Advances in Assessment of Students’ Intuitive Understanding of Physics through Gameplay Data

Mario M. Martinez-Garza, Vanderbilt University, Nashville, TN, USA
Douglas Clark, Vanderbilt University, Nashville, TN, USA
Brian Nelson, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Temple, AZ, USA

ABSTRACT

In this paper, the authors present advances in analyzing gameplay data as evidence of learning outcomes using computational methods of statistical analysis. These analyses were performed on data gathered from the SURGE learning environment (Martinez-Garza, Clark, & Nelson, 2010). SURGE is a digital game designed to help students articulate their intuitive concepts of motion physics and organize them toward a more normative scientific understanding. Various recurring issues of assessment, which pervade assessment of learning in games more generally, prompted the authors to consider whether gameplay (actions of learners in the context of the game) can be analyzed to produce evidence of learning. The authors describe their approach to the analysis of gameplay in terms of qualitative assessment that the authors believe may lay the groundwork for the application of similar computationally-intensive techniques in other educational game contexts.

Keywords: Assessment, Game Play Data, Hidden Markov Modeling, Science Learning, Sequential Pattern Analysis, Qualitative Analysis, Video Games

INTRODUCTION

The goal of this paper is to describe the development of novel approaches to assessment of learning in games. Often, researchers that design experiments around games for learning must rely on post hoc instruments to measure the progress that students make. This approach, while necessary and fruitful, does not leverage the potentially rich store of evidence that students provide about their own learning while they play. While the means exist to collect complete records of the actions and decisions that learners make while they play, no widely-accepted techniques or tools for making sense of this data stream currently exist. In this paper,

DOI: 10.4018/ijgcms.2013100101
we present our initial forays into analysis of game play using data-driven statistical and visualization techniques, and provide examples using data from the SURGE project (Martinez-Garza, Clark, & Nelson, 2011). Furthermore, we provide a rationale and framework that for the use of these techniques and argue for their appropriateness and applicability in other contexts of educational games research.

THEORETICAL FRAMEWORK

The potential of video games to support science learning is generally agreed upon (Gee, 2007; Mayo, 2009; Squire et al., 2003), but the analysis and structuring of evidence for game-based learning remains a challenge. This, in turn, has supported a mixed view of the effectiveness of games as tools for learning (Foster & Mishra, 2008; O’Neil, Wainess, & Baker, 2005). We believe, however, that this conclusion may be premature. The past fifteen years have seen great advances both in the sophistication of game designs and also in the supporting technology; there simply has not been enough time for a commensurate evolution in appropriate research methods. One central methodological difficulty involves capturing and measuring game-induced learning, which tends to be strongly situated within the game context, in out-of-game contexts such as post-tests. More advanced game designs compound this problem by supporting complex player actions that are challenging for learners to summarize and express, difficult for instruments to reliably capture, and resistant to conventional analytical methods. In addition, the use of formal assessments alongside games can compromise a game’s capacity for engagement and immersion, thus potentially reducing the efficacy of both the learning experience and the assessment.

The use of assessments of learning which reside outside a game used to measure learning that happens inside a game presents issues and vulnerabilities that merit careful consideration. Assessment is, after all, not a neutral activity. All assessments carry assumptions about the nature of learning, the nature of knowledge, and the purpose of assessment itself (Willis, 1993). The action of assessment places premiums on certain forms of knowing and understanding while de-emphasizing others. In the case of games for learning science, for example, an assessment may privilege declarative forms of knowledge, e.g. definitions and abstract principles, while the game itself might be more productive in reinforcing tacit knowledge or qualitative understanding of relationships. This insight becomes even more salient given the contrast between different types of games for learning: those in which the curriculum concepts are embedded in the game environment in a manner such that the game environment is structured mainly as context (“conceptually-embedded” games) and those in which the material to be learned is integrated into the core game-play mechanics with which the player is in constant interaction (“conceptually-integrated” games) (Martinez-Garza, Clark, & Nelson, 2012). It follows that these two kinds of games would favor different assessment strategies, given the differences in how they engage the learner, how they gauge success in the game, and how they represent knowledge. These nuances are not necessarily well captured by traditional assessments of learning, which traditionally favor summative declarations of concepts, articulated in discipline-specific forms and language (Sutton, 1996; Fang, Lamme & Pringle, 2010).

Other researchers have expressed similar views about the shortcomings of external assessments in capturing learning that happens in games and interactive media settings in general. De Jong & Van Joolingen (1998) argued that one of the difficulties that research has in achieving unequivocal findings in favor of the learning outcomes of unstructured learning environments lies in data interpretation. Quite often in educational games studies, researchers rely exclusively on outcome measures because there are simply no developed frameworks for interpreting and evaluating process data. This
Creating Virtual Alter Egos or Superheroines? Gamers’ Strategies of Avatar Creation in Terms of Gender and Sex
[www.igi-global.com/article/creating-virtual-alter-egos-superheroines/3955?camid=4v1a](www.igi-global.com/article/creating-virtual-alter-egos-superheroines/3955?camid=4v1a)

BioShock and the Ghost of Ayn Rand: Universal Learning and Tacit Knowledge in Contemporary Video Games
Chris Richardson and Mike Elrod (2016). *Contemporary Research on Intertextuality in Video Games* (pp. 92-107).
[www.igi-global.com/chapter/bioshock-and-the-ghost-of-ayn-rand/157027?camid=4v1a](www.igi-global.com/chapter/bioshock-and-the-ghost-of-ayn-rand/157027?camid=4v1a)