they were given further details, ensuring they had the appropriate context before more information was provided.

Encapsulating: The first and final modules we found to be clear examples of enclosing a specific topic of ideas (slavery in the first and Gettysburg in the second) and creating a focused and targeted module based on it. These two, as hypothesized, were the most successful in terms of effective learning gains.

Relating Ideas: In several of our modules, we placed content that our players would have seen in previous courses as a means of extending familiar content. For example, historical figures such as Lincoln or Harriet Tubman were designed to be recognizable, while their in-game purpose was to extend the existing knowledge about their lives.

Clustering: As a whole, we found that our modules did a good job in the most successful cases clustering related items together spatially in the game world. In the final module, all of the abolitionist characters and quests were in one area of the map while all of the soldiers were in another thereby not only separating game-related items, but also concepts and ideas.

Labeling: For our game, many historical elements assisted in implicitly labeling our content. For example, Union soldiers all appear in blue uniforms, and Confederates appear in gray. This helps players understand aspects of the content immediately, in this case which side of the war the characters were on, and can be an incredibly useful tool.

Clearly Defined Goal: Every module we developed had an end game phase, but we found that the modules for the Underground Railroad and Gettysburg were better about actually informing players of their objective. The modules that were less clear often led to players being confused about what they should be doing next, increasing aimless wandering and reducing the time spent going through the actual educational content, as well as adding needless frustration to the gameplay experience.

Motivating Goal: While most of the end game phases for the modules were well received, the secession module was not. The phase lacked any real game elements, and was closer to a traditional quiz than any other evaluation content we designed. This, coupled with the fact that the segment

was quite long and gave poor feedback on incorrect answers, led some players to stop trying to finish it, indicating that goals must be not only well understood, but motivating.

Difficulty Gaming the System: In earlier modules, it was very possible to complete quests without truly understanding the content that was mapped specifically to those quests just by scrolling through text and picking responses. In our final module, we purposefully designed conversations that made it incredibly difficult to get the final result (completing the quest) without knowing the knowledge that the assessment revolved around.

3.1.1 A Nation Divided Methodology

For each module, students from a small elementary school in Virginia were taken out of the classroom for an hour long period of time. Each student was a 5th grader who was learning the material discussed in the game that year in their classroom. Each student was given a school laptop computer to access the game as well as pencils and pens to answer the pre and post test questions. Some students participated in multiple test sessions while others did not. The number of students that participated in each test session varied from module to module. This information is provided in the Results section under table 1.

The test sessions were divided into three main sections: the pre-test, the game playing session, and the post test. Between each of these sections the students were given instructions on what he or she would be doing by the researchers and the teacher. During each section, the teachers and researchers would provide clarification to any student asking questions or provide help to the student when asked if it did not directly influence their ability to complete the task (e.g. revealing the location of an item within the game). Each student was given a unique identifying number that was held constant between all three sections. The teacher knew which student was tied to each number, but the researchers did not.

During the pre and post-test sessions, each student was given a test that included a number of multiple choice questions related to the material discussed in the game. These questions were all modeled after Virginia Standard of Learning (SOL) style questions for the 5th grade Virginia Studies SOL and some were direct questions from previous SOL exams. The students were able to answer the questions in any order they chose and were given 15 minutes to complete the assessment. In some sessions, some questions appeared on both the pre and posttest for that session and some questions appeared on multiple test sessions.

Between the pre and posttest sessions, the students had up to 30 minutes to play the game. Students were allowed to ask the teachers and researchers clarifying questions but were not given information on exactly where to go or how to complete a task. Any technical difficulties that happened during this period were handled by the researchers. The students were not strictly discouraged from collaborating with other students playing the game unless a student was explicitly telling other students how to complete certain tasks. If a student completed the entirety of the game, he or she was encouraged to continue exploring within the game until the time limit for gameplay was up.

3.1.2 A Nation Divided Results

Table 1 below summarizes the pre-test to post-test score changes over all four modules. The pre-tests means (X) and standard deviations (S) are labeled with the subscript 1 and post-tests values with subscript 2. The P-Values (P) were calculated through T-Tests between pre and post test scores. Three modules led to improvements in scores (modules 1, 3, and 4). Modules 1 and 4 were the only two that show statistically significant results ($p < 0.05$). Both of these modules also contain medium to high effect sizes.

Table 1. Results comparing pre-test to post-test scores across all four test sessions (one test per module)

	<i>n</i>	<i>mean</i> (<i>control</i>)	<i>std. dev.</i> (<i>control</i>)	<i>mean</i> (<i>exp.</i>)	<i>std. dev.</i> (<i>exp.</i>)	<i>p</i>	<i>Cohen's</i> <i>D</i>
<i>Underground Railroad</i>	17	0.614	0.176	0.739	0.147	*0.012	0.766
<i>Secession</i>	20	0.611	0.224	0.572	0.208	0.384	0.180
<i>Mid-war</i>	15	0.550	0.210	0.633	0.208	0.191	0.398
<i>Gettysburg</i>	17	0.346	0.225	0.438	0.234	*0.008	0.398

We also performed an analysis of learning gains across various question types, in order to obtain some sense of student learning across categories. Table 2 below summarizes these results. Modules 1 and 4 performed well on questions regarding people (important players during various times of the war). Additionally, module 1 displayed improvements on analysis questions. Two modules (3 and 4) show significant improvements in questions regarding events. There were no significant improvements on questions regarding viewpoints or facts & terms.

Table 2. P-Values of T-Tests from pre to post-tests broken down by question type. Red highlights indicate that the mean decreased. Bold indicates statistically significant changes

Module	N	People	Geography	Viewpoints	Facts/Terms	Events	Analysis
Underground Railroad	17	0.007**	N/A	0.999	0.168	N/A	0.013*
Secession	20	0.090	0.517	0.209	N/A	0.825	N/A
Mid-war	15	0.178	N/A	0.334	N/A	0.006**	N/A
Gettysburg	17	.003**	0.668	0.104	0.999	0.049*	0.496

3.2. Game 2: Rashi: A Medical Diagnosis Game

Rashi is a medical diagnosis game designed for use within high school and college classrooms [13][14]. The system presents users with one of many cases involving a virtual patient who begins by describing his or her symptoms. The goal of the game is to explore various hypotheses and data collection methods (simulated to be consistent with the work of real professionals) to produce a diagnosis for the patient. The system, in its original form, contains the following core features:

- A student notebook provides users with a space for summarizing competing hypotheses and collected data. The notebook also allows users to establish relationships between collected data and hypotheses (e.g., heart palpitations support the theory that the patient is suffering from Grave's Disease). The students can also mark hypothesis as being the top theory, pending, ruled-out, etc.
- Various data collection tools are provided to users. These include a widget for asking the patient (or related stakeholders) questions, performing pertinent lab tests, or performing a physical examination.

The system provides functionality for organizing and accessing outside data sources, including websites, online textbooks, medical forums, etc.

One might quickly realize that Rashi is lacking features that some may argue are essential in games. Most notably, the system never comes to a formal conclusion (i.e., there is no way to present a final diagnosis and receive feedback on whether the player has succeeded or failed). Because this nascent form of Rashi lacks several potentially interesting game elements, we are provided with a useful control for studying potential improvements derived by more tightly coupling the system to the framework presented in this paper. To study this effect, we have analyzed the addition of three primary features:

First, a patient information and feedback panel was added to improve feedback and direct connections to / with the patient (the primary motive and goal for the player). This panel provides effective feedback to the player through a basic emotive response from a pedagogical agent. The agent's emotion has five states ranging from very upset to neutral to very happy. While this emotional model is extremely simple, it provides the opportunity to show how such simple interventions might lead to noticeable gains. The patient information panel also contains a patient health bar, displaying the relative health of the patient. The health of the patient deteriorates (very slowly) over time while students work on the case.

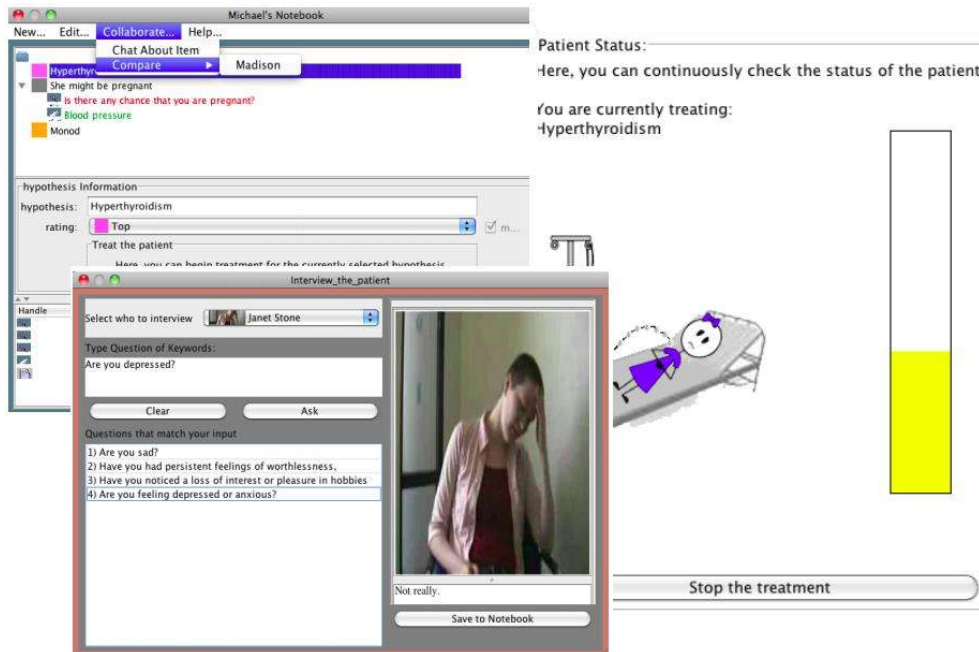


Figure 4: Rashi interface. Students can examine the case (bottom left), organize arguments in a notebook (upper left) and treat the patient (right).

Lastly, the new mechanics provide students with a treatment panel, allowing players to administer treatment of a specific hypothesis. If the treatment is consistent with the correct diagnosis for the case, the student will notice the health of the patient slowly improving. This process is neither timely nor monotonic, forcing players to busy themselves while waiting for results.

We argue that these additional mechanics explicitly address aspects of our framework missing from the control instance of Rashi. It is important to note that this analysis was done retroactively, and only provides evidence that our framework could be useful in explaining past results within serious games literature. We, unfortunately, cannot argue that our model was predictive in this particular scenario. With that being said, the new features of Rashi bolster the following framework elements

- *Motivating Goal:* By providing the student with a visualization of the patient and their overall health, while also providing a mechanism for improving that health, the game explicitly improves the motivation for the goal. This goal is to help the patient, in a general sense, but is now more directly linked to a piece of visual feedback (the health bar).
- *First-Hand Interaction:* The treatment functionality allows a deeper layer of direct interaction, in the sense that players can now administer treatment directly onto the patient and view the results of such actions
- *Difficulty Gaming the System:* Because treatment of the patient is slow, it does not provide the student with immediate results (e.g., good job, you got it right!). Therefore, any strategy involving a guess-and-check will prove overly cumbersome. For this reason, we posit this feature forces students to more carefully consider their choices, as they realize the first-order optimal strategy is to investigate the case as intended.
- *Clearly-Defined Goal:* The treatment panel finally provides Rashi with an interactive, and clear ending resulting in success or failure, as the user can visualize the patient getting better or continuing to worsen.

	FRAMEWORK ELEMENTS								
	Experience			Experiment			Evaluate		
	Firsthand / Interactive Mechanic	Encapsulating	Taxonomizing Information	Relating Ideas	Clustering	Labeling	Clearly Defined Goal	Motivating Goal	Difficult to Game the System
Rashi Old	Yellow	Green	Red	Yellow	Red	Yellow	Yellow	Yellow	Yellow
Rashi New	Green	Green	Red	Yellow	Red	Yellow	Green	Green	Green

Figure 5: Heatmap for Rashi and how well it conforms to our framework

Figure 5 (top) shows a comparison of how Rashi and its core features fit our proposed framework. In addition, figure 5 (bottom) shows how this mapping changes after introducing the additional mechanics described above (patient information, emotive responses, and treatment options). We can see that the features added couple this game more tightly to our framework, which we posit can explain the increases in learning outcomes (see results below).

3.2.1 Rashi Methodology

Past studies comparing a control instance of Rashi (basic features with no patient health bar, status, or treatment functions) to an experimental instance of the game containing the game elements have proven to be successful. The methodology and results of these studies are summarized in [15] and [16], but we wish to re-discuss them here, particularly in the context of how the results provide evidence of the utility of our proposed framework.

In [15], data was collected from a large rural university in 2011 (control group) and 2012 (experimental group). The participants were all students in the university's Biology 101 course. The Rashi activity was one of many labs required for the course. The students all worked on the same case, and were all given the same amount of time (~2 hours) to work under TA supervision. Researchers were present to clear up any confusion regarding the actual use of the system.

The goal of this study was to measure the quantity and quality of student work contributed in each group. Because log data is stored for all student interactions, quantity data is quite simple to collect. This includes the number of hypotheses the student is considering, the amount of data observations made during the investigations, and the number of explicit relations articulated within student arguments.

Additionally, it was important to measure / estimate the quality of student contributions, to ensure that additional contributions did not necessarily include poor quality information. The formula used to measure such quality combined several factors including the students content matched to an expert knowledge base as well as the breadth and depth of arguments. The detail on the formula used and how the expert knowledge base referenced works within it can be found in [16].

3.2.2 Rashi Results

When using the instance of Rashi containing game features more tightly coupled to our framework, students quantifiably completed more work within the system. Table 3 below summarizes this information. The number of projects refers to the number of student teams (3 students per team) that participated.

Table 3: Results comparing quantity of work between control and experimental groups in Rashi

<i>Group</i>	<i>Number of Projects</i>	<i>Per Project Data / Relations</i>	<i>Per Project Hypotheses</i>	<i>Total</i>
<i>Control</i>	396	10.965	5.331	16.295
<i>Experimental</i>	539	17.306	8.688	25.994

We see that the number of hypotheses (students manually articulate potential hypotheses within the system, and are asked to explore more than one) more than doubles even though the number of student projects does not go up by even 50 percent. The same trend can be seen for the number of data observations collected (students collect evidence to support their hypotheses in Rashi) and relationships (between data and hypotheses) constructed within the game. The total contributions per project rises from 16.295 to 25.994, an increase of 59.5 percent.

In addition, the quality of work students produce is measurably higher. Figure 6 below summarizes this increase ($p < 0.01$).

Effect of Games on Quality of Work in Rashi

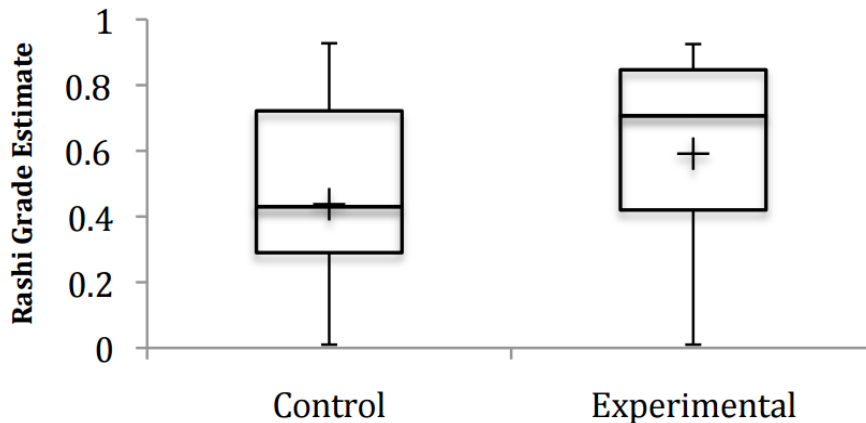


Figure 6: Results of quality of Rashi submissions across both groups (* $p < 0.01$)

3.3 Game 3: *The Legend of Zelda: Ocarina of Time™*

We now look at a game, traditionally viewed as non-educational, under the lens of our framework, both as an academic exercise as well as the basis for the argument that many traditional games are educational in their own right. While the Legend of Zelda™ series involves traditionally entertainment-focused mechanics (exploring, battling enemies, finding items, solving puzzles) the game effectively teaches players its puzzle-solving strategies with minimal verbal communication, and constructs a world specifically designed to “train and test” the player. In this regard, it is a prime example of games as learning machines [21] and demonstrate the implicit teaching that is done even in a non-pedagogically-focused domain.

Experience

First Hand Interaction: Most interactions in the game are quintessentially first-hand. These include speaking to NPC characters, using acquired items directly in the environment (bow-and-arrow, boomerang, bombs, etc.) and solving puzzles. The player has a constant NPC companion in the form of Navi, which will occasionally provide text assistance for challenges by suggesting actions or revealing details about the game, which acts as a “third-hand” experience for the player, and is easily skipped or ignored and consequently not very effective as a teaching tool. This is consistent with our findings in ‘A Nation Divided’, that explanatory text was often ignored by the player.

Encapsulation: Most dungeons in the game introduce players to a new item early, or make use of an item the player has recently acquired, and then continually encourage use of that item, encapsulating the idea. Consider the slingshot in the first dungeon of Ocarina of Time™. The entire dungeon becomes a contained exercise in learning how to use this new tool in the player’s arsenal, by introducing the tool itself, then introducing more complex challenges over time to encourage the player to experiment with new ways of using the tool. After the player has completed all the challenges, theoretically having learned all they needed to learn, there is a final test of mastery for the material, a boss fight. The boss, an armored spider monster, requires the player to use their newfound knowledge to exploit a weakness and defeat it, in this case using the slingshot on the boss’s eye (shown in image below). After beating a dungeon, there is a transition to the open world, making it clear that this particular experience is over.



Figure 7: Link using the slingshot to defeat the spider monster (the slingshot was the focus of instruction for this dungeon and the player must achieve mastery of this item in order to complete the challenge).

Taxonomy: Ocarina of Time™ does an excellent job categorizing the information within the game and communicating clearly-related elements. The game is actually quite complicated, often involving upwards of 10-20 special items that are each useful in different situations. Over the course of the game, the player is tasked with using each of these, often combining their effects to complete more complex challenges. Of course, the game uses progression to slowly introduce these elements to the player, and ensures that each item must be learned before the next can be attempted. Typically, unlocking the ability to start one dungeon is completely reliant on having the previous item and knowing how to use it. In this way, the game itself enforces the order of the player's learning, allowing the designers to ensure players have all requisite background knowledge to try new tasks.

Experiment

Relating Ideas: Ocarina of Time™ encourages players to incorporate prior knowledge into new challenges by trying out different ways to make use of combinations of items. For example, some challenges require the player to use the iron boots to stay underwater, and make use of the hookshot while doing so. Such challenges are presented after the player has prior understanding of both concepts, and create a challenge out of figuring out how to employ both tactics together. More generally, the game creates relationships between concepts by grounding new ideas in existing mechanics. For example, once the player uses the slingshot to learn about Z-targeting and free aim, later learning to use the same strategies with the Bow or the Hookshot is easier.

Clustering: When a new item is acquired, the player is almost always required to use that item immediately, thus showcasing right away the purpose of this item (temporal clustering). The player immediately knows why the item was given to them. For example, the hookshot helps reach areas that are initially out of reach. In addition, spatial clustering is used nicely in that target blocks do not appear in the game until the areas after the hookshot is acquired and usable and/or necessary. This alleviates the player from acquiring an item and floundering as to its use (say if the item was not usable for another three or four levels). Placing the objectives close to the location where a player acquires an item also naturally encourages them to test out their new item and try different approaches to using it.

Labeling: In addition, the artifacts on which the item can be applied are always clearly marked (e.g., the hookshot can grab onto large blocks that look like targets). This supports the clear relationship that the hookshot is used on target blocks, thus simplifying the lesson for the player. The hookshot is a recurring item in these games, and no game in the series allows you to grab onto any wall or any non-marked wall, for this would cause much confusion for the player who is trying to understand the scope and applicability of this new item. Another strategy that Zelda games employ is to use shape and color to make specific information stand out. For example, keywords in

conversation are often bolded or highlighted, and room markings often indicate with their designs some form of hint about where the player needs to go.



Figure 8: The block is clearly labeled with a target, informing the player that the hookshot can be used on the object.

Evaluate

Clear Goal: In order to prevent the player from becoming frustrated with wasted effort, which would get in the way of learning as well as reducing enjoyment of the game, goals must always be clearly defined, such that players can focus less on determining what challenge the designers want them to complete and more on the challenges themselves. Legend of Zelda™ games make use of a couple strategies for this. For instance, at times a player needs to visit a specific character, such as Saria. When this is the case, other characters will still interact with the player, but often drop a hint that the player needs to visit a different character. In dungeons, the game might take control of the camera for a moment to highlight a specific locked door, indicating that the player needs to find a way to unlock it. This, coupled with the map the game often makes available, helps the player know what their objective is, even when the means of completing it are intentionally obscured as part of a puzzle.

Motivating Goal: While it is essential that the game define goals clearly enough for players to work toward them, it is also important that players feel completion of the goal is meaningful. Failure to do so will likely make the objectives feel more like chores than challenges, and will reduce player enthusiasm for the tasks and thus for learning what they need to. Zelda games make use of many different tactics to make the in game goals motivating. The first is by establishing a strong narrative. From saving the Princess to helping out a friend, most quests in the game are grounded in some narrative explanation that motivates story-driven players to complete them. Players who are less influenced by story but care more about completion or power are rewarded for completing goals with new abilities such as ocarina songs, items such as the heart containers, or the in-game currency, Rupees. For the players that are not motivated by rewards or the narrative, the games are typically structured in such a way that new areas require items and skills gained in previous quests to unlock them. Thus players who only want to explore what the game has to offer are motivated to complete their goals, in order to continue exploring. While these strategies won't motivate all players, they are sufficiently broad to catch a good portion of the game playing population and keep them engaged throughout the experience.

Difficulty Gaming the System: Typically, with the exception of glitches, the dungeons in game cannot be beaten without demonstrating capability in the skills the designers wanted a player to learn in the dungeon. Some, such as the Shadow Temple, can still be done, but are far more difficult, relying on clever use of the abilities a player already has. As such, they are more popular in the community as challenges than as an alternate way to play through the game normally. This difficulty is important in an experiential game, because if a player can complete a challenge using new concepts or making use of material they already have mastered, it will be easier to use the skills they have than to try something new. The argument could be made that learning some skills should be optional, and not required to “win” but in general if time and effort is put into building the systems in game, the designers will most likely want learning the material to be mandatory for players. If glitches and workarounds exist, designers will likely want to correct them where possible, as these could be easier than the desired process would be (such as by glitching through a wall to avoid overcoming an obstacle) and allow players to circumvent learning the skills they need.

4. Discussion

While the EEE framework can be applied as a tool for assessing experiential games, insights gained while developing it have led us to believe it may be useful for guiding the development of experiential games in practice. The need for a well-defined set of experiences to mirror the content to be taught has reinforced the essential nature of designers working closely with subject matter experts (SMEs). We propose that SMEs should, before development of the game begins, identify the content that players should learn, encapsulate that content into distinct ideas, and taxonomize those encapsulated ideas to make it clear what knowledge is required and what concepts are related. This can be done using additional framework support from general serious game frameworks (e.g. Activity Theory-based Model of Serious Games). Having this information documented can help designers with categorizing and clustering information, as well as creating meaningful labels in game to assist players. From this outline, SMEs can work with the designers to brainstorm experiences for the game, including how the experiences should be changed to provide additional challenges as players learn, how they should be adapted to introduce related concepts intuitively, and what sorts of challenges would work well to evaluate how much players have actually learned.

Regardless of whether its application is in the development or assessment of a system, we have produced a framework, EEE, that is targeted towards a specific yet very popular genre of game: experiential games. While further research and application of this framework is necessary for drawing stronger conclusions, the three discussed games' natural adherence to topics found in our framework give some initial promise to its validity and its potential use. One important limitation however of our analysis is in how few of our tests were predictive in nature. The success of module 4 (Gettysburg) was predicted and validated through testing, but our other results are merely retroactive in nature. It is important that future work focus on applying the framework in an attempt to predict the success and failure of particular serious games, so as to uphold the stated theory. Only then can we gain true confidence in these ideas and, when necessary, adapt them.

Another important avenue of inquiry involves providing specific metrics to formulate precise predictive outcomes through the framework. Currently, our conceptual adherence to EEE is mostly qualitative, and then mapped onto a Likert scale. In the future, we hope to develop a core quantitative model for measuring adherence to the framework for a more useful predictive measure.

We believe that because of the experiential nature of many educational games in use today, our framework would serve as a useful additional asset to researchers and game developers. Through our examples of 'A Nation Divided' and 'Rashi' respectively, we demonstrate that this framework has potential as both a guide for development as well as an assessment of end quality. Even if a system being developed is not truly "experiential" as we and others have defined it, components of our framework such as the ability to game the system among others are still useful for developers to consider in the creation of their games. While not a strict guideline for development as other frameworks that have been developed, we argue that EEE's usefulness is in focusing the attention during design decisions to game and learning component changes that have the best outcome for producing effective systems.

In conclusion, we believe EEE to be a straightforward and useful tool for dissecting, understanding, and predicting the success of serious games. We developed our model after studies performed with three modules of our game 'A Nation Divided'. Our model then informed our design of a fourth module, which had great success. We then analyze some past results (Rashi, Legend of Zelda) under the lens of EEE in an attempt to contextualize and explain these results as a function of our model, and argue that in general, EEE adherence increases a game's chance of success. We believe EEE is useful both in designing serious games and for assessing their successes and failures.

References

- [1] Amory, A. "Game object model version II: a theoretical framework for educational game development." *Educational Technology Research and Development* 55.1 (2007): 51-77. <https://doi.org/10.1007/s11423-006-9001-x>
- [2] Winn, B. "The design, play, and experience framework." *Handbook of research on effective electronic gaming in education* 3 (2008): 1010-1024.
- [3] Hunnicke, R, LeBlanc, M., and Zubek, R. "MDA: A formal approach to game design and game research." *Proceedings of the AAAI Workshop on Challenges in Game AI*. Vol. 4. 2004.
- [4] Kelly, H., et al. "How to build serious games." *Communications of the ACM* 50.7 (2007): 44-49. <https://doi.org/10.1145/1272516.1272538>
- [5] Westera, W., et al. "Serious games for higher education: a framework for reducing design complexity." *Journal of Computer Assisted Learning* 24.5 (2008): 420-432. <https://doi.org/10.1111/j.1365-2729.2008.00279.x>
- [6] Burgos, D., Tattersall, C., and Koper, R. "Re-purposing existing generic games and simulations for e-learning." *Computers in Human Behavior* 23.6 (2007): 2656-2667. <https://doi.org/10.1016/j.chb.2006.08.002>
- [7] Dondlinger, M. J. "Educational video game design: A review of the literature." *Journal of applied educational technology* 4.1 (2007): 21-31.
- [8] Akkerman, S., Admiraal, W., and Huizenga, J. "Storification in History Education: A Mobile Game in and about Medieval Amsterdam." *Computers & Education* 52.2 (2009): 449–459. Web. 22 Feb. 2016.
- [9] Carvalho, M. B., et al. "An activity theory-based model for serious games analysis and conceptual design." *Computers & Education* 87 (2015): 166-181. <https://doi.org/10.1016/j.compedu.2015.03.023>
- [10] Rongas, B. "A Model-driven Framework for Educational Game Design." *International Journal of Serious Games* 3.3 (2016): 19-37. <https://doi.org/10.17083/ijsg.v3i3.126>
- [11] Nicholson, S. "Completing the Experience: Debriefing in Experiential Educational Games." 11.6 (2013): 27–31. Print.
- [12] Arnab, S., et al. "Mapping learning and game mechanics for serious games analysis." *British Journal of Educational Technology* 46.2 (2015): 391-411. <https://doi.org/10.1111/bjet.12113>
- [13] Dragon, T., et al. "Coaching within a domain independent inquiry environment." *Intelligent Tutoring Systems*. Springer Berlin Heidelberg, 2006. https://doi.org/10.1007/11774303_15
- [14] Murray, T., Woolf, B. and Marshall, D. "Lessons learned from authoring for inquiry learning: A tale of authoring tool evolution." *Intelligent Tutoring Systems*. Springer Berlin Heidelberg, 2004.
- [15] Floryan, M., and Park Woolf, B. "Improving the efficiency of automatic knowledge generation through games and simulations." *Artificial Intelligence in Education*. Springer Berlin Heidelberg, 2013. https://doi.org/10.1007/978-3-642-39112-5_36
- [16] Floryan, M. *Evolving expert knowledge bases: Applications of crowdsourcing and serious gaming to advance knowledge development for intelligent tutoring systems*. Diss. University of Massachusetts Amherst, 2013.
- [17] Baker, R., et al. "Why students engage in "gaming the system" behavior in interactive learning environments." *Journal of Interactive Learning Research* 19.2 (2008): 185-224.
- [18] Baker, R S., et al. "Off-task behavior in the cognitive tutor classroom: when students game the system." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2004. <https://doi.org/10.1145/985692.985741>
- [19] Cocea, M., Hershkovitz, A. and Baker R. S. "The impact of off-task and gaming behaviors on learning: immediate or aggregate" (2009): 507-514.
- [20] Lytle N., Floryan. M., "A Design Framework for Experiential History Games." *Games and Learning Alliance Conference*. Rome, Italy. December 10, 2015.
- [21] Gee, J. P., *Good video games + good learning: Collected essays on video games, learning, and literacy*. New York: P. Lang, 2007. <https://doi.org/10.3726/978-1-4539-1162-4>