Assessing the effectiveness of digital game-based learning: Best practices

Anissa All[*], Elena Patricia Nuñez Castellar, Jan Van Looy

Department of Communication Sciences, iMinds-MICT-Ghent University, Belgium

ARTICLE INFO

ABSTRACT

In recent years, research into the effectiveness of digital game-based learning (DGBL) has increased. However, a large heterogeneity in methods for assessing the effectiveness of DGBL exist, leading to questions regarding reliability and validity of certain methods. This has resulted in the need for a scientific basis to conduct this type of research, providing procedures, frameworks and methods that can be validated. The present study is part of a larger systematic process towards the development of a standardized procedure for conducting DGBL effectiveness studies. In a first phase, the variety in methods that are used for sampling, implementation of the interventions, measures and data analysis were mapped in a systematic literature review using Cochrane guidelines. The present paper reflects the second stage, where this variety in elements are presented to experts in psychology and pedagogy by means of semi-structured interviews, in order to define preferred methods for conducting DGBL effectiveness studies. The interview was structured according to five dimensions that were used in the literature review: 1) participants (e.g., characteristics of the sample involved) 2) intervention (e.g., contents, format, timings and treatment lengths, intervention(s) in control group(s)) 3) methods (sampling, assignment of participants to conditions, number of testing moments) 4) outcome measures (e.g., instruments used to measure a certain outcome) and 5) data-analysis. The interviews were transcribed and analyzed using qualitative software package nVivo. Our results show that areas for improvement involve the intervention dimension and the methods dimension. The proposed improvements relate to implementation of the interventions in both the experimental and control group, determining which elements are preferably omitted during the intervention (such as guidance by the instructor, extra elements that consist of substantive information) and which elements would be aloud (e.g., procedural help, training session). Also, variables on which similarity between experimental and control condition should be attained were determined (e.g., time exposed to intervention, instructor, day of the week). With regard to the methods dimension, proposed improvements relate to assignment of participants to conditions (e.g., variables to take into account when using blocked randomized design), general design (e.g. necessity of a pre-test and control group) test development (e.g., develop and pilot parallel tests) and testing moments (e.g., follow up after minimum 2 weeks). In sum, the present paper provides best practices that cover all aspects of the study design and consist of game specific elements. While several suggestions have been previously made regarding research design of DGBL effectiveness studies, these often do not cover all aspects of the research design. Hence, the results of this study can be seen as a base for a more systematic approach, which can be validated in the future in order to develop a standardized

* Corresponding author. Korte Meer 7-9-11, 9000 Gent, Belgium.
E-mail address: anissa.all@ugent.be (A. All).

procedure for assessing the effectiveness of DGBL that can be applied flexibly across different contexts.

1. Introduction

Digital games encompass a variety of types and genres that can be played using a multitude of digital technologies such as computers, (handheld) consoles and mobile devices. Based on a literature review on digital games definitions, Juul (2003) defines a digital game as

> … a rule-based formal system with a variable and quantifiable outcome, where different outcomes are assigned different values, the player exerts effort in order to influence the outcome, the player feels attached to the outcome, and the consequences of the activity are optional and negotiable (p.5).

Digital game-based learning (DGBL) refers to the usage of the entertaining power of digital games to serve an educational purpose (Prensky, 2001). DGBL is the consequence of a balance between learning and gaming elements (Nussbaum & Beserra, 2014). DGBL contains two important elements: fun/entertainment and an educational component (Bellotti, Kapralos, Lee, Moreno-Ger, & Berta, 2013). Consequently, in the DGBL literature and published effectiveness studies both learning and player engagement/motivation are considered relevant to assess (Bellotti et al., 2013).

Two types of games can be distinguished in DGBL: special purpose games which have been developed with an educational purpose and Commercial-Off-The-Shelf games that have been developed for entertainment purposes, but that are being deployed in an educational context. Note, however, that this does not mean that special-purpose DGBL games cannot be commercially available (Stewart et al., 2013).

Based on the projected primary learning outcomes, three types of special-purpose games can be distinguished. They aim to achieve knowledge transfer (cognitive learning outcomes), skill acquisition (skill-based learning outcomes), and/or attitudinal/behavioral change (affective learning outcomes) (Stewart et al., 2013). Games that are developed with the primary aim of achieving knowledge transfer are typically implemented in education, in order to teach math (Castellar, All, de Marez, & Van Looy, 2015) or language (Yip & Kwan, 2006), for example. Digital games that primarily aim to support skill acquisition are generally used for training, for example in a corporate or military context. For instance, several studies have examined the impact of playing games to develop managerial skills (Corsi et al., 2006; Kretschmann, 2012). Games that are developed to achieve attitudinal change are sometimes used by governments and NGOs to raise awareness about a certain topic, such as poverty (Neys, Van Looy, De grove, & Jansz, 2012). Games with a behavioral change intention are typically found in the health sector. For example, some games promote healthy food and physical activity to children (Baranowski, Buday, Thompson, & Baranowski, 2008). While DGBL can primarily aim to achieve a certain type of learning outcome, this does not exclude secondary learning outcomes (Kraiger, Ford, & Salas, 1993). For instance, a game that primarily aims to teach children English (cognitive learning outcomes) can also result in a more positive attitude towards learning English or English as a subject (affective learning outcomes).

Although meta-analyses have proven the effectiveness of DGBL (Backlund & Hendrix, 2013; Clark, Tanner-Smith, & Killingsworth, 2015; Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012), certain authors have pointed out elements that jeopardize reliability and validity of some findings. This includes comparisons with control groups that did not receive an educational intervention (Hays, 2005), time-on-task differences between experimental and control groups, and validity of research instruments (Randel, Morris, Wetzel, & Whitehill, 1992). Moreover, some studies do not provide enough information about the implementation of the intervention (Clark et al., 2015; Sitzmann, 2011). This makes it hard for readers to know if the reported results are a consequence of the different methods, and not a cause of circumstantial factors that differed between conditions (Randel et al., 1992). Rigorous assessment is required to improve the quality of DGBL, to support resource allocation, and to gain insight in the most effective way to use games to support learning (De Freitas, 2006; Kirriemur, 2004).

1.1. Studies about DGBL effectiveness

Two types of evaluation of educational interventions can be distinguished. A first type is formative evaluation which aims to determine areas for improvement and is thus an evaluation of the process of the intervention itself. This type of evaluation is conducted by using a naturalistic design with observational data collection, which describes an ongoing process in its natural setting. A second type is summative evaluation, which aims at to determine whether or not an educational intervention succeeds in attaining its goals, thus evaluating the outcomes (Calder, 2013). Summative evaluations are conducted by using an experimental design (Hutchinson, 1999). In the present study, we focus on summative evaluation and will concordantly discuss experimental design.

An earlier content analysis on the effectiveness of DGBL approaches, conducted by the current authors, showed that there is a large diversity in the way that experimental research on DGBL effectiveness assessment is conducted, making comparison of results across studies difficult. This heterogeneity can be found on all four dimensions of the study design, as defined by