Towards a Process to Integrate Learning Analytics and Evidence-Centered Design for Game-based Assessment

Yoon Jeon Kim, José A. Ruipérez-Valiente, Philip Tan, Louisa Rosenheck, and Eric Klopfer
Massachusetts Institute of Technology
{yjk7, jruipere, philip, louisa, klopfer}@mit.edu

ABSTRACT: To fully leverage data-driven approaches for measuring learning in complex and interactive game environments, the field needs to develop methods to coherently integrate learning analytics (LA) throughout the design, development, and evaluation processes to overcome the downfalls of a purely data approach. In this paper, we introduce a process that weaves three distinctive disciplines together--assessment science, game design, and learning analytics--for the purpose of creating digital games for educational assessment.

Keywords: Evidence-Centered Design, Learning Analytics, Game Learning

1 INTRODUCTION

Digital games are gaining popularity in educational settings in addition to playing an essential role in young people's everyday lives. To fully understand what young people learn from playing these games and how they do so, the educational research community has developed ways to unobtrusively measure students' learning using log data generated from their engaging and authentic experiences in the game itself (Shute, 2011). This approach, called stealth assessment, promises that robust interferences can be made without interrupting the flow of gameplay, while at the same time reducing learners' anxieties about assessment (Kim & Shute, 2015). Evidence-centered design (ECD) is a framework commonly used by assessment designers to establish coherence across all aspects of stealth assessment development. Because much of the ECD effort is focused on formalizing assessment models (e.g. competency models), this can make the process less iterative. On the other hand, pure data-driven modeling approaches often found in fields like LA can be subject to bias, when they forget the human nature of the field. By integrating LA with ECD, the application of the assessment design process can avoid these pitfalls by intentionally incorporating expert-informed design decisions. We have been applying this process in our game-based assessment project called Shadowspect (see Figure 1 and a video online), which aims to measure common core geometry standards (e.g. visualize relationships between 2D and 3D objects) and relevant reasoning skills (e.g. spatial reasoning).

2 OVERVIEW OF THE PROCESS

The process describes the steps for game designers and developers to ensure that the assessment needs are well-balanced with the goal of a playful and engaging gameplay experience. Figure 1 on the right illustrates the three iterative phases of this framework—design, development, and evaluation—and shows how the LA, game design, and assessment disciplines inform one another to build coherence across all aspects of game-based assessment. Next, we provide a short description of each phase:

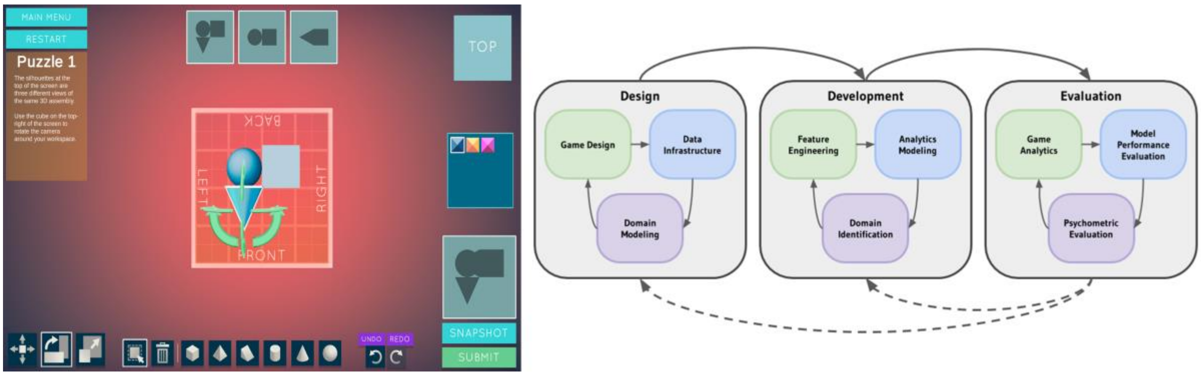

Figure 1. Screen capture of the Shadowspect game and overview of the process.

1. **Design**: The first step for design is to accomplish domain modelling by reviewing existing literature to further define what skills, knowledge, and attributes constitute the competencies that we want to measure. After the game designer has a reasonable understanding of the target competency (i.e. the competency model) and what evidence of that competency might look like (i.e. the evidence model), then the designer is able to explore the design space to come up with game mechanics that are compatible with the assessment mechanic. Additionally, this phase can include design of a data infrastructure that can accommodate the game events, algorithmic and assessment machinery, and reporting tools for the instructor. This can be a challenge due to the open-ended nature of game environments.
2. **Development**: The first step in developing assessment machinery is to work on feature engineering to create variables related to the target competencies. This is a step that entangles a mix of domain and analytics expertise, thus it is good to be carried out by a multidisciplinary team. Then, we use those features as the input of our analytics modeling. Where traditional ECD applies fixed rules based on human experts by tightly controlling the game elements for each target competency, this step can be improved with LA to discover learners' attributes or behaviors that are related to the competencies by generating automatic scoring rules or applying machine learning models. Then, we start mapping which evidence of the game and algorithmic outputs are linked to the domain that we aim to evaluate.
3. **Evaluation**: The last phase is evaluation of the analytic model in terms of both construct validity as well as performance metrics, and then embedding it within the assessment machinery. As part of this process, we also want to invest time in evaluating our game through analytics, so that we can identify game elements that could directly affect the psychometrics of the assessment and further iterate ways to improve player experience (fun) while using that data to identify "random patterns" or "off-task behaviors" that are the product of game design flaws. Finally, once evaluated algorithms and analytics models are incorporated as part of the assessment machinery, we then need to evaluate the psychometric qualities of the overall game-based assessment regarding generalizability, reliability, validity, and fairness (AERA, 1999).

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*

Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education*, *87*, 340-356.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer games and instruction*, *55*(2), 503-524.