An Approach to Holistic Development of Serious Games and Learning Simulations

Aleshia T. Hayes

University of Central Florida, Synthetic Reality Lab, Orlando Florida
Aleshia.prof@gmail.com

Abstract. This discourse is an argument for a holistic approach to developing learning games and computer mediated experiences through the intersections of the areas of efficacy, effectiveness, and user experience in designing and developing serious games and simulated learning experiences. Some examples are explored in which reasonably effective design approaches could have been improved by a more holistic and iterative approach. The approach includes the integration of learning objectives, outcomes, usability, motivation, experience, ludus, aesthetics, cost and sustainability of the systems based on research within the fields of education, learning theory, game design theory, and simulation. These constructs explain the need for an iterative and holistic approach to designing and developing learning games. Embracing iterative and learning centered design of serious games will perpetuate development of effective educational technology.

Keywords: Educational Technology, Efficacy, Engagement, Serious Games, Evaluation.

1 Introduction

Millions of dollars and countless hours are invested in the development of serious games for education internationally every year. While it is widely accepted that games can be effective and engaging learning tools, there is apprehension about the use of technology. This apprehension appears in the discourse on high level discussions of education in technology, such as when Education Nation Town Hall Talks moderator, Alex Witt, posed a question about the value of technology in the classroom with, "…we don't have a lot of concrete evidence that the billions and billions of dollars that schools are using to try and introduce technology into the classrooms are working. We don't have the evidence so far to say that test scores are being raised, that students will not drop out [10]". In addition to concerns about the potential that certain serious games or simulations may be ineffective, the potential of negative training, in which students learn the wrong thing, must be considered in development of learning systems [12]. In the US Department of Education's National Education Technology Plan, there is an explicit call to begin making more "data driven decisions" about how to acquisition and use of educational technology.

P. Zaphiris and A. Ioannou (Eds.): LCT 2014, Part II, LNCS 8524, pp. 42–49, 2014.

1.1 Gap in the Existing Approach

A gap exists in the comprehensiveness of each of these individual evaluation procedures, as the user experience analysis does not necessarily quantify the effectiveness and efficacy of the specific game or experience being evaluated. In his call for more purposeful game design, Dr. Konstantin Mitgutsch of MIT Singapore writes that "Serious games are mainly assessed in terms of the quality of their content, not in terms of their intention-based design" [17]. This void in assessment of intentionality of design largely explains the apprehension that is arising in public forums about the value of technology in the classroom.

This paper is a call for a holistic iterative approach to design and development as well as assessment and evaluation of serious games and learning environments. For the purposes of this paper, holistic is defined as, relating to or concerned with wholes or with complete systems rather than with the analysis of, treatment of, or dissection into parts [16]. Because of the iterative nature of the design and consumption of games, this discourse aims to inform the processes of assessment (formative and summative), evaluation, and design. While the two terms are often erroneously used interchangeably, assessment and evaluation are referred to as distinct processes. Evaluation relates to placing a summative value judgment on an entity or assigning worth to it [1], whereas, assessment refers to a formative process of improvement that is often diagnostic in nature [1], [15], [17]. Test-driven development resembles this approach, but is still missing elements of the specifications, requirements and objectives necessary in educational technology.

1.2 Partnerships Yield Holistic Design and Assessment

Partnership between designers, educators, developers, testers, funding sources, and consumers will resolve many of the questions of efficacy and effectiveness of serious games and simulations as a shared lexicon creates a productive discourse that contributes to the evolution of educational technology. Professional evaluators of serious games and learning systems (e.g. DOE administrators, school districts, software company employees, and bloggers) have a myriad of constructs to consider when evaluating a system.

Not only do they need to answer multiple questions relating to learning objectives, outcomes, usability, motivation, experience, ludus, aesthetics, cost and sustainability of the systems they evaluate, but they also have to look at the intersection of these variables. Does the system do what it says it will do? How hard is it to learn and use? How fun is it? How much and how often will the children play it? At what level will they learn from this system? How sustainable is the investment into this system? Because these considerations are complex and range multiple disciplines, evaluation is often distilled to the lowest common denominators of fun, learning outcomes, and cost. Because of this, design is often also limited to these superficial factors, in an attempt to satisfice the needs of stakeholders [21].

1.3 Constructing a Holistic Design and Assessment Approach

Design should be objective oriented, assessment, on the other hand, pertains to gaining information, such as the needs or accomplishments of an entity [1]. If designers and developers are creating technology to meet the needs of the clients, it behooves these clients to explicitly determine and express their expectations in a holistic way. While both assessment and evaluation are beneficial to game design and measurement, the distinction is important. Parents deciding whether or not to spend $60 on a learning game may wish to evaluate that game or view the evaluation of a professional evaluator. The funding agency supporting development may request a needs assessment or a formative assessment before providing thousands to fund development of a learning system. That same funding agency may follow up by assessing or evaluating the game at certain checkpoints or upon completion.

2 Elements of the Holistic Approach

Developers and designers often evaluate or assess the learning systems that they develop through some objective measure as determined by their goals and values; typical approaches to collecting metrics on systems include usability analysis, user experience, efficacy, effectiveness measures [6], [22], [23]. Some prevalent ways to evaluate the efficacy and effectiveness of learning tools include performance improvement assessments, blind coder ratings, qualitative and quantitative self-reports of social presence, questionnaires, and ultimately performance tests that measure improvement in desired knowledge, skills, or abilities [3], [4]. This discourse calls for the integration of these approaches with innovations in design approaches and subsequent assessment of systems include the learning objectives, MDA (Mechanics, Dynamics, and Aesthetics) Framework, and Game Flow [5], [18]. While measures are often internally designed or imposed upon the development team by the sponsoring agency, this framework can ensure thorough consistency when comparing similar education tools.

Table 1. Elements of a Holistic Approach to Development, and Evaluation

Construct	Formative or Summative
Learning Objectives	Formative
Learning Outcomes	Summative
Usability	Summative
User Experience	Summative
Motivation	Formative & Summative
Ludus	Formative & Summative
Aesthetics	Formative & Summative
Cost	Formative & Summative
Sustainability	Formative & Summative

2.1 Learning Objectives

Whether learning objectives in the serious games are explicit and didactic, or more discovery or inquiry based, there are too few explicit objectives for any serious game that are based on learning outcomes. "The content of a serious game could be well presented, adequately formulated, "correct" or irrelevant, hard to access or insufficient, and in worst cases, just wrong and biased" [17,4] The Serious Game Design Assessment framework posits an approach that looks at content in terms of whether the game environment and mechanisms reflect the necessary information to adequately transfer, and fidelity. This looks at fidelity in terms of name accuracy, narrative, and rules.

Universal Design for Learning is an example of an approach derived from architectural approach, universal design, that include designing structures to be used by people of varied levels of ability. One instrument that may inform the process of game and simulated environment assessment and usability is a 12-item questionnaire that integrates usability and principles of Universal Design, the (RAPUUD) Rapid Assessment of Product Usability & Universal Design [2]. While this tool is intended for assessment of consumer products, this tool provides a starting point for the underlying concepts to be cross-validated to integrate existing usability techniques with structured universal design for learning (UDL) constructs. The Universal Design for Learning principles that have been accepted by the US Department of Education are:

- Provide multiple and flexible methods of presentation of information and knowledge. Examples include digital books, specialized software and websites, text-to-speech applications, and screen readers.
- Provide multiple and flexible means of expression with alternatives for students to demonstrate what they have learned. Examples include online concept mapping and speech-to-text programs.
- Provide multiple and flexible means of engagement to tap into diverse learners' interests, challenge them appropriately, and motivate them to learn. Examples include choices among different scenarios or content for learning the same competency and opportunities for increased collaboration or scaffolding [9].

2.2 Learning Outcomes

Learning objectives and outcomes can be reciprocally impacted by the universal design for learning. Instructional design principles also inform the development of learning objectives. Learning objectives need to be explicit and measureable. To be more effective, effective learning goals can be broken into tasks and subtasks that each have measurable goals. Similarly, learning outcomes should be measured against appropriate learning objectives. Any analysis of learning outcomes without effective learning objectives will be less potent and potentially inaccurate.

2.3 User Experience

Measurements of user experience conducted by designers are generally positive, indicating user engagement, enjoyment, and preference over other methods of instruction [8], [20]. While this is a valuable finding, which some may rely on to allude to engagement, the often subjective construct of engagement without demonstrated benefits to learning has a diminished generalizable meaning. As Frokjear explains, the correlation between user satisfaction and effectiveness are often negligible and should be looked at separately [12]. Because outcome research results are specific to the samples (or populations from which they were drawn) and the outcomes measured, "it is essential that conclusions from the research be clear as to the population(s) and outcomes for which efficacy is claimed" [11]. Flay goes on to explain that, "Effectiveness trials test whether interventions are effective under 'real-world' conditions or in 'natural' settings. Effectiveness trials may also establish for whom, and under what conditions of delivery, the intervention is effective" [11].

User experience is a general term that often encompasses the perceptions derived through interactions that a user has with a system [22] For the purposes of this discourse, user experience refers to the levels of engagement, presence, ludology, learnability, memorability, and general satisfaction with the interface. The MDA framework addresses much of this in its' synthesized approach to considering of mechanics, dynamics, and aesthetics. Specifically, their objective is to understand and "decompose" the formal process of decision-making during gameplay in order to inform design, criticism, and research (pg. 4).

In the domain of student education, engagement is a construct /that captures the quality of students' participation with learning activities in the classroom, ranging from energized, enthusiastic, focused, emotionally positive interactions with academic tasks to apathetic withdrawal [19]. Skinner et. al. provide an instrument that is effective in measuring a learner's levels of Engagement vs. Disaffection with learning.

2.4 Usability

The International Standards Organization has amalgamated usability as effectiveness, efficiency, and satisfaction in reference to users being able to achieve an intended goal using the specified object in an intended context [22]. Notably, at least one of the intended goal for learning simulations and serious games should be achievement of learning objectives. Usability is often criticized as being simplistic, dichotic and excessively generalized [15]. The use of a single test to determine the viability of a product is often insufficient. Some studies are conducted on as few as five participants and the results are then generalized to the population as a whole [22]. While this approach is very effective for a formative or even summative assessment, the criticisms have arisen from the use of these studies to invalidate systems as having poor usability with limited metrics [15]

These critiques of usability analysis are due to the fact that usability studies were not intended to be holistic assessments or evaluations of systems. Instead they were

to look at effectiveness, efficiency, and satisfaction in terms of standards [2]. Additionally, usability standards are generally applied to many products, rather than being designed to measure games specifically. Mitgutsch explains, "investigating their impact becomes incomplete if the games' purpose and their coherence in relation to their design is not identified beforehand. Therefore, we argue that research on the impact of serious games starts with the analysis and evaluation of their qualities in terms of their purpose based formal conceptual design" [17]. The reciprocal relationship between design and assessment this article combines the assessment and design constructs to give a thorough analysis of the elements involved in this undertaking.

2.5 Ludus and Motivation

While there are other motivations to engage with serious games, the key element that distinguishes these tools from other educational technology and even traditional approaches to learning is the motivation that is generated from the ludic nature of the experience [19]. Unfortunately, this essential element is often overlooked in the evaluation, assessment, and even design of serious games. The capacity to immerse oneself in a ludic learning experience is becoming more powerful as the technology advances, with tools such as AR (Augmented Reality), MR (Mixed Reality), VR (Virtual Reality) [4], [6], [8]. But, the capacity to create ludic learning does not guarantee such an experience. Multiple elements of engagement, immersion, participation, and discovery must be integrated in order to make these experiences anything more than pixels on a screen.

2.6 Examples

There are several examples where this has been done well, and far more examples in which the attempts have been abysmal. One of the more successful efforts includes the BrainPop products which integrate usability and user experience in their deployment with the targeted and explicit learning objectives and measures of learning. Another successful effort was the River City Project, which modeled the approach of successful games, World of Warcraft and The Sims to create a learning experience in a MUVE (Multi-user Virtual environment) [8]. Both of these experiences are known to increase understanding and self-efficacy, which may be correlated to higher ratings of user enjoyment, engagement, and satisfaction.

2.7 Iterative Implementation

The synthesis of the different approaches to assessment and evaluation may be daunting if conducted simultaneously. Fortunately, the alternative to an exhaustive study evaluating a system on every possible level is to take the iterative approach. The ADDIE model used in instructional system design allows for an iterative approach [7]. The iterative approach requires careful consideration of the interaction of all of the elements of a learning system in order to avoid making small changes that

inadvertently change the entire experience. Further research will reveal if there is a correlation or even causal relationship between the holistic constructs within these learning experiences.

3 Conclusion

Because of the nascent nature of this research, the author's recommendation is to consider the synthesis of these principles as heuristics to guide the assessment and design of serious games and learning simulations, rather than as the basis for a finite algorithmic model. In some cases very few of these considerations need to be conducted, while in other cases, the complexity of the learning system or learning objective may demand a more extensive evaluation. This is not intended to drive the development of a rubric by which to judge a learning system, rather, these ideas can be used to shape a more informed and empowered perspective and the discourse that crosses both game designers and consumers or evaluators of those systems. The heuristics that come from a shared discourse about games can lead to a holistic approach to understanding assessing evaluating and designing serious games and simulations.

References

1. Alkin, M.C.: Evaluation Essentials: From A to Z. Guilford Publications (2010)
2. Asundi, K., Lenker, J.A., Nasarwanji, M., Paquet, V., Feathers, D.: A tool for rapid assessment of product usability and universal design: Development and preliminary psychometric testing. Work 39(2), 141–150 (2011)
3. Bailenson, J.: The Effects of Fully Immersive Virtual Reality on the Learning of Physical Tasks. Paper Presented at The 9th International Workshop on Presence Cleveland, Ohio (2006)
4. Botella, C., Bretón-López, J., Quero, S., Baños, R., García-Palacios, A.: Treating cockroach phobia with augmented reality. Behavior Therapy 41(3), 401–413 (2010)
5. Csikszentmihalyi, M.: Flow: the psychology of optimal experience / Mihaly Csikszentmihalyi (1st Harper Perennial Modern Classics ed.). Harper Perennial, New York (1990)
6. Dede, C.: Immersive Interfaces for Engagement and Learning. Science 323(5910), 66–69 (2009)
7. Dick, W., Carey, L., Carey, J.O.: The systematic design of instruction. Merrill/Pearson, Upper Saddle River, N.J (2009)
8. Dieterle, E.: Neomillennial Learning Styles and River City. Children, Youth and Environments 19(1), 245–278 (2009), http://www.colorado.edu/journals/cye (retrieved (date))
9. DOE (Department of Education). National educational technology plan. Transforming American education: learning powered by technology (2010)
10. Education Nation (Producer). Will Technology Transform Learning? (2012), http://www.educationnation.com/index.cfm?objectid=17C1CAAD-FE84-11E1-9448000C296BA163 (retrieved)

11. Flay, B.: Standards of Evidence (2004), http://www.preventionresearch.org/StandardsofEvidencebook.pdf (retrieved)
12. Frokjaer, E., Hertzum, M., Hornbaek, K.: Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? (2000)
13. Harteveld, C.: Triadic Game Design Balancing Reality, Meaning and Play (2011), http://site.ebrary.com/id/10452578
14. Hogue, J.R., Brickman, N.S., Pelz, C.A., Markham, S.: Virtual Reality Parachute Simulation Enhancements for Improved Aircrew Emergency Training (2005)
15. Hornbaek, K.: Dogmas in the assessment of usability evaluation methods (English). Behaviour & Information Technology (Print) 29(1), 97–111 (2010)
16. Merriam Webster. Holistic. from Encyclopedia Brittanica (2012), http://www.merriam-webster.com/dictionary/holistic
17. Mitgutsch, K., Alvarado, N.: Purposeful by design?: A serious game design assessment framework. Paper Presented at the Proceedings of the International Conference on the Foundations of Digital Games, Raleigh, North Carolina, http://gambit.mit.edu/readme/academic_papers/fdg2012_submission_82-1.pdf
18. Robin, H., Marc, L., Robert, Z.: MDA: A Formal Approach to Game Design and Game Research (2004), doi:citeulike-article-id:382810
19. Skinner, E.A., Kindermann, T.A., Furrer, C.: A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom (2009)
20. Thomas, R.G., William John, N., Delieu, J.M.: Augmented Reality for Anatomical Education. Journal of Visual Communication in Medicine 33(1), 6–15 (2010), doi:10.3109/17453050903557359
21. Tullis, T., Albert, B.: Chapter 10 - Case Studies Measuring the User Experience, pp. 237–287. Morgan Kaufmann, San Francisco (2008a)
22. Tullis, T., Albert, B.: Measuring the user experience: collecting, analyzing, and presenting usability metrics. Elsevier/Morgan Kaufmann, Amsterdam, Boston (2008b)
23. Villalta, M., Gajardo, I., Nussbaum, M., Andreu, J.J., Echeverria, A., Plass, J.L.: Design Guidelines for Classroom Multiplayer Presential Games (CMPG). Computers & Education 57(3), 2039–2053 (2011)