

Serious Game Stakeholder Experience Assessment Method (SGSEAM) User Guide

Yongwen Xu
Collaborative Software Development Laboratory
Department of Information and Computer Sciences
University of Hawai'i at Manoa
yxu@hawaii.edu

September 23, 2013

Contents

1	Introduction	2
2	SGSEAM Overview	2
2.1	Stakeholders	2
2.2	Assessment Methodology	3
2.3	Stakeholder Experience Assessment	3
3	SGSEAM User Guide	4
3.1	Player Assessment	5
3.1.1	Pre-Post effectiveness study	5
3.1.2	Self-reported usability metrics	5
3.1.3	Engagement metrics	6
3.2	System Admin Assessment	6
3.2.1	Post-hoc system admin interview	7
3.2.2	In-lab installation study	7
3.3	Game Designer Assessment	7
3.3.1	Post-hoc game designer interview	7
3.3.2	In-lab game design study	8
3.3.3	Game design log data analysis	8
3.4	Game Manager Assessment	8
3.4.1	Post-hoc game manager interview	8
3.4.2	In-lab game management study	9
3.4.3	Game management log data analysis	9
3.5	Game Developer Assessment	9
3.5.1	Post-hoc game developer interview	9
3.5.2	In-lab game development study	9

1 Introduction

One of the benefits of using a serious 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”?

We found no prior work concerning comprehensive assessment for the particular needs of a serious game framework. Serious Game Stakeholder Experience Assessment Method (SGSEAM) describes a method for assessing serious game frameworks from the stakeholder experience perspectives. 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” [4]. 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.

2 SGSEAM Overview

2.1 Stakeholders

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.
- **Game Developers:** those who use the game framework to customize, extend and enhance their games.
- **Researchers:** those who are conducting research using the game framework.
- **Spectators:** those who do not participate in the game play but are interested in the game and the results of game play.
- **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, providing support for energy data collection if the serious game requires energy data, etc)

- **Funding organizations:** the organizations who provide funding for the game or game framework.

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 game developers, which are closely related to software infrastructure.

The following sections describe the methodology used in SGSEAM, followed by the detailed description of assessment methods for each identified stakeholder.

2.2 Assessment Methodology

Creswell [3] categorizes research methods into three approaches: quantitative, qualitative, and mixed methods, according to what knowledge claims are being made and how knowledge is acquired. Quantitative method reflects a post-positivist paradigm where hypotheses are specified *a priori* and tested by experimental design. Qualitative method reflects a constructivist or participatory paradigm where knowledge would be acquired by observation and open-ended design. SGSEAM employs the mixed methods approach which based on pragmatic knowledge claims and assumption that collecting diverse types of data provides better understanding of the research problem: assessing the strengths and shortcomings of a serious game framework.

In SGSEAM, the concurrent triangulation strategy described in Creswell’s mixed method approach is used. Data collection and analysis involves both quantitative information (instrument and analytical data recorded by the system such as website logs, interaction database, etc), as well as qualitative information (interviews and questionnaire responses).

2.3 Stakeholder Experience Assessment

SGSEAM follows closely with the "Goal-Question-Metric" (GQM) approach [2] in software engineering research. GQM defines a software measurement model on three levels: a goal of the measurement, a set of questions to assess the goal, and a set of metrics associated with each question. There are many metrics related to user experiences [5], SGSEAM focus on the metrics that is useful to provide insights about the strengths and weaknesses of a serious game framework.

In SGSEAM, the assessment goals are the experiences of the identified stakeholders. For each stakeholder, a set of questions is used to assess the strengths and shortcomings from the stakeholder’s perspective. For each question, a set of alternative assessment approaches is proposed.

Table 1 provides an overview of the assessment method:

There are usually multiple assessment approaches for a specific question. The approaches listed above can be generally categorized into in-vivo and in-vitro assessments. The in-vivo approaches, such as pre-post test, in-game surveys and interviews, assess the real world instance of the game. The in-vitro approaches use in-lab experiments in a simulated environment. Different assessment approaches will have different levels of rigor or validity. For example, the in-lab experiments (in-vitro) can enlist several subjects to perform the same pre-defined tasks and collect comparable data in a more controlled setting, while in-game surveys or interviews in the in-vivo approach typically collect data from different settings but the data reflect the real world interaction between the stakeholders and the framework.

Stakeholder	Assessment Goal	Assessment approaches
Players	To what extent does the framework affect players? To what extent does the framework engage players?	Pre-Post effectiveness study, Self-reported usability metrics, Engagement metrics
System admins	How easy is it to install and maintain the system?	Post-hoc system admin interview, In-lab installation study
Game designer	How easy is it to design a game?	Post-hoc game designer interview, In-lab game design study, Game design log data analysis
Game managers	How easy is it to manage a game?	Post-hoc game manager interview, In-lab game management study, Game management log data analysis
Game Developers	How easy is it to enhancing the system?	Post-hoc game developer interview, In-lab game development study

Table 1: Overview of SGSEAM

The details of the individual assessment approach for each stakeholder are described in the following sections. These approaches for each stakeholder can be additive. The more approaches applied, the higher confidence of the assessment for the stakeholder.

3 SGSEAM User Guide

This user guide section describes the three steps in using SGSEAM to assess a serious game framework:

- ***Step 1: Identify all the potential stakeholders of the framework and categorize them into SGSEAM stakeholders.***

For each stakeholder, identify the population, the name and contact if possible. For example, the player stakeholder could be identified as the users interact with the game interface, perform certain tasks given by the interface, or winning the prize if there is one. It is important to be able to contact the stakeholders in some way, either via email or phone, to get the feedback from their experiences with the framework.

- ***Step 2: Identify the important tasks that need to interact with the framework for each stakeholder.***

For example, the players interaction with the game produced by the frameworks; the system admins install, backup, monitor the software system; the game designers create the content for the game, design or decide what game mechanics to used; the game managers manage the game during the game period; and the game developers develop enhancement, customization using the framework.

- ***Step 3: Determine the appropriate assessment approaches for each stakeholder.***

The appropriate assessment approaches should be determined according to the resource available. Sometimes it is impossible or really hard to implement a certain approach. The more approaches applied, the higher confidence of the assessment can be achieved. The following sections describe in detailed the different approaches for each stakeholder. Each assessment approach describes what data to collect, how to collect the data and how to analyze the data to obtain insights about the strengths and weaknesses of the framework from each stakeholder's perspective.

3.1 Player Assessment

The goal of player assessment is to determine the effectiveness of the game framework from player's perspective. It is essential that a game produced by a serious game framework could achieve its intended "serious" purpose. The intended purposes of serious games are always subject specific. For example, the desired effect of a serious game for energy education and conservation is to increase players' energy literacy and reduce their energy consumption during (and, hopefully, after) the game. A serious game for language learning would have a very different desired effect.

3.1.1 Pre-Post effectiveness study

Users of SGSEAM could use domain-specific questions to assess the desired effects of their serious game. For example, the following question could be used to assess a serious game for sustainability education: To what extent does the game increase player's literacy in sustainability?

One approach to assess the question of the effectiveness is a quasi-experimental pre-post study. A set of survey questionnaires can be presented to a random selection of the players before the game (pre-test). After the game ends, the same survey (post-test) is presented to the players who responded the pre-test survey. These two sets of survey response data are compared to understand if the game has had an impact on the survey subjects. The extent of the changes reflected in the survey result could indicate the degree of effectiveness of the serious game for this subject.

Other measurements, for example, the energy consumption data in a energy challenge serious game, could be collected before and after the game to determine the extent of changes that may be caused by the participation in this serious game.

3.1.2 Self-reported usability metrics

Another approach for assess the players' experience is to interview players about their self-reported experience with the game. The interview could be administrated via a face-to-face conversation or through online survey. We found that the online survey is more cost effective than face-to-face conversation. In additions, the online survey could be potentially implemented as part of an activity inside the game, as in the case of the Kukui Cup serious game.[\[1\]](#). Some of the sample interview questions are included in the followings:

Open-ended questions:

- What did you like most about the website/game?
- What did you found confusing?
- What issues did you have while using the site/game?

- What was the thing you liked the least about the site/game?
- What can we do to improve the site/game?

Close-ended questions with Likert scale from "Strong disagree" to "Strongly agree":

- It was easy to find what I was looking for on the website
- The website was responsive
- The website provided adequate help in teaching me how to play
- I understood how to play
- this is something my friends should participate in

3.1.3 Engagement metrics

Player engagement is an important measure for understanding the effectiveness of a serious game. By investigating the degree of engagement, we can determine to what extent individuals are participating in the game, as well as to what extent the community population is participating in the game. On the other hand, engagement has a subtle relationship to the overall effectiveness of a serious game. It is possible for the game to be played by only a subset of the target population, but have an impact on those not playing by virtue of their contacts with players. Gaining better insight into this diffusion effect could be an interesting research area.

To obtain engagement data, SGSEAM analyzes the following measures based upon system log data provided by the framework:

- participation rate
- number of players per day
- play time of a player per day
- submissions of all player per day
- social interaction of all player per day
- website errors per day

The participation rate measures the percentage of users who used the game based on the total eligible players. In the serious game context, it indicates the level of involvement or awareness of the serious matters. The number of players and play time per day measure how frequently the players interact with the game. The submissions per day measures the rate of serious game specific activities (online or real world) that players completed, while the social interaction per day measures the rate of social interactions happened in the game between the players. At last, the website errors per day measures the rate of errors encountered by the players while using the game website. In general, with the opposite of website error measurement, the higher value these measurements are, the higher engagement level the game has.

3.2 System Admin Assessment

System administrators are responsible for installing and maintaining the software infrastructure for the game. Their tasks include the framework and dependency installation, maintain the database, backups, and so forth.

3.2.1 Post-hoc system admin interview

One approach to assess the question of how easy it is to install and maintain the system is a post-hoc interview. The actual system admin(s) are asked about their experience after their installation in the production system. The interview includes the following questions:

- How much time did you require to install the system and the dependencies?
- How much time did you require to maintain the system?
- What problems did you encounter?
- Did you find it difficult to admin the system? What was difficult?

After the interview data is acquired, the assessor will perform qualitative data analysis, which involves transcribing (if the interview data is in audio format), categorizing and coding the description of reported problems or difficulties.

3.2.2 In-lab installation study

Another approach to assess the question is to use an in-lab experimental study. A group of system admins will be asked to install the system, record the time spent and problem encountered as they complete each step. The qualitative data (i.e., the descriptive problems reported by the participants of the study) will need to be categorized and coded. The assessor will triangulate the reported time data and the problem categories to identify the area of strength (less time spent) and weakness (problems and difficulties).

The level of confidence of the above two assessment approaches varies. The experimental study approach is more rigor because of the generality achieved from the larger population of participants under study. The data collected during the step by step experimental study is more accurate than the one collected in the post-hoc interview.

3.3 Game Designer Assessment

A game designer uses the serious game framework to design and create a serious game. A serious game framework always provides certain tools or interfaces to game designers with the hope that these will simplify the design of a game. Such tools might involve configuring global settings for the game, such as how long will the game run, who are the players, and how to design individual game elements.

SGSEAM assesses the game designer stakeholder by addressing the following two questions: (a) How much time is required to design an instance of a serious game using the framework? and (b) How many, and how problematic are the errors that designers encounter during the design process?

There are three approaches for game designer assessment:

3.3.1 Post-hoc game designer interview

One approach is to interview the actual game designer(s) after they had completed the design in a production system. The following questions will be asked:

- How much time did you spend to complete each design task?
- What problems did you encounter?

- Did you find it difficult to configure? What was difficult?
- Did you find it difficult to design a specific game? Which one, and what was difficult?

The interview data will be transcribed (if audio recording), categorized and coded to identify the strengths and weaknesses.

3.3.2 In-lab game design study

Another approach is an in-lab experimental study, where a group of participants is asked to use the system to perform a same set of design tasks. The time spent and problems encountered are recorded for each task. The assessor will triangulate the reported time data and the problem categories to identify the strengths and weaknesses.

3.3.3 Game design log data analysis

A third approach is to collect the system log data related to the game designing tasks. When available, the time spent and error encountered can be queried from the system logs. Although these system generated data might be easier to gather in some systems, it might not provide the same depths or insights than the other two approaches where the experiences are provided by the participants directly. On the other hand, these system data can be supplemental to the other approaches. They could be correlated with the data gathered from the other assessment approaches to increase the confidence of the assessment.

3.4 Game Manager Assessment

A game manager uses the serious game framework to manage the serious game that the game designers created. It is possible that a game manager is also the game designer. Serious game frameworks normally provide certain interfaces for the managers to manage the game. This may involve managing player submissions, monitoring the game state, entering manual resource data, notifying winners of the game, etc.

SGSEAM assesses the game manager stakeholder with the following questions: (a) How much time is required to manage an instance of a serious game using the framework? and (b) How many, and how problematic are the errors that managers encounter during the design process?

Similar to the assessment of game designer experience, SGSEAM proposes three approaches.

3.4.1 Post-hoc game manager interview

The post-hoc interview approach gather data from the game manager(s) by asking the following questions:

- How much time did you spend to complete each managing task?
- What problems did you encounter?
- Did you find it difficult to manage? What was difficult?

3.4.2 In-lab game management study

The experimental study approach gather data from a group of participants about the time spent and problems encountered for each task of managing the serious game.

3.4.3 Game management log data analysis

The log data analysis collects system log data related to the game managing tasks. The time spent and error encountered can be deducted from the system log and reveals strengths and weaknesses of the game managing interface.

3.5 Game Developer Assessment

The game developer stakeholder is different from the game designer stakeholder, in that the game designer stakeholder tailors the framework without requiring any software development, while the game developer stakeholder enhances, corrects, and extends the system by manipulating code.

To investigate how easy it is to understand, extend, and debug a serious game framework from a developer's perspective, SGSEAM assesses how much time it takes to develop an enhancement to the game framework, and how many errors are encountered during the process.

3.5.1 Post-hoc game developer interview

This assessment approach is accomplished by interviewing the actual developer(s) to answer the following questions:

- How much time did you spend developing a customization using the game framework?
- What problem(s) did you encounter?
- Did you find it difficult to understand, extend and debug the system? What was difficult?

3.5.2 In-lab game development study

The experimental study assessment approach asks a group of developers to develop a same set of enhancements to the system, and ask them to record the time spent to develop and problems encountered.

Similarly, the descriptive data will be categorized and coded. The time data will be correlated to the problem data to identify the areas of strength and weakness.

References

- [1] Robert S. Brewer. *Fostering Sustained Energy Behavior Change And Increasing Energy Literacy In A Student Housing Energy Challenge*. PhD thesis, University of Hawaii, Department of Information and Computer Sciences, March 2013.
- [2] Victor R Basili, Gianluigi Caldiera and H Dieter Rombach. The goal question metric approach. *Encyclopedia of software engineering*, 2:528–532, 1994.

- [3] John W. Creswell. *Research design: qualitative, quantitative, and mixed methods approaches*. Sage Publications, Thousand Oaks, California, 2nd ed. edition, 2003.
- [4] Marthe Hurteau, Sylvain Houle, and Stéphanie Mongiat. How legitimate and justified are judgments in program evaluation? *Evaluation*, 15(3):307–319, 2009.
- [5] Thomas Tullis and William Albert. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Morgan Kaufmann, 2010.