SGSEAM: Assessing Serious Game Frameworks from a Stakeholder Experience Perspective

Yongwen Xu, Philip M. Johnson, Carleton A. Moore,
Robert S. Brewer, Jordan Takayama
Collaborative Software Development Lab, University of Hawaii at Manoa, Honolulu, HI 96822 USA
{yxu, johnson, cmoore, rbrewer, jktakaya}@hawaii.edu

ABSTRACT
Assessment of serious game frameworks is emerging as an important area of research. This paper describes an assessment mechanism called the Serious Game Stakeholder Experience Assessment Method (SGSEAM). SGSEAM is designed to provide detailed insights into the strengths and shortcomings of serious game frameworks through a stakeholder perspective based approach. In this paper, we report on the use of SGSEAM to assess Makahiki, an open source serious game framework for sustainability. Our results provide useful insights into both Makahiki as a serious game framework and SGSEAM as an assessment method.

Author Keywords
serious games; framework assessment; sustainability

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g., HCI): Miscellaneous K.8.0. Personal Computing Games

General Terms
Serious Game; Assessment; Game Design; Case study.

INTRODUCTION

Serious games (games with additional goals beyond just entertainment) have been a topic of academic research for decades [11]. The recent phenomenon of gamification [4] also calls for evaluation research in areas beyond traditional entertainment purposes.

One fundamental question in evaluating a serious game or a gamified application is the extent to which the game or application achieves its "serious" purpose. This is quite different from traditional entertainment games. There is an increasing focus on the evaluation methodology in the field of serious games [9] [6]. These approaches focus on evaluation of a single game, as opposed to a game *framework*. One of the benefits of using a game framework is that, if correctly designed, it will provide useful and reusable "building blocks" with which to develop a variety of serious games. Yet how are we to know if a serious game framework has been "correctly designed"?

Gamification'13, October 2-4, 2013, Stratford, ON, Canada.

There exists some assessment tools such as GEQ (Game Engagement Questionnaire)[1], QUIS (Questionnaire for User Interaction Satisfaction)[5]. We found no prior work concerning comprehensive assessment for the particular needs of a serious game framework. To help answer this question, this paper proposes a method for assessing serious game frameworks, called the Serious Game Stakeholder Experience Assessment Method (SGSEAM). We consider SGSEAM as an assessment method instead of an evaluation method. The main purpose of an evaluation is to "determine the quality of a program by formulating a judgment" [7]. An assessment, on the other hand, is nonjudgmental. SGSEAM does not try to judge a framework according to a standard, or to compare one framework against another. Instead, it is used to identify the major strengths and shortcomings of a framework from the perspectives of major stakeholders. The benefits of SGSEAM assessment are for the developers of serious game frameworks to learn from the findings of the assessment.

SERIOUS GAME STAKEHOLDER EXPERIENCE ASSESSMENT METHOD (SGSEAM)

The goal of SGSEAM is to identify (a) major strengths of a serious game framework, which aids the community by indicating features of the framework to emulate, and (b) major shortcomings of the framework, which aids the community by indicating features to avoid. The target audiences of SGSEAM are the developers of the serious game framework.

The approach that SGSEAM uses is to assess the experiences of various important stakeholders when they interact with the serious game framework. In the full life cycle of a serious game framework there are a great variety of potential stakeholders, including:

- *Players*: those who participate in the game produced by the framework.
- *System admins*: those who install and maintain the technological game infrastructure.
- *Game designers*: those who design the content and game mechanics. They include content experts, instructional designers, etc.
- *Game managers*: those who manage the game during the period of game play.
- *Developers*: those who extend, enhance and debug the game framework.
- *Community partners*: those who partner with the game organizers to help run the game (such as coordinating real-world events as part of the game).
- *Funding organizations*: the organizations who provide funding for the game or game framework.

Stakeholder	Assessment Goal	Assessment approaches
Players	To what extent does the system affect players? To what extent does the system engage players?	experimental study, interviews, engagement metrics
System admins	How easy is it to install and maintain the system?	experimental study, interviews
Game designer	How easy is it to design a game?	experimental study, system logs, interviews
Game managers	How easy is it to manage a game?	experimental study, system logs, interviews
Developers	How easy is it to enhance the system?	experimental study, interviews

Table 1. Overview of SGSEAM

The scope of SGSEAM is to assess serious game frameworks as software infrastructure. While the overall success of a serious game depends on the individual success of all of these stakeholders, SGSEAM only assess the experiences of the players, system admins, game designers, game managers, and developers.

SGSEAM employs the concurrent triangulation strategy [3] of the mixed method of quantitative and qualitative data collection and analysis including instrument and analytical data recorded by the system such as website logs, user interaction database, as well as interviews and questionnaire responses.

Similar to "Goal-Question-Metric" (GQM) approach [2] in software engineering research, The assessment goals in SGSEAM are to identify the strength and weakness of each identified stakeholders. For each stakeholder, a set of questions and alternative assessment approaches are proposed.

Table 1 provides an overview of the assessment method:

There are usually multiple assessment approaches for a specific question. Different assessment approaches will have different levels of rigor. In experimental design terms, rigor refers to external and internal validity. The assessment approaches for a question can be additive. The more approaches applied, the higher confidence of the assessment.

CASE STUDY OF ASSESSING MAKAHIKI USING SGSEAM

This section presents a case study of how we applied SGSEAM to assess Makahiki, an open source serious game framework for sustainability[8]. Makahiki is an serious game framework with the purpose of education and behavioral change regarding energy and water consumption. It can be tailored to the needs of different organizations. Makahiki had been used to create four different Kukui Cup Energy Challenges at the University of Hawaii (UH) in 2011 and 2012, Hawaii Pacific University (HPU) in 2012, and East-West Center (EWC) in 2012. The successful creation of several serious game challenges by different organizations provides evidence that Makahiki framework can be successfully tailored.

Measurement	MIN	AVG	MAX
Participation rate	13%	37%	74%
Number of players per day	43	85	147
Play time per day	1 min	27.7 mins	8.5 hours
submissions per day	32	266	1110
social interactions per day	51	208	468
website errors per day	0	0.6	4

Table 2. Makahiki Engagement Metrics

Besides the real world usage of Makahiki in the series of Kukui Cup challenges, we performed in-lab assessment experiments using SGSEAM. Makahiki was used in a serious game development course in Spring semester of 2013 at the Information and Computer Sciences Department of the University of Hawaii at Manoa. There were a total of 8 students who participated in the experiments. The participants were either senior undergraduates or graduate students majoring in Computer Science. During the course, the students installed Makahiki, configured and designed a serious game instance with Makahiki, and finally developed an enhancement to the Makahiki framework. We asked the students taking the course to voluntarily participate in the assessment experiments of Makahiki.

Player Assessment

To assess the effectiveness of the framework for designing games that improve player literacy in sustainability, we conducted two energy literacy surveys, one before the challenge (pre-game) and one after the challenge (post-game). 24 players completed both surveys. Out of the total 19 energy literacy questions, the average number of questions answered correctly is 7.54 before the challenge, and 8.96 after the challenge. This result indicates an 18% ($p=0.056$) improvement on the energy literacy. We also surveyed non-players as a control condition, and found that their literacy did not change, indicating that the improvement in player literacy was indeed due to the game.

To assess the effectiveness of the framework for designing games that produce positive change in sustainability behaviors, we recorded and analyzed energy consumption data before, during and after the challenge. Before the challenge, an energy usage baseline was established. During the challenge, compared to the baseline, 12 out of the total 20 teams reduced their energy consumption, with the highest reduction of 16.1%. However, 3 teams actually increased their energy consumption, with the highest increase of 11.7%. Overall, the average reduction of the 20 teams was very low—approximately 2%.

To assess player engagement of the game, we calculated a variety of engagement metrics. The results are shown in Table 2:

The average participation rate of this challenge is 37%, which is good compared to other sustainability challenges. Over the course of the challenge, an average player spent about 27.7 minutes per day on the website.

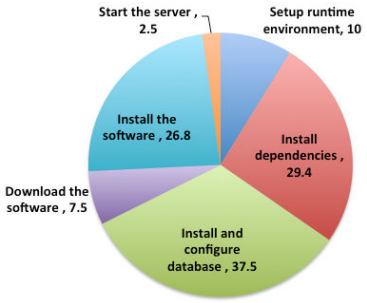

Figure 1. Average time (minutes) for installation steps (n=8)

In summary, SGSEAM indicates that Makahiki can be successful in achieving player engagement and literacy improvement. SGSEAM could not provide evidence of positive change in behavior.

System Admin Assessment

System admin assessment was done using an in-lab experiment. Students in the serious game class were tasked with installing the Makahiki system into their local computers. In order to understand how much time it takes to install Makahiki and what problems might be encountered, we designed a Google Form explaining the steps required to install Makahiki. We asked the students to record the time they spent completing each step and the problems they encountered. We also asked the students to provide feedback about their installation experiences in the form of blog posts. [10] describes in detailed the Google Form that is used in this assessment.

The results from the Google Form responses show that the average total time to successfully install Makahiki was 1.4 hours, with a maximum time of 2 hours and the minimum time of 0.9 hour. Figure 1 shows the average time for each installation step.

We coded and categorized the descriptive problems reported by the students in both the Google Form and their blog posts. We identified that the "Install and configure database" step has the longest average time. It is also has the most participant reported problems. This assessment determines the areas for future improvement are (1) to improve documentation on DB installation, and (2) to improve the install script to automate more installation tasks.

In summary, SGSEAM identified database installation as a weak point in installation. Otherwise, SGSEAM indicates generally positive results regarding Makahiki with respect to installation.

Game Designer Assessment

We also used the in-lab experiment to assess the game designer experience of Makahiki. The students in the experiment was tasked to design a serious game using the Makahiki framework. We asked the students to follow specific design steps and record the time required and any problems encountered during their design process, using a Google Form similar to the one used for the system admin assessment. In addition, students were asked to provide feedback about their design experiences in the form of blog posts. [10] describes in detailed the Google Form that is used in this assessment.

Problem encountered	Number of participants
Difficulty in understanding predicate system and unlock condition	7
A bug that prevented users with usernames containing capital letters from logging in	2
A bug in the processing of Ajax queries	1
Difficulty in generating event attendance codes for game activities	1

Table 3. Makahiki Game Design Analysis, (n=8)

The game designer assessment was generalized into 7 tasks corresponding to distinct types of administrative tasks and game design planning. The time for each task is calculated from the Google Form results. The most time consuming task is "Smart Grid Game Design", which took average 107.9 minutes (56% of total time) to complete, while the least time consuming tasks is "Raffle Game Design", which took average 7.9 minutes (7% of total time) to complete.

We aggregated the problems reported in the feedback of the 8 students that participated in the experiment. Table 3 shows the result of the analysis:

In summary, SGSEAM revealed two shortcomings with Makahiki configuration: "Smart Grid Game Design" and "Configure Challenge Settings". Issues encountered in "Smart Grid Game Design" included 1) difficulty and lack of documentation on the predicate system used to define dependencies between game activities, and 2) difficulty in generating event attendance codes for game activities. Issues encountered in "Configure Challenge Settings" included 1) a bug in the processing of Ajax queries caused by consecutive clicks on the same interface button, and 2) a bug that prevented users with username containing capital letters from logging in.

Game Manager Assessment

We used the 2012 Kukui Cup Challenge at the Hawaii Pacific University (HPU) to assess the game manager experience of Makahiki. We interviewed the game manager of the HPU Kukui Cup challenge, who is also the game designer of the challenge. We asked him about his game management experiences using the Makahiki admin interface.

The interview took place after the challenge and was audio-recorded. We transcribed the audio recording. The data shows that the game management interface was easy for him to use. He also discovered a useful feature in the approval interface without help from the Makahiki support team. The only problem he reported was that after the competition ended, he discovered that some of the analytics data disappeared. This was identified by the Makahiki development team as a software bug and has since been fixed.

In summary, SGSEAM uncovered few problems with Makahiki game management using the interview approach. We realized that the confident level of this assessment ap-

proach is low because of availability of only one data point. An experimental study approach or perform interviews to multiple game managers will increase the confidence level of the assessment.

Developer Assessment

We assessed developer experience using an in-lab experiment. One of the class assignments for the students in the experiment was to develop an enhancement to Makahiki. This involved setting up a development environment, following the tutorial to create a "Hello world" widget using Makahiki, and finally, developing an enhancement to extend the functionality of Makahiki. The students were asked to submit their development source code to the public source code repository and write a blog post to discuss their efforts to complete the development activity.

All 8 students reported that the first task of creating the simple "Hello world" widget was easy, while the enhancement development was hard. Only one student successfully completed all 5 required features, while the rest successfully completed 1 or 2 features. The main problem students reported was the lack of documentation for the development libraries. One student stated in his blog that he decided to choose Makahiki framework to develop his own serious game because of Makahiki's features and possibility of reducing development effort by using the framework.

In summary, SGSEAM reveals significant problems with developer efficiency. Analysis is still ongoing regarding the specific causes of problems and how best to address them.

DISCUSSION

We have developed a serious game framework assessment method called Serious Game Stakeholder Experience Assessment Method (SGSEAM). SGSEAM assesses serious game frameworks from the perspective of the major stakeholders' experiences. These experiences are assessed qualitatively and quantitatively to identify the strengths and shortcomings of a serious game framework. We hope that, by using SGSEAM, developers of a serious game framework will gain insights on the areas to improve on, and produce better serious games or gamified applications.

The results of the SGSEAM assessment case study on Makahiki show both strengths and weaknesses in the framework. The assessment has provided actionable insight into how to improve the framework for system administrators, developers, and game designers. We now understand Makahiki far better than we did before the application of SGSEAM.

Our use of SGSEAM also reveals concerns with the assessment method itself. For certain stakeholders, we took advantage of a course on serious game design to obtain fairly detailed quantitative data about, for example, game design assessment. While we feel confident of these results, the effort required to collect the data was substantial. On the other hand, for other stakeholders such as game managers, we only had access to a single person who could provide insight from that perspective. While easier to collect, the small sample size limits our confident in the data. We are considering ways to augment the method with a "confidence" value that helps others better interpret the findings.

We also realized the needs of applying SGSEAM to other serious game frameworks in order to understand the effectiveness of SGSEAM. It is an ongoing research for us.

ACKNOWLEDGMENTS

This research is supported in part by grant IIS-1017126 from the National Science Foundation.

REFERENCES

1. Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology 45*, 4 (2009), 624–634.
2. Caldiera, V. R. B. G., and Rombach, H. D. The goal question metric approach. *Encyclopedia of software engineering 2* (1994), 528–532.
3. Creswell, J. W. *Research design: qualitative, quantitative, and mixed methods approaches*, 2nd ed. ed. Sage Publications, Thousand Oaks, California, 2003.
4. Deterding, S., Dixon, D., Khaled, R., and Nacke, L. From game design elements to gamefulness: Defining "gamification". In *Proceedings of MindTrek* (2011).
5. Harper, B. D., and Norman, K. L. Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference* (1993), 224–228.
6. Harteveld, C. Triadic game evaluation: A framework for assessing games with a serious purpose. In *Workshop of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (2010).
7. Hurteau, M., Houle, S., and Mongiat, S. How legitimate and justified are judgments in program evaluation? *Evaluation 15*, 3 (2009), 307–319.
8. Johnson, P. M., Xu, Y., Brewer, R. S., Moore, C. A., Lee, G. E., and Connell, A. Makahiki+WattDepot: An open source software stack for next generation energy research and education. In *Proceedings of the 2013 Conference on Information and Communication Technologies for Sustainability (ICT4S)* (February 2013).
9. Mayer, I. Towards a comprehensive methodology for the research and evaluation of serious games. *Procedia Computer Science 15*, 0 (2012), 233 – 247.
10. Xu, Y. Approach to access system admin and game designer experiences in makahiki. Tech. Rep. CSDL-13-04, Department of Information and Computer Sciences, University of Hawaii, Honolulu, Hawaii 96822, June 2013.
11. Zyda, M. From visual simulation to virtual reality to games. *IEEE Computer 38*, 9 (Sep 2005), 25 – 32.