Entertainment Computing 27 (2018) 209–224

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: www.elsevier.com/locate/entcom

Specification and evaluation of an assessment engine for educational games: Empowering educators with an assessment editor and a learning analytics dashboard

Yaëlle Chaudy*, Thomas Connolly

University of the West of Scotland, Paisley, UK

ARTICLE INFO

Keywords:
Educational games
Assessment
Learning analytics
Assessment editor
Assessment engine

ABSTRACT

Assessment is a crucial aspect of any teaching and learning process. Educational games offer promising advantages for assessment; personalised feedback to students and automated assessment process. However, while many teachers agree that educational games increase motivation, learning and retention, few of them are ready to fully trust them as an assessment tool. We believe there are two main reasons for this lack of trust: educators are not given sufficient information about the gameplays, and many educational games are distributed as black-boxes, unmodifiable by teachers. This paper presents an assessment engine designed to separate a game and its assessment. It allows teachers to modify a game's assessment after distribution and visualise gameplay data via a learning analytics dashboard. The engine was evaluated quantitatively by 31 educators. Findings were overall very positive: both the assessment editor and the learning analytics dashboard were rated useful and easy to use. The evaluation also indicates that, having access to EngAGe, educators would be more likely to trust a game's assessment. This paper concludes that EngAGe can be used by educators effectively to modify educational games' assessment and visualise gameplay data, and that it contributes to increasing their trust in educational games as an assessment tool.

1. Introduction

Games-based learning (GBL) is increasingly used as a supplementary tool for education. GBL offers a variety of advantages to assist traditional teaching. They can, for instance, allow students to learn at their own pace and they are a safe and controlled environment for students to learn through trial and error. Various institutions use GBL for learning and training, ranging from schools [32] and higher education institutions [13] to healthcare [34,51] and the army [29,55].

However, while many teachers agree that GBL increases motivation towards learning [43] and despite the evidence that games are valid assessment tools [24], there seems to be a lack of trust in an educational game's assessment [43,48]. Teachers need to feel in control before introducing a new tool in the classroom and there is a need for ownership over the game [30]; without control, educators might feel threatened by a game rather than supported by it.

One of the main limitations of GBL is that educational games are too often distributed as "*black-boxes*"; they are closed and self-contained systems, making it difficult to modify or retrieve data from [47]. This can mean that the potential of the game and its attractiveness to educators are reduced. Indeed, in traditional teaching, improvisation and adaptation to students represent a key aspect of the educator's role [27], however, teachers tend to lose this capacity with the introduction of a tool they cannot modify to suit the needs of their students. Then, they cannot retrieve data about the gameplays to appreciate whether their teaching goals have been met. Educators and researchers have very little insight about what the students learn through a computer game and how they interact with it. Learning Analytics (LA) is an emerging field based on data mining processes [50] that can provide such detailed reports about the gameplays; data from the gameplays of several educational games are collected and data mining algorithms allow conclusions to be drawn about the games and the players. However, due to the novelty of the field, presently very few papers exist on LA and its application in GBL and LA is still beyond the reach of most teachers [28].

Various platforms such as < e-Adventure > [53] or e-CLIL [21] provide educators with the ability to create and modify their own computer games; < e-Adventure > even includes a learning analytics module [36]. These games engines externalise content and assessment integration from the game's code and partially address the problems

* Corresponding author.
E-mail addresses: yaelle.chaudy@uws.ac.uk (Y. Chaudy), thomas.connolly@uws.ac.uk (T. Connolly).

https://doi.org/10.1016/j.entcom.2018.07.003
Received 27 September 2017; Received in revised form 6 July 2018; Accepted 12 July 2018
1875-9521/

identified previously. However, these engines were created for educators alone; they are not meant to be used when working with game developers and therefore only provide limited options in terms of game genres and assessment integration. Teachers sometimes lack the time to develop the games themselves or there is a need for a type of game not offered by such platforms.

To summarise, computer games are a powerful tool for learning and assessment but they are often underused by educators, particularly for assessment. We propose three key improvements that could be made for GBL to be more teacher-friendly. First, teachers should be given more control over the game and they should feel a sense of ownership toward the game. Second, the games should be made more flexible, allowing educators to modify and adapt them. The third improvement is the introduction of more detailed reports on the gameplays through LA that will provide teachers with an insight into the appropriateness of the assessment regime used in the game and their students' learning outcomes. It would be optimal to look at all three improvements from a general point of view, addressing assessment integration as well as all other facets of a game such as content integration, story line, graphics and sounds. However, this paper focuses on assessment as it is central in the learning and teaching process.

The aim of this study is to develop and evaluate an assessment engine that would facilitate integrating these three improvements to educational games. In this paper, we present an assessment engine, EngAGe (an Engine for Assessment in Games), that is used by developers during the development of an educational game and it provides tools for educators after distribution of the game. Our approach is based on the externalisation of the assessment. The resulting modularity offers the possibility to modify the assessment logic via an online editor without interfering with the game mechanics and to retrieve information about the gameplays through an LA dashboard.

This paper is divided into five sections as follows. In Section 2, we present a summary of the literature on LA associated with educational games. In Section 3, we explain how EngAGe is used by educators, detailing the design for the assessment editor and the LA dashboard. In Section 4, we present the findings of an evaluation of the tool carried out with 31 educators. Finally, Section 5 draws conclusions and discusses future directions of our research.

2. Previous research

This section presents the findings of a literature review performed for this research and reviews the different approaches to using LA in GBL. No restriction was imposed on the dates of the papers, however, the oldest relevant study identified was published in 2011 reflecting how recent the topic of LA in games is. The following search terms were used: "*learning analytics*" AND *game.* The search was performed on 15 databases relevant to education, information technology and/or social science: ACM (Association for Computing Machinery), ASSIA (Applied Social Sciences Index and Abstracts), BioMed Central, Cambridge Journals Online, ChildData, Index to Theses, Oxford University Press (journals), Science Direct, EBSCO (consisting of Psychology and Behavioural Science, PsycINFO, SocINDEX, Library, Information Science and Technology Abstracts, CINAHL), ERIC (Education Resources Information Center), IngentaConnect, Infotrac (Expanded Academic ASAP), Emerald, Springer and IEEE (Institute of Electrical and Electronics Engineers) Computer Society Digital Library (CSDL). Relevant papers were identified based on two criteria: papers discussing learning analytics in games and papers presenting a framework for learning analytics in educational games. Papers presenting learning analytics outside of a game environment were excluded. Where possible, the search was based on abstract, titles and keywords to focus on relevant papers. A total of 364 papers were returned published between 2011 and 2016, 22 of these papers were relevant to this review, as summarised in Table 1. These papers are comprised of five book chapters, 14 conference papers and three journal papers. The studies presented in these papers differed in three main aspects: the data collected, the type of analysis applied, and the target users of the tool. These three aspects were categorised and are described in this section. Eleven of the relevant papers were used in real life situations and four showed empirical evidence of the usefulness of the system presented. None presented evidence of its usability.

2.1. Different types of data collected

The first obvious challenge to integrating learning analytics in educational games is deciding what data to collect. The literature review identified five types of useful data game developers and educationalists should consider monitoring when using GBL: time-related data, counts, game actions, scores and player data. These are described below.

- *Time-related data:* Some of the studies identified in the literature monitored data related to time. This can range from the total time spent on an activity [39] to the time the player took to perform a particular action or achieve a level, or the time of day the player played [14].
- *Counts:* Some of the systems monitored data in terms of numbers. In his paper, Duval [14] collected the number of logins and assignments finished while Holman et al. [26] also collected the number of content views and Greller et al. [20] the number of questions answered.
- *Game interactions/actions:* This type of data gives an insight into the player's actual interactions with the game. It can be very general, such as a player's state in a game [45] or more specific such as clicks or answers given to a question [20,35,45]. Piech et al. [39] even describe how they logged snapshots of code whenever a program was saved or compiled.
- *Scores:* The scores of a player are a very important and relevant measure. The performance of the player can be monitored [1,40,45] as well as the badges he/she earned [26]. Score can also be associated with time to visualise its evolution throughout the gameplay and across gameplays.
- *Player data:* In order to refine the data collected, it is useful to have information about the user. The information can be demographic (e.g. age, gender, language), academic [45] or technical with system configuration and session data being logged [18].

2.2. Types of data analysis

Once the data is collected, an analysis process is needed in order to transform it into useful information. There are two different techniques that could be used: Information Visualisation (IV) that describes the data and Data Mining (DM) that makes predictions based on more complex algorithms.

2.2.1. Information visualisation (IV)

According to Card et al. [7], IV is "*the use of computer-supported, interactive, visual representations of abstract data to amplify cognition*". Card [6] defines its aim with the following analogy: "*The purpose of information visualization is to amplify cognitive performance, not just create interesting pictures. Information visualizations should do for the mind what automobiles do for the feet*". IV is a tool for humans to draw conclusions about the data available and the visualisation process could be described in six key steps: (i) Mapping – how is information visually encoded? (ii) Selection – among the data available, what is relevant to the considered task? (iii) Presentation – how is the visualization laid out on the available screen space? (iv) Interactivity – what tools are provided to explore and rearrange the visualization? (v) Human factors – are human perceptions and cognitive capabilities being taken into account? (vi) Evaluation – has the effectiveness of the visualization been tested on users? [12].