Diary Studies: A Method For Games User Research "In The Wild"

Marina Kobayashi
Electronic Arts
209 Redwood Shores Parkway
Redwood City, CA 94065
mkobayashi@ea.com

Joseph Iloreta
Electronic Arts
209 Redwood Shores Parkway
Redwood City, CA 94065
jiloreta@ea.com

CHI'13, April 27– May 2, 2013, Paris, France.
ACM 978-1-4503-1016-1/12/05.

Abstract
While lab usability studies are one of the most widely used methods for gathering feedback on games in-development, there are many limitations and tradeoffs to the method that compels us as games user researchers to ask what methods can we apply beyond usability? This talk will discuss Diary Studies as an alternative method with some of its own counters and tradeoffs to usability studies, as well as best practices for successful diary studies and case studies from applying the method to games in development.

Author Keywords
Games user research, methods, usability, playtest, field research, remote research, diary study, video games, game design, game development, builds, staging, surveys, closed beta, open beta, game development lifecycle

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Evaluation/methodology.

General Terms
Design, Measurement

Introduction

This talk will first briefly examine the purpose of usability testing, common study constructs, and limitations or cons to the method. The later portion of the talk will explore diary studies as a means to go beyond usability, addressing some of the limitations of lab studies, but also having its own set of tradeoffs. Finally, the talk will conclude with recommendations for conducting a successful and actionable diary study.

Games Usability Testing Overview

The traditional definition of usability testing is that it involves observing "real users" interacting with the product or user experience in order to uncover errors and areas for improvement by measuring the users' performance in key areas: efficiency, accuracy, recall, and emotional response [1]. For games, the focus of usability testing is expanded to include playability aspects such as: fun, challenge, and frustration.

Games Usability Study Basics

- **Who:** recruited participants from anticipated target audience. Individuals not matching the target audience profile would potentially give feedback not appropriate to the product and a mismatch for the actual target users (for example: a puzzle gamer may not give relevant or appropriate feedback on a first-person shooter experience).
- **What:** participants have hands-on interaction with the user experience in at least one version of the game design via: wireframes, flash prototypes, development builds, and any other applicable representation.
- **How:** observations are recorded: metrics and counts of behaviors, performance, and utterances.
- **When:** ideally conducted iteratively over the course of the product development lifecycle, but can also be conducted post-release for determining improvements for later versions such as patching the software or web rollouts.
- **Why:** to eliminate aspects of the product that may produce errors or a lesser user experience, resulting in a positive and desirable product.

Figure 1. A participant in the usability lab plays the development build at this station equipped with test kits and a PC for surveys. The station is also set up with screen capture,

a webcam and microphone for capturing the gameplay, the participant's facial expression, and any utterances.

Games Usability Testing Limitations

Usability studies in labs are a go-to method for many reasons. Primarily, it allows for direct observation of a player interacting with the product. However, there are drawbacks to the method:

1) Practically speaking, usability studies can be resource intensive as they require participant recruiting, 1:1 moderation, facilities and equipment.

2) Scientifically speaking, usability studies are low on experimental realism. Participants are invited to a lab either at the company's facilities, or an external facility - the context and environment are different from where the participant would typically play games. They may be asked to interact with the new or existing product for a specific duration (shorter or longer than they would on their own), or complete specific tasks in an order they might not do them or with new or unfamiliar features and widgets.

3) Users that participate in the study may also have selection biases: geographic proximity to the test facilities allowing them to participate in the study, availability during the times that the studies are scheduled, for example: college students or unemployed persons may have more flexibility to attend during weekday business hours vs. daytime working adults potentially only being available for evening or weekend studies.

All of these differences have the potential to affect the way the participant interacts with the product (behaviors), their attitudes and opinions regarding the product (affect), and their thinking and understanding of the product (cognition).

While usability studies often identify a large amount of issues regarding user interactions with the product, the biggest detractor may be that they are not able to predict how the user may interact with the product in their own context [2] and at their own inclination.

Diary Study Overview

The basic construct of the Diary Study is that the user goes about their normal lives except that they report what they've done or experienced (via questionnaire vs. open-end diary entries, and either in a digital or physical diary). The diary studies design determines whether participants will report at intervals, given a signal, after a trigger event, and the frequency with which they will report [3,4]. Some diary studies also use artifact documentation such as taking photos, screenshots, or collecting objects and items that were used and may inform the researcher's understanding of the workflow and user experience.

Diary Study Basics

- **Who:** recruited participants. If the diary study is remote or online, players may be recruited via advertisements on relevant websites or games, or email blasts to participants from a database query.
- **What:** participants have hands-on interaction with the product in whatever format the developer is able to make available, for example: an open or closed software beta, an alternate version available on a

staging or test area of a website, or post launch study of newly released software.

- **How:** the participants report their experiences via the method or tool that the researchers have elected to use for diary entry collection (ex: online survey, mail-in journal, mail-in disposable camera, or online video logging to name a few).
- **When:** ideally conducted somewhat later in the development lifecycle so that the build is relatively polished and ready for participants to interact with it without a moderator's guidance (i.e. stable build, includes tutorial or any usage instructions expected to be included at time of release) [2]. Diary studies can also be conducted post-release for understanding the actual adoption and pattern of play.
- **Why:** to understand how players will interact with the game "in the wild," answering questions such as how often will they play and for how long? What triggers may cause them to perform some behaviors (quitting the game, making in-game purchases, inviting friends to play, etc).

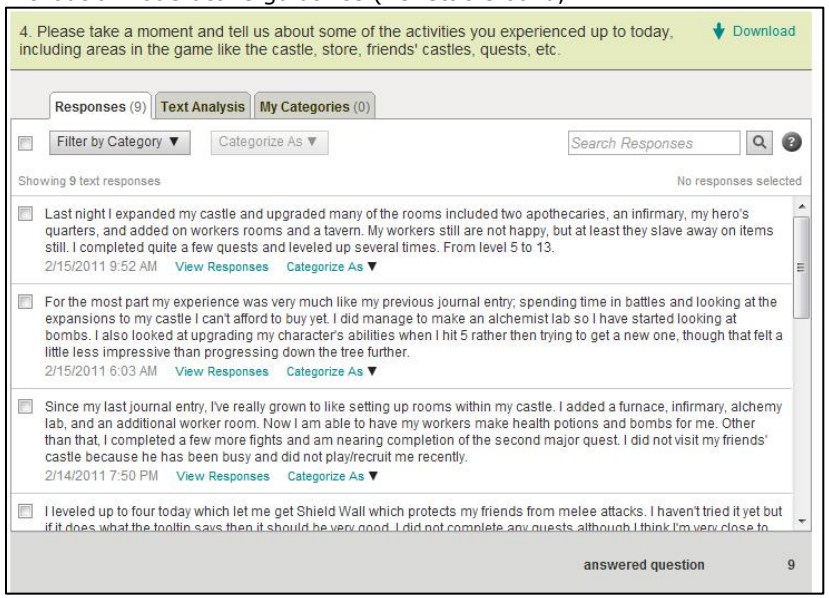

Figure 2. Using online survey tool surveymonkey.com we collected diary entries over the course of 2 weeks. Players were geographically dispersed throughout the US, completed diary entries at their convenience once every 3 days.

Diary Study Pros & Cons
Initially, Diary Studies can have fewer resource requirements (not having in-person participants, not needing a facility or lab equipment, and not requiring a moderator for hours of 1:1 sessions), but the data analysis can be lengthy depending on the number of participants, number of diary entries, and duration of the study.

Designing and conducting Diary Studies has been somewhat more open to variety such as differing formats and contexts. This may be due to the fact that the method is intended for use "in the wild" and therefore must be tailored for such a variety of contexts and workflows. This may yield very unique feedback and interesting artifacts. However, the variance of contexts and the participants' reporting practices may result in low experimental control and potentially less reliability of results across the participants.

Some user researchers regard retrospective data collection as faulty given that memory and recall of an experience may degrade over time, but in some cases, this can be of interest for designing the user experience. Game designers may be hoping to understand how the players' thoughts and emotions about the game change over time, what they are looking forward to experiencing in their next gameplay session after they have further digested previous gameplay, or even after the memory has blurred and faded somewhat. Another reason for capturing this potentially lossy data is to understand what a player may have forgotten between gaming sessions, thus where they may need reinforcement, reminders, and prompts when they return to the game.

Some researchers in the games industry have cited the difficulty in finding actionable data through Diary Studies as a weakness and a poor return on investment (ROI) for the method. However, in the next section, we will propose recommendations for ensuring that the Diary Study results in actionable data and findings.

The biggest strength of the Diary Study method is that it can give insight into how a "real user" is likely to interact with the game "in the wild." This includes understanding aspects like how frequently the player is likely to play as games are elective entertainment and there are many activities competing for the player's time (work, other entertainment, socializing, etc.), or how the player may incorporate the game into a rotation or routine. The Diary Study allows for capturing the first-time user, learning, product adoption, and progression at the user's own pace that a lab study may compress, accelerate, or stunt.

Diary Study Recommendations
Based on our own experiences designing and conducting diary studies for games and other user experiences, we've developed some practical recommendations to help you conduct a successful diary study.

1) **Recruiting:** Screen the participants as thoroughly as you would for an onsite usability study. Their fit with the target audience is important for being confident that their behaviors and gameplay patterns with the game are relevant. If the player does not match the target audience, then we cannot be confident that our target players will play at the same frequency or duration, or react to various triggers in the same ways (for ex: a casual

Facebook social games player participating in a study of a new Multiplayer Online Battle Arena (MOBA) may report frustration at extended match duration). This is also the first opportunity to set expectations and confirm with the participant that they understand what participating in a diary study will entail.

2) **Hypotheses:** It is as important as in a usability study for the games user researcher and game team to formulate hypotheses of what they expect to occur with the game "in the wild." Without clearly articulating these hypotheses or predictions, it is difficult to come up with what questions to ask the player of their experience and what metrics to collect or have the player report on. The game team should have their idea of what a good experience positive adoption (frequency and duration) in the game would be vs. a poor experience and adoption.

3) **Study design:** Carefully consider how many participants should be included in the study for the results to be digestible yet reliable. Determine when participants should complete diary entries, whether they should respond to specific prompts and questions, or if they should be free-form entries. We have found that to collect data in diary studies that will help game designers improve the "real world" game experience it may be necessary to set expectations for players on what data to report.

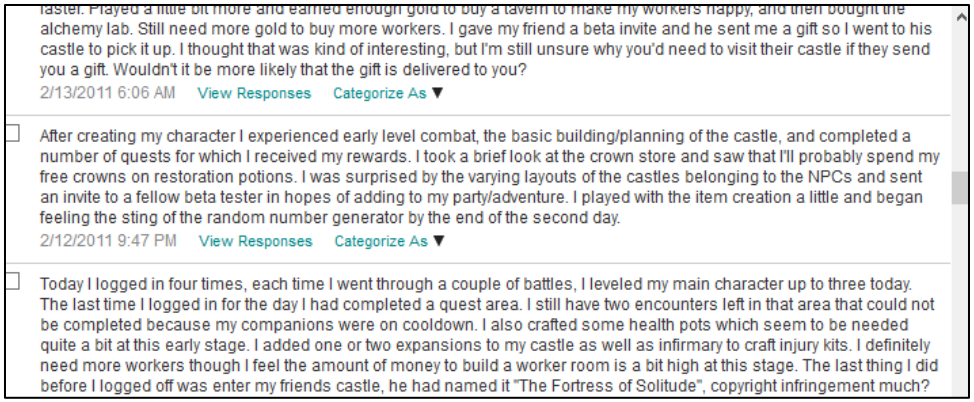

Figure 3. Over the duration of a study, a game team wanted to see whether players would elaborate on their castle – a base-building component of the game – without prompt from the game. Several of the participants reported visiting their NPC's castles and discovering other interesting rooms of castle building, which in turned incentivized them to invest in their own castles.

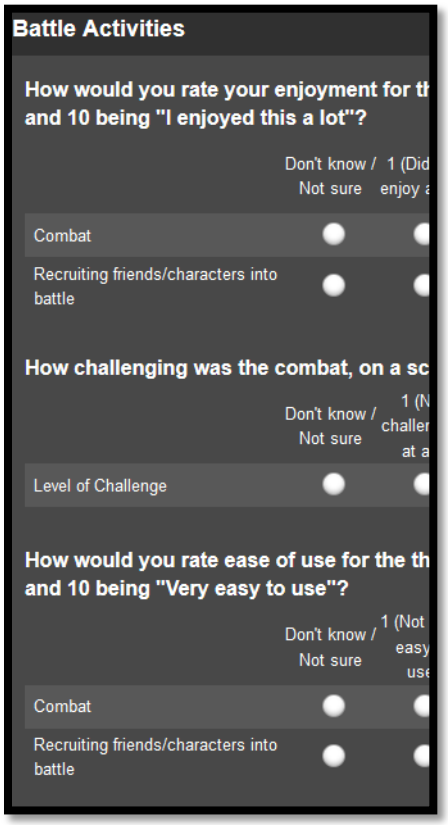

Figure 4. In this study, daily questions were designed to track perceptions around core loop activities and gleam difficulty and complexity over the course of the study.

4) **Pilot testing:** We recommend pilot testing the Diary Study either by playing the game and responding to with Diary entries oneself, or having a colleague do so. This will give the researchers and the game team a preview of the type of data they can expect to receive and determine if questions or instructions need to be modified and if the prompts for completing diary entries should be different (time intervals & frequency vs. triggers & events). This is the best opportunity to verify before conducting the study whether the feedback collected will be actionable, or if the game team in-fact needs other types of data collected to understand the "real user" experience with the game "in the wild" and make improvements.
5) **Builds:** The same if not higher standards must be applied: builds must be stable (not crashing or freezing during gameplay), if there are errors or crashes players must be able to recover their session and proceed unaided, and the build version must be kept consistent throughout the diary study. We have had game teams roll out updates to web-based games that had Diary Study participants opted-in – the result was a lot of confused players reporting that the game had changed since their last visit.
6) **Data analysis:** During the Diary Study it can be helpful to review entries as they come in to verify that the feedback is actionable and meaningful - in terms of the questions asked, but also if there is a bad participant they can be excused without wasting time (ex: participants that do not follow instructions, not completing reports, or fill text fields with obscenities), to check for any red flags (reports of a bad build or other problems that may require researcher or game team intervention), previewing early patterns and themes, and to break up the data analysis into smaller more manageable chunks rather than waiting until the conclusion of the study.

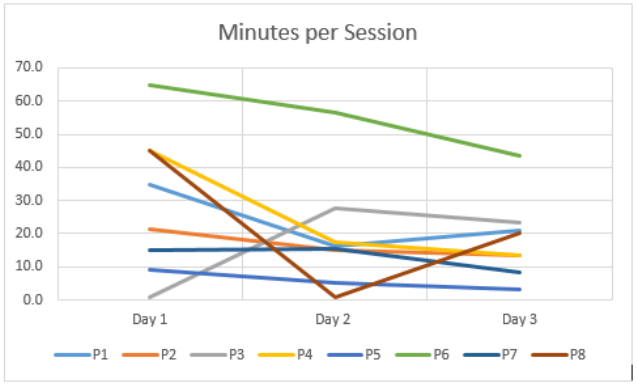

Figure 5. In the study of a Facebook title, players were asked to log the number of sessions and the length of time spent per session and provide reasons as to why. Here, the difficulty ramp as well as cool-downs required of NPC party members led to players spending less time per session by the end of the Beta, to the disappointment of some players.

7) **Collaboration:** Involve the game team in reviewing the diary entries as frequently as possible by sending interval summaries, or providing them with view access to the diary tool if they can allocate time to assisting with the theme and pattern recognition. The richness of the Diary Study data can inspire action and impress upon the game team the reality their game being played by real players in the wild.

Conclusion

Diary Studies can be a powerful method for Games User Research, but like all other methods there are tradeoffs and best practices. Diary Studies should not replace Usability Studies, and neither is the perfect method for all aspects of the player experience. As described earlier, Usability's strength is in providing direct observation of players' hands-on interactions with games benefitting from a moderated lab study context. Diary Studies benefits differ in that the purpose of the diary study is gathering the patterns of play with the game taken "into the wild."

Authors

Marina Kobayashi joined the EA Game Lab in 2009 and is now a Senior Games User Researcher. In the last 4 years she's consulted for DICE, BioWare, and EA Partners with a focus on AAA titles and HD gaming experiences. Marina pushed the envelope on ways to adapt traditional user research methods for gaming contexts. Marina earned her Masters in Human-Computer Interaction, B.S. in Policy & Management, additional major in Decision Science, and minor in Business Administration all at Carnegie Mellon University.

Joseph Iloreta is a Games User Researcher for the EA Game Lab. He leads studies with development teams across EA's various studios working in the social and mobile space. Previously, Joseph worked as design researcher/user experience researcher for Belkin in Los Angeles, leading teams into the field and informing design decisions for Belkin's range of consumer electronic products. Joseph obtained both a B.A. in Psychology and HCI and a Masters of Design in Interaction Design from Carnegie Mellon University.

Acknowledgements

Authors Marina Kobayashi and Joseph Iloreta thank all the members of the EA Game Lab for their support, and a special thank you to Tim Toy for recruiting for our diary studies. They gratefully acknowledge Electronic Arts for supporting and growing the EA Game Lab, its practice of Games User Research including the application and adaptation of methods such as Diary Studies. They would also like to acknowledge their management: Scott Taylor and Jim Lewis. They would also like to acknowledge the game teams' whose studies are mentioned here, including *Dragon Age: Legends* and *Dragon Age 2*.

References

[1] Wikipedia: Usability testing. http://en.wikipedia.org/wiki/Usability_testing

[2] Isbister, K. and Schaffer, N. *Game Usability*. Morgan Kaufmann (2008).

[3] Wikipedia: Diary studies in User Research. http://en.wikipedia.org/wiki/Diary_studies_in_User_Research

[4] Karis, D. *Diary Studies in HCI & Psychology: Why They're Useful and How to Conduct Them.* UPA Boston's Tenth Annual Mini UPA Conference (2011).

[5] Davis, J; Steury, K; and Pagulayan, R. *A survey method for assessing perceptions of a game: The consumer playtest in game design.* The International Journal of Computer Game Research, (2005). http://www.gamestudies.org/0501/davis_steury_pagulayan/

[6] Fulton, B. *Beyond Psychological Theory: Getting Data that Improves Games*, Gamasutra (2002). http://www.gamasutra.com/gdc2002/features/fulton/fulton_01.htm