Serious Game Stakeholder Experience Assessment Method (SGSEAM) User Guide

Yongwen Xu, Philip M. Johnson, Carleton A. Moore, Robert S. Brewer

Collaborative Software Development Laboratory
Department of Information and Computer Sciences
University of Hawai'i at Manoa
{yxu, johnson, cmoore, rbrewer}@hawaii.edu

October 22, 2013

Contents

1. SGSEAM Overview

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"?

Serious Game Stakeholder Experience Assessment Method (SGSEAM) describes a method for assessing serious game frameworks from the stakeholder experience perspectives. 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 benefits of SGSEAM assessment are for the developers of serious game frameworks to learn and improve from the findings of the assessment.

SGSEAM is 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. 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 to benefit the developers of the framework.

Figure 1.1 outlines the steps of the process of applying SGSEAM to a framework.
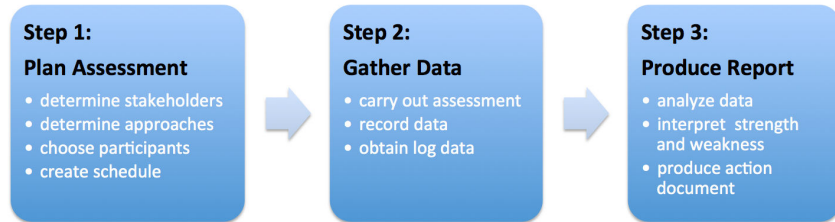

Figure 1.1: Applying SGSEAM to a framework

There are three steps in the process of applying SGSEAM. Step one is to plan the assessment, including identifying the stakeholder and participants and creating the assessment plan. The deliverable for this step is the assessment plan document. Step two is to gather data. The deliverable for this step is the assessment data repository.

Step three is to produce the strength and weakness report. The deliverable for this step is the action document for framework improvement. The following chapters describe the steps in details.

2. Plan Assessment

2.1 Identify stakeholders

> Identify the stakeholders in each SGSEAM stakeholder class, write down their names and roles.

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. 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.

Stakeholder class	Definition	Examples
Players	participate in the game produced by the framework.	students, residents
System admins	install and maintain the technological game infrastructure.	system admin, IT staffs
Game designer	design the content and game mechanics.	instructional designers, content experts
Game managers	manage the game during the period of game play.	sustainability coordinators, residential staffs
Game developers	develop customization, extend and enhance the game.	programmers, internal developers

Table 2.1: SGSEAM Stakeholders

In the full life cycle of a serious game framework there are a great variety of potential stakeholders in addition to the ones listed above. Some other stakeholders are: Researchers who conduct research using the game framework; Spectators who do not participate in the game play but are interested in the game and the results of game play; Funding organizations who provide funding for the development and execution of the game.

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

stakeholders described in Table 2.1, which are closely related to software infrastructure.

2.2 Determine assessment approach

> 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.

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 2.2 provides an overview of the assessment method:

Stakeholder	Assessment Goal	Assessment approaches
Players	How does the framework affect and engage players?	1.Pre-Post effectiveness study 2.Self-reported usability metrics 3.Engagement metrics
System admins	How easy is it to install and maintain the system?	1.Post-hoc admin interview 2.In-lab installation study
Game designer	How easy is it to design a game?	1.Post-hoc designer interview 2.In-lab game design study 3.Design log data analysis
Game managers	How easy is it to manage a game?	1.Post-hoc manager interview 2.In-lab game management study 3.Management log data analysis
Game developers	How easy is it to enhancing the system?	1.Post-hoc developer interview 2.In-lab game development study

Table 2.2: SGSEAM approaches

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.

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.

2.2.1   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 increases players' energy literacy and reduces their energy consumption during (and, hopefully, after) the game. A serious game for language learning would have a very different desired effect.

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 set 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.

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.[**?**]. 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

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.

2.2.2 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.

Post-hoc 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.

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.

2.2.3   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:

Post-hoc 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.

In-lab game design study

Another approach is an in-lab experimental study, where a goup 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 tasks. The assessor will triangulate the reported time data and the problem categories to identify the strengths and weaknesses.

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 confident of the assessment.

2.2.4   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.

Post-hoc manager interview

The post-hoc interview approach gather data from the game manger(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?

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.

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.

2.2.5   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.

Post-hoc 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?

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 during the development.

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.

2.3   choose participants

Choose participants from each stakeholder class.

2.4   create assessment schedule

Create a schedule for each assessment.

3. Gather Data

3.1 carry out the assessment

Carry out the assessment.

3.2 record data

3.3 refine assessment plan

3.4 obtain log data

4. Produce Strength and Weakness report

4.1 analyze data

4.2 interpret strength and weakness of framework

4.3 produce artifact action document