Chapter 17

A Brief Methodology for Researching and Evaluating Serious Games and Game-Based Learning

Igor Mayer
Delft University of Technology, The Netherlands

Geertje Bekebrede
Delft University of Technology, The Netherlands

Harald Warmelink
Delft University of Technology, The Netherlands

Qiqi Zhou
Delft University of Technology, The Netherlands

ABSTRACT

In this chapter, the authors present a methodology for researching and evaluating Serious Games (SG) and digital (or other forms of) Game-Based Learning (GBL). The methodology consists of the following elements: 1) frame-reflective analysis; 2) a methodology explicating the rationale behind a conceptual-research model; 3) research designs and data-gathering procedures; 4) validated research instruments and tools; 5) a body of knowledge that provides operationalised models and hypotheses; and 6) professional ethics. The methodology is intended to resolve the dilemma between the “general- ity” and “standardisation” required for comparative, theory-based research and the “specificity” and “flexibility” needed for evaluating specific cases.

DOI: 10.4018/978-1-4666-4773-2.ch017
INTRODUCTION

The growing interest in digital and other forms of game-based learning (GBL), serious games and simulation gaming (both abbreviated as SG) is accompanied by an increasing need to know the effects of what we are doing and promoting (Mayer, Bekebrede et al., 2013; Mayer, Warmelink, & Bekebrede, 2012). Meeting this need requires proper methods, tools and principles that can be agreed upon, validated and applied by the fragmented GBL and SG communities. In other words, we must move towards a ‘science of game-based learning’ (Sanchez, Cannon-Bowers, & Bowers, 2010). Considerable efforts and resources are currently being devoted to researching and evaluating SG and GBL, thereby increasing both the number and the quality of such evaluations (see discussion below). Considerable weaknesses remain, however, including the following:

- A lack of comprehensive, multipurpose frameworks for comparative, longitudinal evaluation (Blunt, 2006; Meyer, 2010; Mortagy & Boghikian-Whitby, 2010; Vartiainen, 2000.)
- Few theories with which to formulate and test hypotheses (Mayer, 2005; Noy, Raban, & Ravid, 2006.)
- Few operationalised models with which to examine ‘causal’ relations (e.g. in structural equation models) (Connolly, Stansfield, & Hainey, 2009; Hainey, 2011; Hainey & Connolly, 2010.)
- Few validated questionnaires, constructs or scales, whether from other fields (e.g. psychology) or constructed especially for SG and GBL (Boyle, Connolly, & Hainey, 2011; Brockmeyer et al., 2009; Mayes & Cotton, 2001.)
- A lack of proper research designs that can be used in dynamic, professional learning contexts other than that of the less preferable randomised controlled trials (RCT), which are impractical, unethical and uncommon in almost every domain except medicine, therapy and related fields (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Kato, Cole, Bradlyn, & Pollock, 2008; Knight et al., 2010; Szturm, Betker, Moussavi, Desai, & Goodman, 2011; van der Spek, Wouters, & Van Oostendorp, 2011; van der Spek, 2011.)
- The absence of generic tools for unobtrusive (‘stealth’) data-gathering and assessment in and around SG (Kickmeier-Rust, Steiner, & Albert, 2009; Shute, Masduki, & Dommez, 2010; Shute, Ventura, Bauer, & Zapata-Rivera, 2009; Shute, 2011.)

In short, despite a promising increase in publications, methods, tools and findings, we continue to lack an overarching methodology for SG research. The fragmented SG and GBL community faces the enormous challenge of aligning itself in order to evaluate and research gaming for learning in a comparative, systematic fashion using procedures, frameworks and methods that can be validated, checked and reproduced. It is valuable to identify the effects of playing games through randomised, controlled trials (RCT), many of which involve students in laboratories. Nevertheless, it is also essential to identify the effects of GBL and SG in uncontrolled circumstances and within the context of objectives that truly matter for real-life performance (e.g. emergency management, leadership), as is usually the case in professional learning and training. The challenge is to gather data on the quality, application and outcomes of a broad range of SG on different topics and with different objectives, within and for different institutional contexts, at different times and under uncontrolled conditions.

Research Objective

The objective of this chapter is to present a condensed methodology for researching and evaluat-