Chapter 7
Evaluation of Serious Games

Stefan Göbel
Technische Universität Darmstadt, Germany

Michael Gutjahr
Technische Universität Darmstadt, Germany

Sandro Hardy
Technische Universität Darmstadt, Germany

ABSTRACT

Comprehensive evaluation studies are necessary to "prove" the benefit of Serious Games (SG). This is also extremely important for the commercial success of SG: Best practice examples with profound, well-recorded positive effects will provide relevant arguments to invest into SG for training/education, sports and health, and other application domains. On the other hand, it is not easy to prove the benefit of SG and to measure its effects (e.g. learning effects or medical effects) and affects (user experience factors such as fun during play). Evaluation methodologies might be split into observation, self-evaluation (e.g. questionnaires, interviews), associative methods, performance analyses, and psychophysiology measurement. Technology-enhanced evaluation methods, for instance, facing expression measurement are in the centre of attention. This chapter provides an overview of these methods and describes current interdisciplinary research and technology development achievements in that field.

INTRODUCTION

The main target of an evaluation study for Serious Games is to prove an impact of a SG (fulfillment of overall serious goal plus entertainment factors). Therefore, two important questions have to be addressed: *What* effects and affects are going to be measured? And *How* is it done? "What" defines the target of the game "what are the main effects and affects the game should fulfill and be proved by an evaluation study?" (e.g. improvement of knowledge, performance, well-being). "How"

DOI: 10.4018/978-1-4666-3673-6.ch007

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

defines the method, includes aspects of evaluation methods design and instruments. "How has the study to be designed (setting)" and "how can the effects been measured (instrument)".

To decide what to measure, a first classification can be conducted by dividing the "impact" of a game in three categories: emotion, cognition, and performance. Emotions are all variables that can be used to draw a conclusion of the emotional state of the user and its experience of the game. But the term "user experience" (Lennart Nacke, 2009) represents a complex construct including aspects of positive emotions, negative emotions, effectance, motivation, immersion, flow, arousal, curiosity, usability and not at last cognitive load. Cognition includes all variables showing some changes in cognition and its structure like problem solving, decision making, memory and available knowledge. Performance includes all variables indicating an advancement in the person's ability to conduct a task like reaction time, persistence, strength.

Concerning 'what to measure', a lot of research has been done to show effects and to prove the benefit of Serious Games in particular application domains. Examples include work in the field of game-based learning, e.g. Egenfeldt-Nielsen (2005) conducted a study at a Danish high-school involving 72 students and two teachers. In this study the use of Europa Universalis II, a commercial historical strategy game was examined. De Freitas and Oliver (2006) introduced a framework for helping tutors evaluating the potential of using games- and simulation-based learning and applied their framework to two Serious Games. Mitchell, Savill-Smith, and Britain (2004) provide a comprehensive overview over the use of computer and video games for learning. They argue to use those games for learning as they engage players and draw them into virtual worlds, because they are fun and provide a challenge and immediate feedback, and because they handle huge amounts of content and can be instantly updated and customized. Law, Kickmeier-Rust, Albert, and Holzinger (2008) provide an evaluation framework based on a holistic understanding and a formal ontological representation of interacting processes involved in active and dynamic learning processes, which is able to measure performance outcomes, attitude outcomes, and programmatic outcomes. Shen, Wang, and Ritterfeld (2009) performed a study about the entertainment value of Serious Games. Their findings reveal that Serious Games can be reasonably enjoyable compared to non-serious games. Further work in the field of game-based learning comes from Malone and Lepper (1987), Klabbers (2003), and Prensky (2003). Work in the field of games for health and sport shows some effectiveness of games e.g. a reduction of "diabetes-related urgent and emergency visits by 77 percent after diabetic youngsters had the game at home for six months, compared to […] a control group" (Lieberman, 2001). Whitehead, Johnston, Nixon and Welch (2010) "survey a number of quantitative exergame studies to define a general set of elements that make exergames effective from a physical standpoint", Kato, Cole, Bradlyn and Pollock (2008) evaluated the effectiveness of a game, Re-Mission, for cancer therapy and lined out that "the video-game intervention significantly improved treatment adherence and indicators of cancer-related self-efficacy and knowledge". Further work in the field of games for health and sport comes from Baranowski, Buday, Thompson and Baranowski (2008), Papastergiou (2009) and Kretschmann, Dittus, Lutz, and Meier (2010).

But *how* can emotion, cognition and performance be measured? Here it is necessary to differ between the *method* (the setting/the design of a study) and the *instrument* (tool) that should be used. The method is the design of the study (the setting in which the data are collected). Most important for the possibility to interpret empirical data is the methodical design of the research study. An experimental design with a control group allows a causal interpretation and therefore is mostly desirable. Conducting a pretest and a posttest in both groups (the experimental group

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/evaluation-serious-games/75810

Related Content

Designing a 21st-Century Learning Module With CGScholar: An Example of Design for Brazilian Educators
Vania Carvalho de Castro (2023). *Promoting Next-Generation Learning Environments Through CGScholar (pp. 297-311).*
www.irma-international.org/chapter/designing-a-21st-century-learning-module-with-cgscholar/326793

Quality Enhancement in Higher Education Institutions in India: Challenges Ahead
Katta Rama Mohana Raoand Chandra Sekhar Patro (2016). *International Journal of Virtual and Personal Learning Environments (pp. 29-40).*
www.irma-international.org/article/quality-enhancement-in-higher-education-institutions-in-india/188427

E-Learning Systems Requirements Elicitation: Perspectives and Considerations
Shaikha B. AlKhuderand Fatma H. AlAli (2017). *International Journal of Virtual and Personal Learning Environments (pp. 44-55).*
www.irma-international.org/article/e-learning-systems-requirements-elicitation/194035

Toward a Personal Learning Environment Framework
Mohamed Amine Chatti, Mohammad Ridwan Agustiawan, Matthias Jarkeand Marcus Specht (2010). *International Journal of Virtual and Personal Learning Environments (pp. 66-85).*
www.irma-international.org/article/toward-personal-learning-environment-framework/48222

A Framework for Effective Online Professional Development
Heidi Marie Rock (2019). *Student-Centered Virtual Learning Environments in Higher Education (pp. 210-232).*
www.irma-international.org/chapter/a-framework-for-effective-online-professional-development/214524