Digital Studies / Le champ numérique, 2014

Assessing Serious Games: The GRAND Assessment Framework

Domini Gee, University of Alberta: gee@ualberta.ca

Man-Wai Chu, University of Alberta: manwai@ualberta.ca

Simeon Blimke, University of Alberta: simeon@ualberta.ca

Geoffrey Rockwell, University of Alberta: geoffrey.rockwell@ualberta.ca

Sean Gouglas, University of Alberta: sean.gouglas@ualberta.ca

David Holmes, University of Alberta: dmholmes@ualberta.ca

Shannon Lucky, University of Alberta: slucky@ualberta.ca

Abstract / Résumé

The videogame industry is a considerable market: in 2012, the industry was worth over $86 billion USD and about seventy-two percent of American households play videogames. It is unsurprising, then, that commercial and educational developers and/or researchers have sought to capitalise on videogames. Games and simulation technologies have been used for educational purposes for thousands of years prior to the digital era (Gee 2007). Digital games, however, offer many new affordances including increased accessibility, reinforced automation (i.e., fair and consistent application of rules), embedded data-gathering for assessment, dynamic adaptation to student needs, the ability to simulate complex situations for student inquiry in a safe context, and reduced overall costs (Jin and Low 2011). However, it is difficult to assess the process of serious game development and effectiveness of educational play. Many serious games retrofit assessment late into the project, creating a gap between original intents and the game's current uses, limiting effectiveness of measuring and meeting the project's goals. As such, we propose an assessment framework that synthesises work from various fields (educational assessment, game design, usability, project management) that aims to guide researchers and game developers through a project from its inception to the end by presenting specific topics to address and questions to answer throughout the game design phase of the project. By building assessment into the game development from the get-go, original intents and a game's current uses can more closely align, allowing for stronger, purposeful games.

L'industrie du jeu vidéo est un marché appréciable. En 2012, elle dépassait 86 milliards de dollars US et environ soixante-douze pour cent des ménages américains jouent aux jeux vidéo. Il n'est pas surprenant alors que des concepteurs et des chercheurs de la sphère commerciale et éducationnelle ont cherché à tirer profit des jeux vidéo. Bien avant l'avènement de l'ère numérique, jeux et technologie de simulation étaient utilisés à des fins pédagogiques (Gee 2007). Le jeu vidéo offre toutefois de nouvelles affordances : accessibilité accrue, automatisation renforcée (c.-à-d. l'application juste et systématique des règles), collecte de données intégrée pour l'évaluation, adaptation dynamique aux besoins des élèves, possibilité de simuler des situations complexes pour l'expérience de recherche de l'élève dans un contexte sécuritaire et, finalement, réduction des coûts globaux (Jin and Low 2011). Il est toutefois difficile d'évaluer le processus de conception de jeux sérieux et l'efficacité du jeu éducationnel. De nombreux jeux sérieux intègrent l'évaluation tard dans le projet, créant un écart entre les intentions originelles et les utilisations actuelles du jeu, limitant ainsi l'efficacité à mesurer et respecter les objectifs du projet. À ce titre, nous proposons un cadre d'évaluation synthétisant les travaux dans divers domaines (évaluation pédagogique, conception de jeu, facilité d'utilisation, gestion de projet) qui guiderait chercheurs et concepteurs de jeux dans un projet, de son lancement à sa fin, en présentant des sujets précis à traiter et des questions à répondre pendant la phase conception du jeu du projet. L'incorporation dès le départ de l'évaluation dans la conception du jeu permettrait un meilleur alignement des intentions premières et des utilisations actuelles du jeu, permettant ainsi une expérience de jeux plus forte et plus significative.

KEYWORDS / *MOTS-CLÉS*

Serious games, assessment, educational, game development, project management, framework

Contents

Introduction

Many popular computer and videogames are long, complex, and difficult, yet players welcome the challenges and are motivated to learn and continue through the game (Gee 2007). Well-designed games introduce skills to players while maintaining a high level of engagement, something many educators strive to achieve. Games and simulation technologies have been used for educational purposes for thousands of years prior to the digital era (Gee 2007). Digital games, however, offer many new affordances including increased accessibility, reinforced automation (i.e., fair and consistent application of rules), embedded data-gathering for assessment, dynamic adaptation to student needs, the ability to simulate complex situations for student inquiry in a safe context, and reduced overall costs (Jin and Low 2011). It is no wonder countless resources have been put into combining entertainment and learning in serious games and simulations (Gee 2007; ESA 2012). Research has found that players are able to learn content knowledge and a variety of skills through an entertaining videogame that engages and motivates the player (DiCerbo and Behrens 2012; Molka-Danielson 2009). Learning through videogames is important because it presents players with multidimensional learning environments allowing important 21st century skills (e.g., communication, problem solving, and critical thinking) to be taught and assessed throughout the game (ATCS21 2013).

With the popularity of videogames among school aged children, it is tempting to use videogames to teach students important skills while engaging students in a motivating digital learning environment. The GRAND team at the University of Alberta has created and tested some locative games (e.g., Return of the Magic and the Intelliphone Challenge). However, it has been difficult to find appropriate ways to assess the process of serious game development and the effectiveness of educational play. Research has shown assessment frameworks that are retrofitted to existing videogames limit the assessments' abilities to because specific requirements of the assessment framework may not be possible given the design of the game. As such, retrofitted assessment frameworks may not meet specific assessment framework goals nor have the ability to measure the goals (Gierl, Alves, and Taylor-Majeau, 2010). For this reason the GRAND team has created a framework for assessment during development through an extensive literature review of gaming, assessment, learning, and methods research to guide researchers through a project from its inception to the end by presenting specific topics to address and questions to answer throughout the game design phase of the project.

This paper is split into three parts. First, a discussion of the difficulties around the assessment of serious games and current frameworks that have been developed to assess serious games will be presented. Second, the assessment framework the GRAND team has developed from our literature review will be discussed. Finally, a case study involving assessing a locative game will be used to highlight the GRAND team's assessment framework.

Assessing serious games

In 2012, revenue for the videogame industry reached $86 billion USD with an estimate of seventy-two percent of American households playing videogames (ESA 2012). Of the videogamers in the US, thirty-two percent are school-aged children under 18 years of age (ESA 2011). According to one estimate, by age 18 the average young person will have invested roughly the same number of hours playing videogames as they will have devoted to formal schooling (Prensky 2006).With such a large population of students playing videogames, educators should capitalise on the opportunities to educate through games as students are both expending large amounts of time on videogames and motivated to tackle challenges presented in games.

The opportunities for educational games have not escaped the attention of developers and industry. Mizuko Ito in Engineering Play documents the history and commercialization of children's educational games (2009). More recently there has been a movement to make educational games a serious subject under the rubric serious games. Serious games took off as a field of academic development with the foundation of the Serious Games Initiative at the Woodrow Wilson Center for International Scholars in 2002. This Initiative has the goal to help usher in a new series of policy education, exploration, and management tools utilising state of the art computer game designs, technologies, and development skills. As part of that goal the Serious Games Initiative also plays a greater role in helping to organise and accelerate the adoption of computer games for a variety of challenges facing the world today. (Serious Games Initiatives 2002).

Serious games are defined as games designed for purposes other than pure entertainment. The purposes of these games may include learning in many fields such as defense, health care, city planning, and engineering, just to name a few (Adams 2013). As serious games gain traction in a variety of industries, usually for training and educational purposes, it is important for researchers and game developers to validate the educational claims of these games. (Messick 1989).

Many commercial educational games, such as *Math Blasters* and *Remission*, use videogame features to enhance students understanding of content knowledge and skills in a variety of domains, from mathematics to cancer treatments (Beale et. Al. 2009; Knowledge Adventure 2007). Although many videogames are created to make revenue, non-profit organisations, such as Canada's Center for Digital and Media Literacy, have also begun to use videogames as a means to teach students about cyber bullying and safe internet use (Media Smarts 2012). Research reports has found these educational serious games enhancing students' motivation and engagement as measured by students' self-reports (DiCerbo and Behrens 2012; Hsu and Chiou 2011; Molka-Danielson 2009). Additionally, empirical research has shown

students who learn through videogames gain a better conceptual understanding of the content and are better able to explain why responses are correct (Gee 2007).

Since many educational games are often designed to enhance students' knowledge and skills in hopes of increasing their academic achievement, research studies often focus on student scores on post-tests when compared to pre-tests (Gillispie, Martin, and Parker 2009). One problem with these pre- and post- tests is that they are often administered using a paper-and-pencil format while the learning environment included the use of technology-rich videogame simulations. This discrepancy between learning environment and test format is problematic because students need to learn a new educational format during the test which may hinder the measure of their knowledge and skills.

For example, Math Blaster allow students to learn algebra in a technology-rich learning environment using an interactive videogame where they are able to score points, lose "life" points when they make mistakes, "heal" themselves by solving more complex algebra problems, and get instant feedback throughout the semester (Knowledge Adventure 2007). However, at the end of the semester the teacher administered a paper-and-pencil multiple choice test where mistakes lower the score, feedback is given days after the exam was administered, and no opportunities are provided for students to make up for their mistakes. As such, when teachers teach the curriculum and students learn using videogames, it is important for the tests associated with these teaching and learning environments to also utilise videogame technologies to ensure an alignment is present between the three components of education: teaching (curriculum), learning, and testing (Pellegrino, Chudowsky, and Glaser 2011).

This discrepancy between the three components of education - teaching, learning, and testing - creates a need for serious videogames to assess students through embedded testing. Embedded tests are not administered at the end of a teaching lesson or unit, thus they break the mould of always being administered at the end of a learning session. Instead, embedded tests, as the name indicates, are embedded within an interactive digital learning environment and measure acquisition of knowledge and skills as the student is learning the tasks within the environment (Zapata-Rivera 2012). Embedded tests are designed to measure key, fine-grained learning objectives. In this way, embedded tests are believed to provide better evidence about claims related to student achievement because they are measuring learning as it is occurring. This aspect of embedded tests makes them different from traditional, standardised tests, which are often paper-and-pencil and reflect a static, one point-in-time measure of what and how much students know at the end of a teaching or learning unit. When embedded tests are used in a videogame, it is important to build the test components into the videogame during the initial design phase to ensure a consistent flow between the learning environment and the tests. However, research (see Gierl, Alves, and Taylor-Majeau 2010) has shown that many educational serious videogames retroactively build in an assessment framework after they have been created. This limits the usefulness of the videogame as a test because a gap exists between the original intent of the videogame and the current uses of the videogame. Thus, it is important to consult and be guided by an assessment framework from the initial game design planning stages to ensure the game is developed to meet specific goals (Chantam 2011). The next section presents some assessment frameworks that have been developed to assess videogames and guide the game design process.

Literature review of assessment frameworks

The assessment of educational serious games have often been linked to student achievement gained from playing the serious games (Tobias et. al. 2011). This idea of assessing the success of an educational serious game using student achievement was further supported by a cost analysis of serious games which showed a gain in 0.75 grade point score for a student would only require $400 using an educational serious game as compared to the same gain of 0.75 grade point score requiring $1, 170 of traditional classroom instruction (Fletcher 2011). However, using students' grades as a measure of a serious game's success has additional problems because many factors may contribute to improved achievement (e.g., personal tutor, increased interest in the topic, etc.) in the situation besides the use of a serious game. Additionally, many skills such as higher-order thinking and problem solving, which are often used while playing videogames, may not be measured using a simple grade point (Marsh 1991). As such, using student achievement as a measure to assess the success of a serious game seems to be misleading because they may not be linked.

The use of rubrics have also been relatively popular in assessing educational serious games. Rice provides a rubric of characteristics for a videogame to measure higher order thinking which asked players to rate the game using 20 true or false questions: e.g., "has a story line" or "avatars are lifelike" (Rice 2007, 93-94). Once players answer each of the 20 true or false questions the sum of the number of true they have answered will indicate the "cognitive viability" of the game. For example, a sum of 15-19 indicates the "game holds several positive characteristics lending itself to higher order thinking". Sauve, Renaud, Elissalde, and Hanca used a rubric of evaluation criteria in conjunction with the Learner Verification and Revision (LVR) methodology to assess online games (Sauve et. al. 2010; Komoski 1979; Komoski 1984). The LVR methodology consist of three phases (i.e., preparation, verification, and decision) which focus on test users' feedback to identify and correct errors and problems during the development process. This LVR method was used as a framework to develop three rubrics to be administered at different stages of the game development process. Using this LVR methodology as a framework with their evaluation criteria (e.g., playability, accuracy of information provided by the game, challenge, and active participation); a list of criteria was provided in a rubric to assess online computer games. Although rubrics have been developed to assess educational serious games, they tend to be very specific (i.e., assessing whether games measure higher order thinking) or they only provide a list of criteria for game designers to consider during the design process. Additionally, the LVR method did not encompass all the vital stages of the game design process. These point-in-time approaches of using a rubric to assessing a game tend to assess a game either after the game is completed or during the test phase of the game development, but have been proven to be problematic in because game development is a process that should be assessed continually (Shute 2011).

An approach popular in commercial game design literature is to inform design using an iterative design process, creating multiple versions of a game for continual assessment and improvement to ensure design goals are being reached (Tobias and Fletch 2011). Based on Chatham's experience as a consultant within the U.S. military's serious games

development programs he raised the issue that "usually government software development ends up shirking testing and assessment. These are the last things on the schedule, so when money inevitably runs short, they get cut, and users get flawed software". He notes one project, DARWARS *Ambush!*, was able to eliminate flaws and "petty annoyances that prejudice users upon first encounter with a product", by hiring professional third-party game usability testers (Chantam 2011, n. pag.).

Professional usability testing is a mechanism developed for commercial games, but because it addresses problems such as cognitive load and user frustration it may increase learning opportunity in educational games (Suave 2010). In contemplating a useful framework for assessing videogame, we considered the detailed guidelines provided by game design books (e.g., planning, designing, development, delivery, etc.). The literature provided advice from industry professionals that were scattered and unfocused on specific aspects of game design. For example, commercial game design literature suggested focusing on appealing characters, iterative designs, and marketing advice for beginners (Saltzman 1999; Michael 2003; Rogers 2010; Rollings 2003). Some literature would devote copious amounts of resources to discuss best practice suggestions or game design "principles" apart from the usual injunctions to use iterative design etc. (Despain and Acousta 2013). Most of these were organised roughly in sequence to typical activities conducted during game design pre-production/production/post-production cycles. These detailed outlines of the game design process did not provide a means to measure whether a game designer had executed each section successfully. Some researchers considered approaching the idea of assessing a videogame by using self-assessment type questions for game designers to answer while developing a game (Schell 2008).

By explicitly listing out these questions, game developers are probed for weak or unaddressed areas in their design concept. Perry and DeMaria (2009) give a modest list of 40 questions that address topics from game design to funding issues. Some of the items listed are less relevant for serious game (e.g., "Does the target audience already respect the developer of this game?" and "Does the game potentially have any collectable value?""Is it part of a series, for example?"), while others are more important for an engaging learning environment (e.g., "Can the game be customised or personalised?" and "Will the game have a fun and interesting learn-as-you-play in-game tutorial?"). The problem with Perry and DeMaria's list of questions are essential topics pertaining to the broad ideas of game design are mixed in with micro-level questions that address minute details of game play. On the other hand, Schell provides readers with literally hundreds of questions (i.e., over 400 pages of questions) to be answered during the game development process. (e.g., "What does the client say he wants, what does the client think she wants? What does the client really want, deep down in his heart?"). Although these detailed guidelines and self-assessment questions are informative in guiding the game design process, providing a checklist of necessary steps in game design, they do not provide us with a tool that could use to assess a videogame and are lacking in areas more specific to our work.

In addition to a focus on the overall game design process, we found it necessary to focus on literature that expanded on specific segments of game design that were lacking in the previous frameworks we explored. In particular, Walker (2003) has two chapters dedicated to feedback he gathered from industry insiders - including public relations representatives, corporate executives, and editors - and fans regarding what aspects sell videogames. Though he did not ask the same set of questions to all industry insiders, some of the questions he consistently asked were:

- What were four things that make games sell?
- Do licenses—such as "Official NBA" and "Star Wars"—enhance the game's sales?
- What creates buzz?
- And what is the most important thing to sell games?

What he asked fans, however, was one consistent set of questions. He asked questions such as what influences their gaming purchases the most, what is their favorite game and why, and whether they would buy games linked to a license they enjoy. Questions such as license and what influences gaming purchases link to some of the questions asked of industry insiders, which creates a linking between how industry and fans view common aspects of game design. However, the biggest flaw with Walker is while his questions can reveal interesting feedback about how industry insiders and players view the industry, he is primarily concerned with the commercial aspect. Most of his questions are related to issues that do not concern us, such as licensing and franchises, media relations, and sell-point connected to success.

This section has provided several assessment techniques from the fields of educational serious games and commercial game design. Although the techniques discussed are generally good for assessing videogames, each technique has serious flaws that could be enhanced by coupling these techniques together. For example, the idea of using rubric criteria is good for assessing a game, but the structure of the rubric should follow a strong game design framework such as the guidelines provided by Michael and Saltzman. The idea of an iterative design was good for providing continuous assessments of the videogame throughout all phases of the game from the beginning to the multiple revisions and enhancement of the game after the first administration. Providing a list of questions was also an effective measure to assess whether a game has accomplished the tasks in a section. From the techniques seen in the assessment of educational serious games and commercial game design, there was a need to combine several of these good techniques currently on published so that the weaknesses of one technique could be enhanced by the strengths of another. As such, there is a need to develop a new framework that encompasses the strengths of all the assessment techniques reviewed. The next section will introduce the GRAND Assessment Framework, which was developed by combining the strengths of each assessment techniques so that future researchers and game designers are able to use this framework to assess their videogame.

The GRAND assessment framework

At the centre of our assessment framework is this question: how do you know your game? We have done an extensive literature review of assessment theories, methods, and frameworks across fields, which include education literature, game design, game studies, usability and heuristic tests, and from our previous experiences building games. Though we have drawn inspiration from others' work and have found some crossover between different disciplines and ours, our

framework was built to accommodate our unique needs and work while being flexible enough to apply across different projects or disciplines. For example, while education is occasionally the primary goal of a project, other values such a fun, usability, design, deliverables, feedback and more, become just as important to creating immersive environments, efficient assessment methods, and overall stronger games.

Our assessment framework consists of a set of questions organised for game designers and researchers to ask at different points in development - from ideation to play, mapped onto assessment. Not all questions apply to all projects, but addressing the appropriate questions at the start of and during the design process helps projects to not only consider assessment early but helps teams design assessment into the project from the beginning so that they can meet project goals (and know what the goals were). With this in mind, we designed our framework around seven overarching areas drawn from the game design literature: stakeholders and expectations, requirements, resources, planning, design, feedback and closure.

Section	Definition	Rationale	Example Questions
Stakeholders and Expectations	The interested parties involved in or affected by the project.	Identifying these parties is important because stakeholders identify who has a stake in the project and asking creates transparency between the parties involved regarding how each party is contributing to the final project. It is important to identify the parties at the beginning of the project.	• (a) Who are the interested parties? • (b) What will stakeholders get out of the project? • (c) How will you prioritise stakeholders, including the audience?
Requirements	The requirements of the project.	Defining the requirements of the project is important because they identify what the project's goals are to begin with. This creates a common understanding among participating parties of what needs to be accomplished, what would be necessary to view the project as a success, and provides a general direction and criteria for the project.	• a) What is the primary purpose of the project? • b) What is the minimal, observable and measurable evidence indicating success? • c) How will interested parties communicate and is there anything like a project charter agreed upon by stakeholders that makes clear what is expected of everyone?
Resources	The availability of resources.	It is important to know before and during the project, especially before actually planning or building your game. What your resources are and how to manage them is crucial to determining if the project can fulfill primary goals, fulfill obligations, and allow for enough methods and alternatives to create a strong game.	• a) What resources do you need to achieve your purpose? • b)What resources do you have? • c)How will you get resources and how will you deal with the loss of resources?
Planning	Planning the project.	Planning takes what developers and researchers have established in previous sections and brings them together to create a project plan. A project plan is important to knowing	• a) Is there a project plan and what level

		who will contribute what at what time, how the project team will finish their project, and contingency planning for setbacks.	of detail does it entail? • b) Is there a method for regularly tracking progress? • c) How will the plan be adapted if you fall behind or lose resource?
Design	The actual building of the game.	Design is one of the most important sections. Though Design includes questions that are similar to previous sections', the key difference is the Design is primarily concerned with the game itself. This identifies what the game needs to succeed, how it should be built, and how to assess the game design process. Assessing the game design stage is important because it allows us to identify if the process is works (and whether adjustment is necessary) and, more importantly, is it fulfilling the goals of the project.	• a) What design process will work best for this project? • b) How can you assess the design process? • c) What resources are needed to develop the game? • d) Does the game meet your goals? Are there unexpected problems? • e) Are there ways to improve the game or platform based on the results of this attempt?
Delivery	Making the game accessible to the intended audience.	Delivery is how you will deliver the game to the audience once you have completed it. This involves not only ensuring they have access to play it but teaching them how to play the game.	• a) How will you deliver the product to your audience? • b) How will they become aware of your project? • c) How will you train the audience to understand and play the game?
Feedback	The feedback and responses.	Arguably one of the most important steps, Feedback identifies what sort of responses parties want and how to handle that feedback. Feedback is important because it identifies how the project has been received, what should be implemented in further and immediate modifications, and if your project is even getting the feedback its looking for. Furthermore, feedback may also highlight unexpected elements that were missed during the design process.	• a) What feedback do you and stakeholders want? • b) How will reports from the audience or other stakeholders will be handled? • c) Can we provide feedback to encourage player

Closure	The end of the project.	This is important because it asks what determines the end of the project and how do you wrap it up. End-State asks you to reflect on your results and how you will communicate these to the other parties once the project is over.	a) What will the end-state look like? b) Have you met your minimal observable and measurable evidence for success? c) How you will know if stakeholders are satisfied? d) Have you communicated to all interested parties that the project is over?

Together, these create our framework, known as GRAND Assessment Framework. The GRAND Assessment Framework was developed to fill the gap in the literature between educational assessments and game design. Additionally, the GRAND Assessment Framework is also practical because it was developed in conjunction with the development of an educational serious game called the Intelliphone Challenge.

The Intelliphone Challenge: the need for the GRAND Assessment Framework

While developing the Intelliphone Challenge, the researchers and game designer continually encountered problems regarding whether the game was meeting the initial goals. The researchers and game designers needed an assessment framework that could measure the success of the game design in meeting their goals.

The Intelliphone Challenge is a locative game developed by the University of Alberta Humanities Computing Department in partnership with the City of Edmonton's Fort Edmonton Park Historical Site. Utilising the FAR-Play geolocative game engine designed at the University of Alberta, the game encouraged patrons to explore Fort Edmonton through a guided narrative, which focused on three of the park's areas that correspond to three important time periods in the City of Edmonton's history: 1885 street, 1905 street, and 1920s street. The Intelliphone Challenge's game narrative follows the story of a fictional Edmonton family as they make their way through growth of both the City of Edmonton and the Canadian west. Each narrative focuses on a different aspect of the growth of the city that the park wanted to stress: security, communication, and community development.

Each time period was tied together by quizzes, each containing a part of the overreaching story. Players would search for scannable QR codes scattered across the park which upon scanning would receive a multiple choice question thematically linked to the location. Players who correctly answered the questions would be notified of achievements upon passing a threshold for correct answers (Play-Pr 2012).

Although the game was deemed a success by the clients, Fort Edmonton Park historical educators, there was still an issue of how the game development and feedback was assessed. Feedback from the design team and players were collected during the first administration of the game so that revisions and enhancements to the game could be made for the second administration of the game. The feedback from the first administration of the game helped bring our attention to the need for an assessment framework because some the feedback regarding the game's original intent and its uses could have been pre-emptively addressed. For example, much of the feedback pertained to the fact that better communication at the beginning would have set better foundations for the project. From the inception of the project, there are questions that could have been addressed immediately such as who the interested parties included. In the GRAND Assessment Framework we referred to this as "affiliation." For the Intelliphone Challenge the main interested parties included researchers and game designers from Humanities Computing and Computing Science from the University of Alberta and the Fort Edmonton Park educators. However, a secondary line of interested parties also included the city of Edmonton, students' teachers, ethics committees, affiliated organisations, and potential game audiences. This secondary line of interested parties have a differing 'stake' that needed to be considered when asking what each stakeholder was getting out of the project. This is why it is important to ask how to prioritise each group of stake holders, especially in relation to the primary aim of the project. After the interested parties were explicitly defined, there was a need to also define the "expectations" of the project. The primary purpose of the project is to create a game that allows Fort Edmonton Park attendees to experience the park through a guided narrative experience. However, the researchers and game designers from the University of Alberta had their own goals, or "stakes", of wanting to test and refine a previously developed gaming platform, named FAR-Play, and examine the user experience specifically focusing on user interface. Similarly, Fort Edmonton Park also had their own goals of whether attendees who play the game would have a better experience of the park through a game and whether this would increase Fort Edmonton Park's admission rates. The two goals of the University of Alberta's researchers and game designers and the

Fort Edmonton Park's historical educators are very different and needs to be attended to during the game design process. However, one of the most important interested party that needs to be considered is the participants, or players. While the game may work on a technical level and accommodate b oth groups' expectations, the Intelliphone Challenge would fail as a game if the participants are not engaged. Given that both groups' desired feedback and conditions of success were linked heavily to participant engagement, it is important to consider the level of engagement during the game.

Once the project started there was a need to track the project so that all interested parties were informed of the progress. The GRAND Assessment Framework referred to this as "planning". For the Intelliphone Challenge regular emails and bi-weekly meetings already planned. When these basic questions are raised during the initial phases of a project, it allows all interested parties to be aware of everyone's role and the final goal of the project. It also establishes an open and transparent foundation for the team to confirm, revise, build, and raise other questions.

Some of the problems encountered included the wireless internet connection at the park was weak and unstable making it difficult for participants to play the continually play the game. Additionally, many park attendees>did not have a smart phone, found the user interface incompatible with their electronic devices, or could not access the game online through their device. These feedbacks lead the researchers and game designers to consider questions such as "Are there ways to improve the game or platform of the delivery?" These issues later sparked the sections named "design" and "delivery" in the GRAND Assessment Framework.

The second administration of the Intelliphone Challenge was administered as a desktop game that could be finished prior to visiting the historical site, Fort Edmonton Park. However, this desktop version of the game preventing the researchers and game designers from the University of Alberta to reach their goal of investigating the FAR-Play gaming platform and user interface. This enhancement of the game decreased the amount of resources required by the team (i.e., participants would play the game on their own computer prior to visiting the park) and increased the return of feedback. Additionally, the second administration featured an enhanced game with further modifications made to the user interface, to reflect the feedback received, and introduced new counters and quizzes for additional feedback to the participants.

Throughout the Intelliphone Challenge, the team encountered many problems and challenges. The main lesson learned by the team was the importance of an assessment framework that would guide the game design process and raise questions to be asked by the team during the initial phase of the game design. As such, the GRAND Assessment Framework was developed to help future researchers prevent making the same mistakes as the Intelliphone Challenge team and to fill the gap in the literature between educational assessments and game design.

Conclusion

Our GRAND Assessment Framework was built to fill a gap based on a literature of previous theories and practices. In previous sections, we established that though there is more interest in serious games it is difficult to find appropriate means of assessing them. Many games retroactively build assessment after they are complete, which creates a gap between original intents and the game's current uses and limits the usefulness of assessment. To rectify this, we have built assessment into the game design process to more closely align original intents and the game itself. This is not an extensive framework but it provides grounds for further work in the future.

Further work on our framework would emphasise three focuses: first, revisions to the sections of the framework and associated questions to ensure these are the most-overreaching and significant questions to game design; second, practical application and testing of our framework on our games to test effectiveness; and third, the linkage of our framework to list of practices and methods. The last focus is not covered in this paper, but we hope that by linking the framework to practices and methods, our framework of questions would not only provide a useful way into assessment, but by mapping it to different practices and methods discussed in the literature (Consalvo and Dutton 2006; Annetta and Bronack 2010), it could serve as a guide for those looking for methods, whether formative or critical. Though answering the framework may give researchers and developers an idea of what their project is for or what they wish to assess for, they may not be aware of what are the best-suited (or ill-suited) methods for the research questions they want to answer.

Ultimately, a more refined framework can be developed to build a sort of toolbox for design and theory so that researchers and game designers can create better games and better assessment methods.

Acknowledgements

Works Cited

Adams, E. 2009. "The Designer's Notebook: Sorting out the Genre Muddle." *Gamasutra*. Accessed August 19, 2013. http://www.gamasutra.com/view/feature/4074/the_designers_notebook_sorting_.php?page=2.

Annetta, L., and S. Bronack. 2010. *Serious Educational Game Assessment: Practical Methods and Models for Educational Games, Simulations and Virtual Worlds.* Rotterdam: Sense Publisher.

ATCS21 (Assessment and Teaching of 21st Century Skills). 2009. *What are 21st Century Skills?*. Accessed August 19, 2013. http://atc21s.org/index.php/about/what-are-21st-century-skills/.

Beale, I.L., P.M. Kato, V.M. Marín-Bowling, N. Guthrie, S.W. Cole. 2009. "Improvement in Cancer-Related Knowledge Following Use of a Psychoeducational Videogame for Adolescents and Young Adults with Cancer." *Journal of Adolescent Health* 41.

Consalvo, Mia and Nathan Dutton. 2006. "Analysis: Developing a Methodological Toolkit for the Qualitative Study of Games." *Game Studies* 6 (2).

Chantam, R. 2011. "After the Revolution: Game-Informed Training in the U.S. Military." In *Computer Games and Instruction*, edited by Sigmund Tobias and J.D. Fletcher. Charlotte: Information Age Publishing.

Despain, Wendy and Keyvan Acousta. 2013.*100 Principles of Game Design*. Berkeley: New Riders.

DiCerbo, K. E, and J. T. Behrens. 2012. "Implications of the Digital Ocean on Current and Future Assessment." In *Computers and Their Impact on State Assessment: Recent History and Predictions for the Future*, edited by R. Lizzitz and H. Jiao, 273-306. Charlotte: Information Age Publishing.

ESA (Entertainment Software Association). 2011. *2011 Sales, Demographics and Usage Data: Essential Facts about the Computer and Videogame Industry*Accessed August 19, 2013. http://www.theesa.com/facts/pdfs/ESA_EF_2011.pdf.

ESA (Entertainment Software Association) 2012. *2012 Sales, Demographics and Usage Data: Essential Facts about the Computer and Videogame Industry.*Accessed August 19, 2013. http://www.theesa.com/facts/pdfs/ESA_EF_2012.pdf.

Fletcher, J.D. 2011. "Cost Analysis in Assessing Games for Learning." In*Computer Games and Instruction*, edited by Sigmund Tobias and J.D. Fletcher, 417-434. Charlotte: Information Age Publishing.

Gee, J. P.2007. *Good Videogames + Good Learning: Collected Essayson Videogames, Learning and Literacy.*New York: Peter Lang Publishing.

Gillispie, L., F. Martin, and M. Parker. 2009. "Effects of the Dimension-M 3D Video Gaming Experience on Middle School Student Achievement and Attitude in Mathematics." In *Proceedings of Society for Information Technology and Teacher Education International Conference 2009*, edited by I. Gibson et al., 1462-1469. Chesapeake: AACE.

Gierl, M. J., C. Alves, and R. Taylor-Majeau. 2010. "Using the Attribute Hierarchy Method to Make DiagnosticInferences about Examinees' Knowledge and Skills in Mathematics: An Operational Implementation of Cognitive Diagnostic Assessment." *International Journal of Testing* 10:318-341.

Hsu, T. Y. and G. F. Chiou. 2011. "Preservice Science Teachers' Prior Gameplay Experience and their Perceptions of Digital Game-Supported Learning." In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2011*, edited by T. Bastiaens and M. Ebner, 3331-3339. Chesapeake: AACE.

Ito, Mizuko. 2009. *Engineering Play: A Cultural History of Children's Software*. Cambridge: The MIT Press.

Jin, P. and R. Low. 2011. "Implications of Game Use for Explicit Instruction." In*Computer Games and Instruction*, edited by Sigmund Tobias and J.D. Fletcher. Charlotte: Information Age Publishing.

Kirkle, Sonny E., Steve Tomblin, and Jamie Kirkley. 2005. "Instructional Design Authoring Support for the Development of Serious Games and Mixed Reality Training." In *The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*. Arlington: National Defense Industrial Association.

Knowledge Adventure. 2007. "About Knowledge Adventure." *Math Blaster*. Accessed August 19, 2013. http://www.mathblaster.com/About.aspx.

Komoski, Ken. 1979. "Counterpoint: Learner Verification of Instructional Materials." *Educational Evaluation and Policy Analysis* 1 (3):101-103.

Komoski, Ken. 1984. "Formative Evaluation: The Empirical Improvement of Learning Materials." *Performance and Instruction Journal* 22 (5): 3-4.

Marsh, Herbert W. 1991. "Multidimensional Students' Evaluations of Teaching Effectiveness: A Test of Alternative Higher-Order Structures." *Journal of Educational Psychology* 83 (2): 285-296.

Media Smarts. 2012. "Educational Games." *Media Smarts*. Accessed August 19, 2013. http://mediasmarts.ca/digital-media-literacy/educational-games.

Messick, S. 1989. "Validity." In "Educational Measurement, 3rd Edition" ,edited by R. L. Linn, 13-103. New York: Macmillan.

Michael. David. 2003. *The Indie Game Development Survival Guide*. Hingham: Charles River Media.

Molka-Danielsen, J. 2009. "The New Learning and Teaching Environment." In *Learning and Teaching in the Virtual World of Second Life*, edited by J. Molka-Danielsen and M. Deutschmann, 13-26. Trondheim: Tapir Academic Press.

Pellegrino, J. W., N. Chudowsky, and R. Glaser. 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington: National Academy Press.

Perry, David and Rusel DeMaria. 2009. *David Perry on Game Design: A Brainstorming Toolbox*. Boston: Charles River Media.

Play-Pr. 2012. "The Intelliphone Challenge: Fort Edmonton Park Geo-Locative Game." Accessed August 10, 2013. http://playpr.hexagram.ca/projects/.

Prensky, Marc. 2006.*"Don't Bother Me Mom, I'm Learning!": How Computer and Videogames Are Preparing Your Kids for Twenty-first Century Success and How You Can Help!*.St. Paul: Paragon House.

Rice, John. W. 2007. "Assessing Higher Order Thinking in Videogames." *Journal of Technology and Teacher Education* 15 (1): 87-100.

Rollings, Andrew. 2003. *Andrew Rollings and Ernest Adams on Game Design*. Indianapolis: New Riders.

Rogers, Scott. 2010. *Level Up!: The Guide to Great Videogame Design*. Chichester: Wiley.

Saltzman, Marc. 1999. *Game Design: Secrets of the Sages*. Indianapolis: Macmillan Digital Pub.

Sauve, Louise. 2010. "Usability Guidelines for a Generic Educational Game Shell." In *Educational Gameplay and Simulation Environments: Case Studies and Lessons Learned*,edited by David Kaufman and Louise Sauve, 416-434. Hershey: Information Science Reference.

Sauve, Louise, Lise Renaud, and Gabriela Hanca Elissaide. 2010. "Formative Evaluation of an Online Educational Game." In *Educational Gameplay and Simulation Environments: Case Studies and Lessons Learned*, edited by David Kaufman and Louise Sauve, 416-434. Hershey: Information Science Reference.

Serious Games Initiatives. 2002. "The Serious Games Initiatives." *Serious Games*. Accessed August 19, 2013.http://www.seriousgames.org/.

Schell, Jesse. 2008. *The Art of Game Design: A Book of Lenses*. San Francisco: Morgan Kaufmann Publishers Inc.

Shute, Valerie. 2011. "Stealth Assessment in Computer-Based Games to Support Learning." In *Computer Games and Instruction*, edited by Sigmund Tobias and J.D. Fletcher, 503-524. Charlotte: Information Age Publishing.

Tobias, Sigmund, J.D. Fletcher, David Yun Dai, and Alexander P. Wind. 2011. "Review of Research on Computer Games." In *Computer Games and Instruction*, edited by Sigmund Tobias and J.D. Fletcher, 112-222. Charlotte, NC: Information Age Publishing.

Tobias, Sigmund, and J.D. Fletcher. "Computer Games, Present and Future." In *Computer Games and Instruction*, edited by Sigmund Tobias and J.D. Fletcher, 525-546. Charlotte, NC: Information Age Publishing.

Walker, Mark. 2003. *Games That Sell!*Plano: Wordware Pub.

Zapata-Rivera, D. 2012. *Embedded Assessment of Informal and Afterschool Science Learning.* Paper presented at the Summit on Assessment of Informal and After-School Science Learning.