MAKAHIKI AND SGSEAM: A SERIOUS GAME FRAMEWORK FOR SUSTAINABILITY AND STAKEHOLDER EXPERIENCE ASSESSMENT METHOD

A DISSERTATION SUBMITTED TO THE
GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI‘I AT MĀNOA
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

COMPUTER SCIENCE

AUGUST 2014

By

Yongwen Xu

Dissertation Committee:

Philip M. Johnson, Chairperson
David Chin
Scott Robertson
Lipyeow Lim
Daniel Port

TABLE OF CONTENTS

1 SGSEAM Design 1

1.1 Overview 1

1.2 Methodology 3

1.3 Plan the Assessment 4

1.3.1 Identify stakeholders 4

1.3.2 Determine assessment approach 5

1.3.2.1 Player Assessment 6

1.3.2.1.1 Pre-Post effectiveness study 6

1.3.2.1.2 Self-reported usability metrics 7

1.3.2.1.3 Engagement metrics 7

1.3.2.2 System admin assessment 9

1.3.2.2.1 Post-hoc admin interview 9

1.3.2.2.2 In-lab system admin study 10

1.3.2.3 Game designer assessment 10

1.3.2.3.1 Post-hoc designer interview 11

1.3.2.3.2 In-lab game design study 11

1.3.2.3.3 Game design log data analysis 12

1.3.2.4 Game manager assessment 12

1.3.2.4.1 Post-hoc manager interview 13

1.3.2.4.2 In-lab game management study 13

1.3.2.4.3 Game management log data analysis 13

1.3.2.5 Game developer assessment 14
1.3.2.5.1 Post-hoc developer interview 14
1.3.2.5.2 In-lab game development study 14
1.3.3 Choose participants 15
1.3.4 Create assessment schedule 15
1.4 Gather Data 16
1.4.1 Carry out the assessment 16
1.4.2 Obtain log data 16
1.5 Produce Assessment Report 16
1.5.1 Analyze Data 17
1.5.2 Determine strength and weakness 17
1.5.3 Produce Report with Actionable Steps 17
A SGSEAM Assessment Guide for Lucid BuildingOS and BuildingDashboard 18
A.1 Overview 18
A.2 Step 1: Plan the Assessment 19
A.2.1 Identify Stakeholders 19
A.2.2 Determine Assessment Approach 20
A.2.2.1 Assess Player Stakeholder Experience 21
A.2.2.2 Assess System Admin Stakeholder Experience 23
A.2.2.3 Assess Game Designer Stakeholder Experience 24
A.2.2.4 Assess Game Manager Stakeholder Experience 25
A.2.2.5 Assess Game Developer Stakeholder Experience 26
A.2.3 Choose Assessment Participants 28

A.2.4 Create Assessment Schedule 28
A.3 Step 2: Gather Data 29
A.3.1 Carry Out Assessments 29
A.3.2 Obtain log data 29
A.4 Step 3: Produce Assessment Report 29
A.4.1 Analyze Data 30
A.4.2 Determine Strength and Weakness 30
A.4.3 Produce Report with Actionable Steps 30
Bibliography **31**

CHAPTER 1
SGSEAM DESIGN

This chapter describes the design of the Serious Game Stakeholder Experience Assessment Method (SGSEAM) for assessing serious game frameworks. It starts with the overview of SGSEAM, followed by the discussion of assessment methodology, and the detailed steps of the assessment method. Finally, Appendix A illustrates an example of SGSEAM assessment guide written for a specific serious game framework.

1.1 Overview

One of the benefits of using a serious game framework such as Makahiki, 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"? As we discussed in the Related Work, there are a few framework or tools for general purpose assessment of a software system, but I have not yet found any prior work concerning the comprehensive approach for the particular needs of a serious game framework assessment. This is the motivation of the Serious Game Stakeholder Experience Assessment Method (SGSEAM). It is designed for assessing serious game frameworks in particular.

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.

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.

Figure 1.1 illustrates the steps of applying SGSEAM to a framework.

There are three steps in the process of applying SGSEAM:

1. Step one is to **Plan the assessment**, including identifying the stakeholders, determining assessment approaches, and creating the assessment schedule. The deliverable for this step is the ***assessment plan*** document.

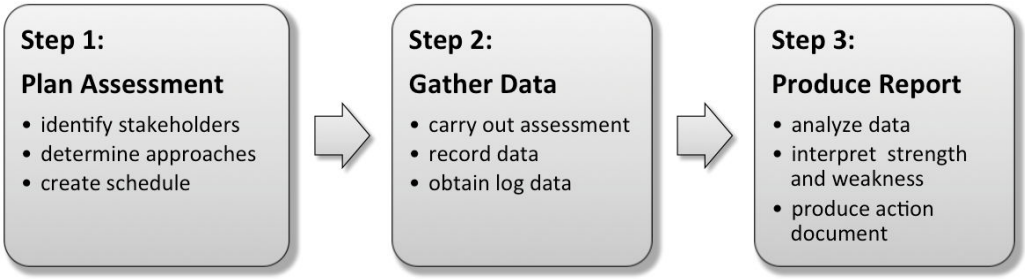

Figure 1.1: Applying SGSEAM to a framework

2. Step two is to **Gather data** by carrying out the assessment, recording and obtaining related data. The deliverable for this step is the assessment ***data repository***.

3. Step three is to **Produce the assessment report** by analyzing the data and interpreting strengths and weaknesses. The deliverable for this step is the ***improvement action*** document.

The following sections describe the methodology used in SGSEAM, followed by the detailed description of steps in applying SGSEAM to serious game frameworks.

1.2 Methodology

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.

Creswell [2] 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 assump-

tion 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).

SGSEAM follows closely with the "Goal-Question-Metric" (GQM) approach [1] 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 [3], 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 described.

1.3 Plan the Assessment

This is the first step of SGSEAM. It first identifies the stakeholders, determines the appropriate assessment approaches according to the available resources, and creates the assessment schedule. The deliverable for this step is the assessment plan document which includes the details of stakeholders, approaches and schedule.

1.3.1 Identify stakeholders

SGSEAM assesses the experiences for the stakeholders listed in Table 1.1.

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. The system admins install, backup, monitor the software system. The game designers create the content for the game and design what game mechanics to used. The game managers manage the game during the game period. Finally the game

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

Table 1.1: SGSEAM Stakeholders

developers develop enhancement and customization using APIs provided by the framework.

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.

1.3.2 Determine assessment approach

There are usually multiple assessment approaches for each stakeholder. Table 1.2 provides an overview of the assessment method and the approaches. The appropriate assessment approaches should be determined according to the resource available. The approaches for a stakeholder is additive. The more approaches applied, the higher confidence of the assessment can be achieved.

The assessment approaches is categorized into in-vivo and in-vitro assessments. The in-vivo approaches, such as pre-post test, in-game surveys and post-hoc 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. It is rigor because of the generality achieved from the larger population of participants under study. On the other hand, in-game surveys or interviews in the in-vivo approach typically collect data from different uncontrolled settings with less rigor. But the in-invo data reflect the real world interaction between the stakeholders and the framework, thus provides better insights in the real world settings.

The following sections describe in detailed the different approaches for each stakeholder. Each

Stakeholder	Assessment goal	Assessment approaches
Player	Determine the extent the framework affect and engage players.	Pre-post effectiveness study(1.3.2.1.1); Self-reported usability metrics(1.3.2.1.2); Engagement metrics(1.3.2.1.3)
System admin	Determine strengths and weaknesses in system install and maintenance.	Post-hoc admin interview(1.3.2.2.1); In-lab system admin study(1.3.2.2.2)
Game designer	Determine strengths and weaknesses in facilitating the game design process.	Post-hoc designer interview(1.3.2.3.1); Game design log data analysis(1.3.2.3.3); In-lab game design study(1.3.2.3.2)
Game manager	Determine strengths and weaknesses in managing the game.	Post-hoc manager interview(1.3.2.4.1); Game management log data analysis(1.3.2.4.3); In-lab game management study(1.3.2.4.2)
Game developer	Determine strengths and weaknesses in developing system enhancement.	Post-hoc developer interview(1.3.2.5.1); In-lab game development study(1.3.2.5.2)

Table 1.2: SGSEAM approaches

assessment approach describes the goal of the assessment, 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.

1.3.2.1 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.

1.3.2.1.1 Pre-Post effectiveness study

We use a quasi-experimental pre-post study to assess the question of the effectiveness of a serious game framework.

This approach requires users of SGSEAM to first determine a set of domain-specific questions to

assess the desired effects of their serious game. For example, a set of questionnaires on sustainability literacy, such as knowledge of power and energy, is used to assess the effectiveness of a serious game for sustainability education.

Once the domain-specific questionnaires are determined and designed, present this questionnaires as a survey to a random selection of the players before the game starts. After the game ends, present the same survey to the same players again. Compare the two set of survey response data to study if the game has an impact on the players regarding to the survey subjects. The extent of the changes reflected in the survey result indicates the degree of effectiveness of the serious game for this subject.

Serious games often engage players with resources of various types (energy, water, waste, etc.). Collect these measurements before, during, and after the game in order to acquire evidence regarding the potential impact upon player use of these resources.

1.3.2.1.2 Self-reported usability metrics

This approach interviews players about their self-reported experiences with the game. Administrate the interview through online survey or face-to-face conversation, although we found that online survey is more cost effective than face-to-face conversation. If possible, implement the online survey as an activity inside the game. For example, the Makahiki serious game framework implements an online survey activity which incentivizes players to complete the survey by rewarding game points for the activity.

SGSEAM proposes to use the usability questionnaires outlined in Figure 1.2 in the online survey or face-to-dace interview:

1.3.2.1.3 Engagement metrics

This approach calculates the engagement metrics to assess the extent of engagement from players and the impact of the game. The more engaging the game is, the more potential impact could be to the players.

1. What did you like most about the game?
2. What did you found confusing?
3. What issues did you have while using the game?
4. What was the thing you liked the least about the game?
5. What can we do to improve the game?
6. It was easy to find what I was looking for on the website.
   Strongly disagree - Disagree - Neutral - Agree - Strongly agree
7. The website was responsive.
   Strongly disagree - Disagree - Neutral - Agree - Strongly agree
8. The website provided adequate help in teaching me how to play.
   Strongly disagree - Disagree - Neutral - Agree - Strongly agree
9. I understood how to play.
   Strongly disagree - Disagree - Neutral - Agree - Strongly agree
10. this is something my friends should participate in.
    Strongly disagree - Disagree - Neutral - Agree - Strongly agree

Figure 1.2: Player self-reported usability metrics questionnaires

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.

SGSEAM proposes to calculate as many as possible the player engagement metrics described in Figure 1.3 by analyzing the data from system log or other channels provided by the framework. The more metrics obtained, the better understanding of the extent of player engagement.

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

Metric	Definition	Mesure
participation	percentage of players who play the game	the level of involvement from players
player	number of players per day	the frequency of players interact with the game
play time	play time of a player per day	the frequency of players interact with the game
submission	submissions of all player per day	the rate of players' completion of game activities
social interaction	social interaction of all player per day	the rate of in-game social interactions between players
game error	game errors per day	the rate of errors encountered by players during the game

Figure 1.3: Player engagement metrics

per day measures the rate of errors encountered by the players while using the game website.

With the exception of the game error metric, the higher value these metrics are, the higher engagement level the game has.

1.3.2.2 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. The goal of system admin assessment is to determine to what extent the framework facilitates the system administration tasks from system admin's perspective. SGSEAM assesses how much time is required to install and maintain an instance of a serious game using the framework and the problems encountered during the system admin process.

SGSEAM proposes two assessment approaches.

1.3.2.2.1 Post-hoc admin interview

This approach assesses the system admin's experience using the post-hoc interview. The system admins are asked about their experience with the framework after they completed the installation and maintanence in the production system. The interview questions are described in Figure 1.4.

The interview should be tape-recorded. Once the interview is completed, qualitative data anal-

1. How much time did you require to install the system and the dependencies?
2. What problems did you encounter when installing the system and the dependencies?
3. How much time did you require to maintain the system?
4. What problems did you encounter when maintaining the system?
5. Did you find it difficult to admin the system? What was difficult?

Figure 1.4: System admin interview questionnaires

ysis is performed against the interview data by doing: (1) transcribing the recordings; (2) coding (categorizing) the time and problems or difficulties encountered. These data reveal the strengths, weaknesses and the areas of improvement for the framework.

1.3.2.2.2 In-lab system admin study

This approach assesses the system admin's experience using the in-lab experimental study. First identify a group of participants who have some levels of system administration experience. Second, provide instructions on each installation steps, ask the participants to install the system according to the instructions, and ask them to record the time spent and problems encountered as they complete each step.

Once the experiment data is collected, categorize the reported problems and correlated with the reported time data to identify the areas of strength (less time spent) and weakness (more time spent and problems or 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.

1.3.2.3 Game designer assessment

A game designer uses the serious game framework to design and create a serious game. A serious game framework normally provides tools or interfaces for game designers to facilitate the design of a game. For example, the framework provides interface to configure the game period, set up players,

and tools to design individual game elements.

The goal of SGSEAM game designer assessment is to determine the strengths and weaknesses of the framework regarding to the game design process. 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 two approaches for game designer assessment:

1.3.2.3.1 Post-hoc designer interview

This approach interviews the game designer(s) after they had completed the design of a serious game using the framework in a production system. The interview includes the questions described in Figure 1.5.

1. How much time did you spend to complete each design task?
2. What problems did you encounter?
3. Did you find it difficult to configure? What was difficult?
4. Did you find it difficult to design a specific game? Which one, and what was difficult?

Figure 1.5: Game designer interview questionnaires

The interview should be tape-recorded. After the interview, transcribed the recordings, code and categorize the reported time and problems to identify the strengths and weaknesses.

In addition, if possible, collect the system log data related to the game designing tasks, analyze the logs to find out the time spent and error encountered during the game designing tasks. Use the log data to verify the findings from the interview data.

1.3.2.3.2 In-lab game design study

This approach assesses the game designer experience using the in-lab experimental study. First identify a group of participants who are somewhat familiar with the subject domain of the game. Second, provide instructions on each designing steps, ask the participants to design the game according to the instructions, ask them to record the time spent and problems encountered as they

complete each step.

Once the experiment data is collected, categorize the reported problems and correlated with the reported time data to identify the areas of strength (less time spent) and weakness (more time spent and problems or difficulties).

1.3.2.3.3 Game design log data analysis

This approach collects 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.

1.3.2.4 Game manager assessment

A game manager uses the serious game framework interface to manage the serious game that the game designers created. It is possible that a game manager is also the game designer. The examples of game management tasks includes managing player submissions, monitoring the game state, entering manual resource data, notifying winners of the game, etc.

The goal of SGSEAM game manager assessment is to determine the strengths and weakness of the framework regarding to the game management process. Similar to the assessment of the game designer, SGSEAM assesses the game manager stakeholder on the time it required to manage an instance of a serious game using the framework and the problems encountered during the managing process.

SGSEAM proposes two approaches for assessment game manager's experience.

1.3.2.4.1 Post-hoc manager interview

This approach interviews the game manager(s) after they had managed a serious game using the framework in a production envrionment. The interview questions are described in Figure 1.6.

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

Figure 1.6: Game manager interview questionnaires

The interview should be tape-recorded. After the interview, transcribed the recordings, code and categorize the reported time and problems to identify the strengths and weaknesses.

In addition, if possible, collect the system log data related to the game managing tasks, analyze the logs to find out the time spent and error encountered during the game managing tasks. Use the log data to verify the findings from the interview data.

1.3.2.4.2 In-lab game management study

This approach assess the game manager's experience using the in-lab game management study. First identify a group of participants who are somewhat familiar with the subject domain of the game. Second, provide instructions on each managing tasks, ask the participants to complete the tasks following the instructions, ask them to record the time spent and problems encountered as they complete each task.

Once the experiment data is collected, categorize the reported problems and correlated with the reported time data to identify the areas of strength (less time spent) and weakness (more time spent and problems or difficulties).

1.3.2.4.3 Game management log data analysis

This approach collects and analyzes the 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.

1.3.2.5 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 development process.

1.3.2.5.1 Post-hoc developer interview

This approach interviews the game developer(s) to assess their experiences of developing the game using the framework. The interview questions are described in Figure 1.7.

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

Figure 1.7: Game developer interview questionnaires

1.3.2.5.2 In-lab game development study

This approach assess the game developer's experience using the in-lab game development study. First identify the general development skills that the framework requires, such as the programming language. Second, identify a group of participants who have some levels of the required development skills. Third, provide requirement specification or instructions on how to develop a new enhancement to the system, ask the participants to complete the task, record the time spent and problems encountered as they works on the task.

Once the experiment data is collected, categorize the reported problems and correlated with the reported time data to identify the areas of strength (less time spent) and weakness (more time spent and problems or difficulties).

1.3.3 Choose participants

Once the assessment approaches are determined for each stakeholder class, the next step is to choose participants. Identify the people from the each stakeholder class that may be willing to participate in the assessment, contact them and get consents for their participation.

For example, in the case of pre-post effectiveness study approach for player assessment, this step randomly chooses a group of players and present a consent form before the online survey. In the case of post-hoc game designer interview approach, the game designer of a real world game instance of the framework should be identified, contacted and consent for the participation in the assessment. When the in-lab game development experiment study is chosen, a group of game developers that meet the required development skills of the framework should be identified and contacted.

1.3.4 Create assessment schedule

Once we decide what the assessment approaches and who the participants are, the next step is to create the assessment schedule. The document should include the detailed assessment plan for each stakeholder class.

Depends on the assessment approach, the actual tasks of the assessment are different. The player pre-post effectiveness study requires the administration of online survey before and after the game. The game designer post-hoc interviews requires administration of interviews to the real world game designers of a production system. Figure 1.8 shows an example of the assessment schedule broken down in the tasks in the plan document.

Game design assessment approach: in-lab experiment study		
Task	**Estimated Start date**	**Estimated End date**
Design the in-lab experiment instruction		
Ask participants to follow the instruction		
Collect response data from participants		
Obtain log data		
Analyze the data		
Interpret strength and weakness		
Produce action document		

Figure 1.8: Assessment schedule in the plan document

1.4 Gather Data

Once the plan has been finalized, the next step is to carry out the assessment, record the data, obtain log data, and (if necessary) refine the assessment plan. The output of this step is a data repository contains all the assessment data that can be analyzed in the next step.

1.4.1 Carry out the assessment

For each assessment approach, complete the tasks outlined in the assessment plan, gather the data when carrying out the assessment. In the case of game designer post-hoc interview approach, record the interview and take notes if necessary. Store all the data into a central data repository.

In the example of the in-lab game design experiment study, a google form is designed to give detailed step by step instructions for the participants to design games using the framework. Participants are asked to record the time they spent completing each step and the problems they encountered. They are also asked to provide feedback about their design experiences in the form of blog posts.

1.4.2 Obtain log data

Talk to the technical staffs of the framework to find out what kind of log data is available. Obtain the log data in a format that is easy to analyze. For example, if the log data is in a database table, ask for the access to the table, or the CSV export of the table data. If the log data is in a log file, ask for the access to the file. Store the log data into the central data repository.

1.5 Produce Assessment Report

In this step, we will analyze the data gathered from previous steps, create an analysis of the strengths and weakness of the framework, and produce an action report with our recommendations as to framework improvements.

1.5.1 Analyze Data

This step performs the data analysis from the data repository obtained from the previous step. For game designer assessment, perform queries from user interaction log data to find out the completion time for a certain user interaction task, for instance, the time for a game designer to complete the configuration of global game settings. For player assessment, calculate the engagement metrics from the game log. For post-hoc interview assessment approach, first transcribe the interview recording into text, code and categorize the responses from the interview questions.

In the example of the in-lab game design experiment study described previously, the assessment data is generalized into 7 tasks corresponding to distinct types of game design tasks. The time for each task is calculated from the Google form responses. The problems reported from the participants are coded and aggregated into the the problems areas.

1.5.2 Determine strength and weakness

This step determines the most important problem areas from our data and summarize them, as well as the areas where the framework appears to be most successful.

In the example of the in-lab design experiment study, there may be a problem area that had been reported by the most numbers of participants, and this problem happened in one of the tasks that took the longest time to complete, we could identify a weakness area of the framework from the perspective of game designer. If there were no problem reported in some game design tasks and the time to complete is short, we could consider those areas are the strengths of the framework.

1.5.3 Produce Report with Actionable Steps

Once the strengths and weaknesses of the framework are identified, an action report should be produced. This report includes the weakness areas that can be improved and actionable steps on how to improve from each stakeholder's perspective. It also includes strengths that the framework needs to maintain.

By producing the report with actionable steps to improve the framework, the SGSEAM assessment is completed.

APPENDIX A
SGSEAM ASSESSMENT GUIDE FOR LUCID BUILDINGOS AND BUILDINGDASHBOARD

This appendix includes the SGSEAM assessment guide written specifically for BuildingOS and BuildingDashboard, with the intension of administrating SGSEAM assessment to the Lucid Design Group's serious game framework. Although the assessment did not actually started, this guide illustrates an example for SGSEAM assessment to a similar serious game framework.

A.1 Overview

This document describes how to assess the Lucid BuildingOS and BuildingDashboard using the Serious Game Stakeholder Experience Assessment Method (SGSEAM).

The goal of this assessment is to identify the major strengths and shortcomings of the software framework using the perspectives of major stakeholders.

The cost of this assessment to Lucid is the requirement for various stakeholders to be available to me for approximately one 30 minute interview.

The benefit of this assessment is the identification of actionable improvements to Lucid BuildingOS and BuildingDashboard.

The SGSEAM assessment method is being developed as part of my Ph.D. research at the University of Hawaii. The assessment of Lucid BuildingOS and BuildingDashboard will help me to identify strengths and weaknesses in SGSEAM. All data about the LucidBuildingOS or BuildingDashboard systems revealed through this assessment will be kept confidential and will not be presented in my research findings.

Table A.1 outlines the steps of the process of applying SGSEAM to a framework.

1. Step one is to **Plan the assessment**, including identifying the stakeholders, determining assessment approaches, and creating the assessment schedule. The deliverable for this step is the ***assessment plan*** document.

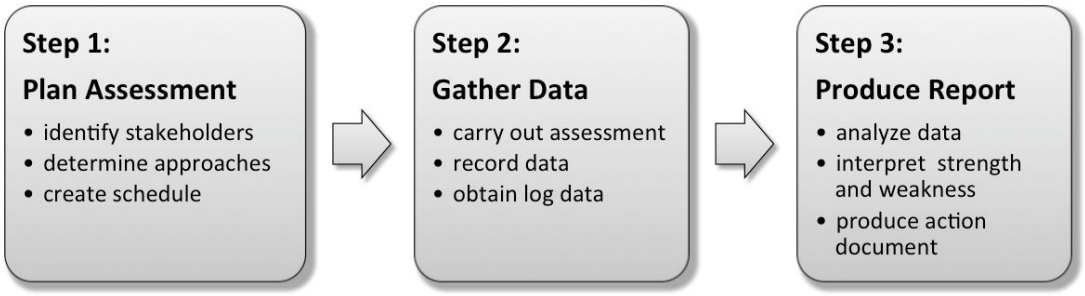

Table A.1: Applying SGSEAM to a framework

2. Step two is to **Gather data** by carrying out the assessment, recording and obtaining related data. The deliverable for this step is the assessment ***data repository***.

3. Step three is to **Produce the assessment report** by analyzing the data and interpreting strengths and weaknesses. The deliverable for this step is the ***improvement action*** document.

The following chapters describe the steps in detail. The Appendix provides additional background material. Each chapter concludes with an "Action Item" shade box, which indicates what you need to do. For example:

> **Action Item:** Read the next three chapters of this document, and determine if this proposed evaluation is feasible. If you identify obstacles, please note them in the spreadsheet so that we can discuss them in an upcoming phone call.

A.2 Step 1: Plan the Assessment

A.2.1 Identify Stakeholders

The first step is to identify the SGSEAM stakeholders and their tasks for the Lucid BuildingOS and BuildingDashboard framework.

According to Campus Conservation Nationals (CCN) Competition Planning Guide, a Competition Organizing Team (COT) will plan and execute the competition. Besides being residents of buildings participating in the competition, they are also users and stakeholders of the BuildingOS framework.

SGSEAM Stakeholder	BuildingOS Users	Tasks	Comments on stakeholder definitions
Player	Building resident	Use BuildingDashboard to view building data and participate in the competition.	
System admin	Internal system admin or developer	Install software, backup, patch, monitor and scale the system.	
Game designer	Behavior Change Manager, Technical Manager, Competition Director, Research Manager	• Decide on a competition format/structure • Set up buildings, meters, and competition in BuildingOS	
Game manager	Technical Manager, Marketing Manager, Building Captain Manager, Events Manager	• Collect, verify baseline and competition data, enter into BuildingOS • Kick-off and other events • Coordinate competition prizes • Manage social media • Monitor competition status	
Game developer	Internal or external Developer	• Develop interface to support other meters • Customize dashboard interface	

Table A.2: BuildingOS Stakeholders

We have converted COT roles into SGSEAM stakeholders and identified their tasks related to BuildingOS and BuildingDashboard, as shown in Table A.2.

Action Item: Review the "Stakeholders" tab in the attached spreadsheet, and provide comments if you believe the set of stakeholders or the mapping needs modification.

A.2.2 Determine Assessment Approach

There are several possible assessment approaches for each stakeholder. Different assessment approaches have different levels of rigor which impacts upon the quality the assessment result. They also require different levels of implementation costs or efforts. Table 1.2 describes the SGSEAM assessment approaches we have developed for each stakeholder category.

While an in-lab experiment has the most rigor, we believe it is too expensive for this assessment. We therefore recommend an interview approach for all stakeholders except players. Table A.3 shows the approaches we recommend for each stakeholder in the case of Lucid BuildingOS and BuildingDashboard.

The following sections describe in detailed the assessment approaches for each stakeholder. description of the recommended approaches.

Action Item: Review the "Approach" tab in the attached spreadsheet. Provide a comment if you believe an approach should be modified, deleted, or added.

Stakeholder	Assessment Approaches	Expected Outcomes	Comments on approach
Player	Pre-post effectiveness study	Determine effectiveness in resource usage reduction.	
	Usability survey	Identify problem areas in game interface	
	Engagement metrics	Determine the extent of engagement	
System admin	Post-hoc admin interview	Identify strengths and weaknesses in the installation and maintenance process.	
Game designer	Post-hoc designer interview	Determine strengths and weaknesses in the game design interface.	
Game manager	Post-hoc manager interview	Determine strengths and weaknesses in the game managing interface.	
Game developer	Post-hoc developer interview	Determine strengths and weaknesses in developing enhancement.	

Table A.3: BuildingOS Assessment Approaches

A.2.2.1 Assess Player Stakeholder Experience

The goal of player assessment is to determine the effectiveness of the game framework from player's perspective as well as the usability of the game interface and the engagement level of the game.

We recommend three approaches for player assessment: pre-post effectiveness, usability survey and engagement metrics. The attached spreadsheet outlines the planned goals and survey questionnaires for these assessment approaches, as shown in Table A.4.

The first approach is **Pre-post Effectiveness Study**. One of the goals of the competition is (but not limited to) the reduction of resource such as energy and water consumption. To assess the effectiveness of this goal, we need to determine the metrics that may be measured before and after the competition (pre-post). Lucid BuildingOS and Dashboard calculates the percentage of reduction of energy and water consumption for each participated building, based on the baseline usage of the previous two weeks. We will use this metrics to measure the effect of the competition. The maximum, minimum and average percentage of reduction of all the buildings are calculated to determine the most, the least and average reduction of the resource usage. This assessment approach reveals the extend of effectiveness of the game produced by the framework, regarding to the resource consumption reduction.

The second approach is **Usability Survey**. We will conduct a player usability survey at the final week or right after the competition to understand the strengths and weaknesses of the game user interface perceived by players. Minimum of 20 players (the more the better) are randomly

1. Player Pre-Post Effectiveness Goals		
Category	Goals	Comments
Education	50% of players can name three things to save electricity	
Behavior Change	50% of players did one thing that reduces their consumption	
	20% of players made at least 1 conservation commitment	
Resource Reduction	3% of average reduction	
2. Player Usability Survey Questionaire		
Question	Response type	Comments
1. What did you like most about the game?	short answer	
2. What did you found confusing?	short answer	
3. What issues did you have while using the game?	short answer	
4. What was the thing you liked the least about the game?	short answer	
5. What can we do to improve the game?	short answer	
6. It was easy to find what I was looking for on the website.	Strongly disagree - Disagree - Neutral - Agree - Strongly agree	
7. The website was responsive.	Strongly disagree - Disagree - Neutral - Agree - Strongly agree	
8. The website provided adequate help in teaching me how to play.	Strongly disagree - Disagree - Neutral - Agree - Strongly agree	
9. I understood how to play.	Strongly disagree - Disagree - Neutral - Agree - Strongly agree	
10. this is something my friends should participate in.	Strongly disagree - Disagree - Neutral - Agree - Strongly agree	
3. Player Engagement Metrics Goals		
Metric	Goals	Comments
participation (percentage of players who participated in the game)	30%	
daily player (average percentage of players per day)	5%	
daily play time (average play time of a player per day)	5 minutes	
submission (average submissions of a player)	2%	
social interaction (average social interaction of a player)	2%	
game error (percentage of players who encountered errors)	1%	

Table A.4: Player Assessment

selected to participate in this survey. The survey is administrated online via survey monkey or other survey tools. We design the survey questionnaire as shown in the section 2 of Table A.4. Once the survey is created online, the survey administrator will email the selected players with the link and instruction to the online survey. After we received all the survey responses, we will code and analyze the response to understand the areas of usability problems in the game interface as well as the areas of strengths. This assessment approach reveals the strengths and weaknesses of the framework regarding the usability of the game interface.

Finally, the third approach is **Engagement Metrics**. This approach calculates the engagement metrics to assess the extent of engagement from players and the impact of the game. The more engaging the game is, the more potential impact could be to the players. We will first obtain the detailed logs of user interaction with the game. These logging includes http web server logs and

user action logs which identify every user click on the web page. Once the log data are available, we will calculate the engagement metrics as described in section 3 of Table A.4. With the exception of the game error metric, the higher value these metrics are, the higher engagement level the game has. Distribution of the above metrics across of the period of the competition also provides insights on the extent of engagement in different time of the competition. For example, it may be typical that the first few days of the competition may have higher number of player and play time metrics because of the launch, or due to the announcement of an interesting real-world event. This assessment reveals the extent of engagement of the players in the game.

> **Action Item:** Review the Player tab in the attached spreadsheet. Provide comments for any Player assessment items that you believe might need to be changed.

A.2.2.2 Assess System Admin Stakeholder Experience

The goal of system admin assessment is to determine to what extent the framework facilitates the system administration tasks from system admin's perspective. SGSEAM assesses how much time is required to install and maintain an instance of a serious game using the framework and the problems encountered during the system admin process.

We consider the tasks of system admin interacting with Lucid's framework are:

1. install the software
2. configure smart meter connectivity
3. backup data
4. monitor performance
5. scaling the system
6. patching

We recommend the post-hoc interview approach for system admin assessment. Once we identify the contact information of the system admins, the interview will be administrated by using an online questionnaire form followed by an optional phone interview if needed. We design the interview with the following questionnaire that is tailored to the specific tasks of the system admins of Lucid's framework. The attached spreadsheet outlines the planned interview questionnaires, as shown in

Table A.5.

System admin Post-hoc Interview Questionnaires	Comments on question
1. How much time did you spend to install the system and the dependencies?	
2. How much time did you spend to configure the meters?	
3. How much time did you spend to maintain the system such as backup, patching, monitoring?	
4. Did you need to scale the system? if Yes, how much time did you spend?	
5. What problems did you encounter?	
6. Did you find it difficult to admin the system? What was difficult?	
7. Can we call you for a short phone interview if we have more questions regarding your experience with the system?	

Table A.5: System Admin Assessment

Once we receive the responses from the system admin, we will code (categorize) the time and problems encountered to find out what are the problem areas if there is any. if we need further explanation to the response, we will administrate a quick phone interview to address the specific response.

These assessment reveals the strengths, weaknesses and the areas of improvement regarding the system admin process for the framework.

Action Item: Review the System Admin tab in the attached spreadsheet. Provide comments for any System Admin interview questionnaires that you believe might need to be changed.

A.2.2.3 Assess Game Designer Stakeholder Experience

The goal of SGSEAM game designer assessment is to determine the strengths and weaknesses of the framework regarding to the game design process. SGSEAM assesses how much time is required to design an instance of a serious game using the framework and the problems encountered during the design process.

We consider the tasks of game designer interacting with Lucid's framework are:

1. decide competition period
2. set up building occupancy, manual or automated meters
3. decide baseline period
4. monitor competition status during the competition

We recommend the post-hoc game designer interview approach for assessing game designer

stakeholder experiences. The interview is administrated by using an online questionnaire form followed by an optional phone interview if needed. We will interview several game designers of different competitions. The more data we collect, the more insights we get. The interview is designed with the following questionnaire that is tailored to the specific tasks of the game designers of Lucid's framework. The attached spreadsheet outlines the planned interview questionnaires for game designer, as shown in Table A.6.

Game Designer Post-hoc Interview Questionnaires	Comments on question
1. How much time did you spend to set up the buildings including meters?	
2. How much time did you spend to setup the competition (competition periods, baseline period, participants)?	
3. How much time did you spend to setup the homepage by deciding which widgets to include?	
4. How much time did you spend to monitor analytical data to understand the state of the game	
5. What problems did you encounter?	
6. Did you find it difficult to use the interface? What was difficult?	
7. Do you agree for us to call you for a short phone interview if we have more questions regarding your experience with the system?	

Table A.6: Game Designer Assessment

After the interview, code and categorize the reported time and problems to identify the strengths and weaknesses. In addition, if possible, collect the system log data related to the game designing tasks, analyze the logs to find out the time spent and error encountered during the game designing tasks. Use the log data to verify the findings from the interview data.

These assessment reveals the strengths, weaknesses and the areas of improvement regarding the game design process for the framework.

> **Action Item:** Review the Game Designer tab in the attached spreadsheet. Provide comments for any Game Designer interview questionnaires that you believe might need to be changed.

A.2.2.4 Assess Game Manager Stakeholder Experience

The goal of SGSEAM game manager assessment is to determine the strengths and weakness of the framework regarding to the game management process. Similar to the assessment of the game designer, SGSEAM assesses how much time it is required to manage an instance of a serious game using the framework and the problems encountered during the managing process.

We consider the tasks of game manager interacting with Lucid's framework are:

1. input data manually
2. manage events, marketing, handing out prizes
3. monitor competition status

We recommend the post-hoc interview approach for game manager assessment. The interview is administrated by using an online questionnaire form followed by an optional phone interview if needed. We will interview several game managers of different competitions. The more data we collect, the more insights we get. The interview is designed with the following questionnaire that is tailored to the specific tasks of the game managers of Lucid's framework. The attached spreadsheet outlines the planned interview questionnaires, as shown in Table A.7.

Game Manager Post-hoc Interview Questionnaires	Comments on question
1. How much time did you spend to enter the meter data manually for the baseline period?	
2. How much time did you spend to enter the meter data manually for the competition period?	
3. What problems did you encounter?	
4. How much time did you spend to monitor analytical data to understand the state of the game	
5. Did you find it difficult to manage? What was difficult?	

Table A.7: Game Manager Assessment

After the interview, code and categorize the reported time and problems to identify the strengths and weaknesses in the game managing process. In addition, if possible, collect the system log data related to the game managing tasks, analyze the logs to find out the time spent and error encountered during the game managing tasks. Use the log data to verify the findings from the interview data.

These assessment reveals the strengths, weaknesses and the areas of improvement regarding the game managing process for the framework.

Action Item: Review the Game Manager tab in the attached spreadsheet. Provide comments for any Game Manager interview questionnaires that you believe might need to be changed.

A.2.2.5 Assess Game Developer Stakeholder Experience

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 development process.

We consider the tasks of game manager interacting with Lucid's framework are:

1. use API to get data in and/or out of the system
2. customize the interface
3. extend the system to support new meters
4. enhancement

We recommend the post-hoc game developer interview approach for assessing game developer stakeholder experiences.

BuildingOS and Dashboard have APIs for developing apps to tie into the framework. We will use the API to develop an extension or customization of the system. Here are the development tasks we proposed to perform using Lucid's API to extend the framework:

1. create a new widget to be available in the home page.
2. support the automated energy data collection from a new type of meter.

We will ask the identified game developers to perform the above development tasks using Lucid's framework. The developer could be Lucid internal developers or some one outside of Lucid. After the development tasks are completed, we will interview the developers to assess his experience for these development tasks. The attached spreadsheet outlines the planned questionnaires for game developer interview, as shown in Table A.8.

Game Developer Post-hoc Interview Questionnaires	Comments on question
1. How much time did you spend to implement the creation of a new widget?	
2. How much time did you spend to implement adding a new type of meter?	
3. What problem(s) did you encounter?	
4. Did you find it difficult to understand, extend and debug the system? What was difficult?	

Table A.8: Game Developer Assessment

Once the interview data is collected, categorize the reported problems and correlated with the reported time data to identify the areas of strength (less time spent) and weakness (more time spent and problems or difficulties) in the process of development.

These assessment reveals the strengths, weaknesses and the areas of improvement regarding the game development process for the framework.

Action Item: Review the Game Developer tab in the attached spreadsheet. Provide comments for any Game Developer interview questionnaires that you believe might need to be changed.

A.2.3 Choose Assessment Participants

Once the stakeholder categories are defined, the next step is to find individuals fitting those categories who will be willing to participate in the evaluation process.

Stakeholder	Person name(s)	Organization(s)	Contact(s)
Player(s)			
System admin(s)			
Game designer(s)			
Game manager(s)			
Game developer(s)			

Table A.9: Choose Participants

Table A.9 shows a sample of the ***Participants*** worksheet.

For each stakeholder, identify the name(s), organization and contact info. 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.

Action Item: Review the Participants tab in the attached spreadsheet, and provide any individuals that you believe might be able to participate at this point in the planning process.

A.2.4 Create Assessment Schedule

Once we know what the assessment approaches and who the participants are, the next step is to create the assessment schedule. We have created a sample schedule based on the sample planning timeline in the CCN Competition Planning Guide, as shown in Table A.10.

Action Item: Review the Schedule tab in the attached spreadsheet. Provide comments for any schedule items or dates that you believe might need to be changed.

Week	CCN Milestone	CCN Task	SGSEAM Task	Comments on schedule
Mar 3 - 7		Set up buildings, meters, and competition in BuildingOS	Finalize stakeholders and assessment approaches	
Mar 10 - 14		Data collection dry run week, troubleshooting and resolve problems	Finalize interview and survey questionnaires	
Mar 17 - 21	Baseline (2 weeks)	Collect & verify baseline data	Choose participants	
Mar 24 - 28				
Mar 31 - Apr 4	Competition (3 weeks)	Kick-off, Collect competition data, Enter data into BuildingOS, manage events, Social media		
Apr 7 - 11				
Apr 14 - 18				
Apr 21 - 25		Collect, verify and enter final competition data into BuildingOS	Obtain log data, System admin post-hoc interview	
Apr 28 - May 2			Player effectiveness & usability study	
May 5 - 9			Game designer post-hoc interview	
May 12 - 16			Game manager post-hoc interview	
May 19 - 23			Game developer post-hoc interview	
May 26 - 30			Analyze data, interpret strength and weakness, produce action document	

Table A.10: Assessment Schedule

A.3 Step 2: Gather Data

Once the plan has been finalized, the next step is to carry out the assessment, record the data, obtain log data, and (if necessary) refine the assessment plan. The output of this step is a data repository contains all the assessment data that can be analyzed in the next step.

A.3.1 Carry Out Assessments

For each stakeholder group, we will complete the tasks outlined in the assessment plan, gathering the data.

A.3.2 Obtain log data

Certain assessments (such as player engagement) depend upon access to certain kinds of log data. We will confer with technical staff as to how to obtain this data.

A.4 Step 3: Produce Assessment Report

In this step, we will analyze the data gathered from previous steps, create an analysis of the strengths and weakness of the framework, and produce an action report with our recommendations as to

framework improvements.

A.4.1 Analyze Data

Our analysis will include qualitative analysis of questionnaire data as well as quantitative analysis of log data. For example, for player assessment, we will calculate the engagement metrics from the game log; for game designer assessment, we will analyze interaction log data to find out the completion time for a certain game design tasks.

A.4.2 Determine Strength and Weakness

We will attempt to determine the most important problem areas from our data and summarize them, as well as the areas where the framework appears to be most successful.

A.4.3 Produce Report with Actionable Steps

Once the strengths and weaknesses of the framework are identified from the data analysis, an action report should be produced. This report includes the weakness areas that can be improved and actionable steps on how to improve from each stakeholder's perspective. It also includes strengths that the framework needs to maintain.

This concludes the SGSEAM assessment for Lucid BuildingOS and BuildingDashboard.

BIBLIOGRAPHY

[1] Victor R Basili1 Gianluigi Caldiera and H Dieter Rombach. The goal question metric approach. *Encyclopedia of software engineering*, 2:528–532, 1994.

[2] John W. Creswell. *Research design: qualitative, quantitative, and mixed methods approaches*. Sage Publications, Thousand Oaks, California, 2nd ed. edition, 2003.

[3] Thomas Tullis and William Albert. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Morgan Kaufmann, 2010.