



Data-Driven Retrospective Interviewing (DDRI): A proposed methodology for formative evaluation of pervasive games [☆]



Magy Seif El-Nasr ^{*}, Shree Durga, Mariya Shiyko, Carmen Sceppa

Northeastern University, United States

ARTICLE INFO

Article history:

Received 6 December 2014

Revised 29 June 2015

Accepted 6 July 2015

Available online 29 July 2015

Keywords:

Serious games

Evaluation

Mixed-methods

Game analytics

Methodology

ABSTRACT

Recently, games have gained much momentum as enablers of health behavior change. Current projects on health games, however, face a serious challenge: evaluation methods used to assess and formatively evaluate such games do not adequately gauge acceptability and integration of the game into participants' lives. These are crucial elements to enable adherence of the participants to the game to induce behavior change. In this paper we present a formative evaluation methodology we developed, called *Data-Driven Retrospective Interviewing (DDRI)*. Used within naturalistic settings, DDRI investigates how participants accept and integrate a game into their lives. The method is comprised of several steps. A game is first instrumented to collect behavioral data, which is then analyzed to inform the construction of interview questions used to retrospectively contextualize and reflect on participants' behaviors. The contribution of this paper is: the methodology and its application to evaluate a pervasive health game called *SpaPlay* (Seif El-Nasr et al., 2011). Findings of this investigation helped designers re-think and augment their design as they uncovered how users used their game, which, in some cases, was different from how they initially intended. Such findings would have not been possible using conventional evaluation methods.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Software evaluation is an important topic that has acquired much attention in the recent years. Substantive progress made in the field of Human Computer Interaction research focused on methods to investigate *usability of software* (e.g., [28,33,40,44,50,63], *requirements gathering* (e.g., [11,31,36], and *software evaluation* [12,27,59,60,66]). While such methods have been developed to target the evaluation of various applications, the development of game specific summative and formative evaluation methods received little attention.

In the recent years, games have gained much attention. The Entertainment Software Association [25] reports that consumers spent \$24.75 billion on video games, hardware and accessories in 2011 [25]. Not only do 72% of American households play computer or video games, the age distribution is versatile: with 18% under the age of 18 years old, 53% 18–48 years old, and 29% 50+ years old. Popularity of games has led the scientific community to

consider the use of games for constructive purposes beyond entertainment. To date, there is a wide range of research and development targeting the use of games for a variety of purposes. Examples of this work include games for promoting physical activity (e.g., [82,89,90]), education (e.g., [1,30,78,91]), training (e.g., [32,77,92]), and music education [2].

With such applications, there is a dire need for rigorous and valid formative and summative evaluation methods that can comprehensively demonstrate and exemplify the impact or efficacy of such games. Evaluating a serious game's efficacy, however, requires testing several aspects, including usability, acceptability, integrability to participants' life style as well as its effect in terms of health or educational outcomes. In this article we focus on acceptability and integrability measures. For measuring outcomes researchers can and often use methods developed by the health sciences, training, and education communities. These include nutrition knowledge questionnaires, and self-efficacy surveys, to mention a few.

While HCI researchers have proposed many methods for usability, acceptability and general evaluation of software applications, these methods are often geared towards evaluating productivity software and cannot be used to test the subjective experience of playing a game. Pagulayan et al. [65] argue that games are different

[☆] This paper has been recommended for acceptance by Pierre Jovelot.

^{*} Corresponding author.

E-mail addresses: magy@neu.edu (M. Seif El-Nasr), shree.durga@gmail.com (S. Durga), M.Shiyko@neu.edu (M. Shiyko), c.sceppa@neu.edu (C. Sceppa).

from productivity software; they outline several reasons, including that games' purchase and use are discretionary, and gamers demand novelty. Most importantly, they argue that user research is an integral process of game design, because, unlike productivity software, games rely on sustained engagement.

To address this issue, the Games User Research (GUR) community has focused on developing tailored methods that are geared towards rigor and depth in game evaluation approaches. To date, best practices within the game user research community include heuristic evaluation, think-aloud, and playtesting methods [19]. Heuristic evaluation is an evaluative approach where experts use a set of guidelines to critique the game in terms of its usability and playability [16,18]. Think aloud [48] is a usability testing method, where users are asked to interact with a piece of software (in this case a game) and utter loudly what they are thinking as they are interacting with the product. Finally, playtesting is a method that involves recruiting 20–25 players to play a game. A combination of interviews, questionnaires and analysis of game log files are used to gauge pacing and game balance problems, as well as players' attitudes and perceptions [35]. More advanced approaches have also been proposed, such as eye tracking to study attention [79,38] or game analytics for analyzing behavior [75,76].

Game analytics is an approach in which game data, automatically recorded with timestamps, is analyzed to derive actionable results for different stakeholders: business, marketing, design, and software development. The field is composed of a large group of theories, methods, processes, and technologies that are used to transform raw data into meaningful information. Game analytics is currently gaining momentum due to its clear promise for game evaluation and assessment [75,76]. As a method it can contribute to understanding frequency and patterns of play through analyses of objective behavioral data recorded in real time for many players. While game analytics can be used to enable better evaluation, the approach is limited to only a recount of actual behavior with little understanding of users' perceptions, needs, feelings, or attitudes – all of which are essential to explain reasons for the observed behaviors.

The field of Game User Research is in its infancy and new approaches are needed to combat the current limitations. In particular, in this paper we target the development of a method to assess acceptability and integrability of pervasive casual games into users' life style. Since we target pervasive games, evaluation methodologies will need to be ecologically valid, i.e. situated in naturalistic/ecological rather than laboratory settings. Some researchers studying usability and adherence of pervasive technologies have discussed the need to improve ecological validity of evaluation techniques [12,59]. For example, in the context of evaluating ubiquitous mobile computing technologies, Carter et al. [6] argued that the evaluation of mobile application is challenging as “the attentional demands of mobile applications cannot be simulated in lab environments, because in realistic environments a plethora of activities interact to constrain severely the continuous periods that participants can attend to mobile devices (p. 64).” They added that controlled lab evaluations could be “very effective at testing issues of aesthetics and standard graphical interface interaction, as well as for comparing possible solutions (p. 81).” As a result, several methods were proposed to help develop rigorous evaluation methods in naturalistic settings. In particular, researchers have used and combined several methods including observation through ethnography or contextual inquiry, interviews, experience sampling or event-contingent experience sampling, and diary-based studies. While some of these research methods present great innovations on the current state of the art, they are limited in that they do not provide an objective record of behaviors with a clear indication of

participants' perceptions of why they made these choices. In addition, most of the work on pervasive computing cannot directly be translated to pervasive games due to the reasons discussed above. Therefore, there is a need for a new research method that can allow us to formatively evaluate pervasive games within naturalistic settings, objectively measuring time use, patterns of play over time, as well as the acceptance and integrability of the game.

As a first step towards developing a solution to address this need, we describe a new method we developed called *Data-Driven Retrospective Interviewing (DDRI)*. Using DDRI, we allow participants to take the game and play it in their natural setting. Studies are conducted longitudinally spanning weeks or months. Data from the game are collected in real time. We analyze and visualize these data and use them to construct interviews, which are often conducted every 7–15 days. We use the visualizations and interview questions to allow participants to retrospectively explain their play behaviors, motivations, engagement with the game, and how the game was integrated into their life context. Results from these individual interviews are then used to re-evaluate the quantitative data collected and make sense of play patterns.

This methodology shares some similarity with other methods and methodologies. For example, it is similar in philosophy to Action Research first proposed by Lewin in his seminal paper “Action Research and Minority Problems” [47]. Lewin construed Action Research as a process of research that consists of “analysis, fact-finding, conceptualization, planning, execution, more fact-finding or evaluation; and then a repetition of this whole circle of activities; indeed a spiral of such circles [71, p. 4].” However, action research has often been used in the context of social problems rather than behavior change, and as Kemmis [41] argues Action Research lacked a formalized and standardized process, which has led to criticism of the methodology in later years. DDRI is similar in philosophy yet uses more standardized approaches combining analytics, data-driven analysis with interviewing procedures for formative evaluation. This methodology and its unique combination of methods constitute the novelty of the approach presented in this paper.

In addition, DDRI is also similar to Experience Sampling Methods (ESM) [31] and Day Reconstruction Method (DRM) [39], in that it collects behavioral data during participants' play. Like ESM and DRM, DDRI collects experiences over time and uses qualitative methods to retrospectively explain the behaviors collected. However, the cyclic nature of DDRI differentiates it from such methods.

In this paper, we will discuss DDRI in detail and outline the results obtained from using it to evaluate a health game, called *SpaPlay* [74]. *SpaPlay* is a social commercial game designed to motivate people to adopt a healthy lifestyle. The game currently has over 400 players actively playing it. It is designed to engage participants in eating healthy and physical activities. It stimulates engagement through intrinsic and extrinsic rewards for making the right food choices and being physically active *outside* of the game. We ran a pilot study with 18 participants using DDRI. We will discuss this study to provide evidence on the effectiveness of DDRI in unpacking how players accepted and integrated the game in their daily lives. The method presents a solid contribution towards developing a method that can be used to investigate questions regarding pervasive games' acceptability and integration within participants' life style.

The paper is organized as follows. First, we discuss previous work in the area of user research, serious game evaluation and HCI usability and evaluation methods. Second, we define the methodology of DDRI delineating different techniques entailed in using this approach. Third, we present a case study conducted

Download English Version:

<https://daneshyari.com/en/article/381804>

Download Persian Version:

<https://daneshyari.com/article/381804>

[Daneshyari.com](https://daneshyari.com)