Serious Assessments in Serious Games

Robert Hubal, RTI International, Research Triangle Park, NC, USA
Jamie Pina, RTI International, Research Triangle Park, NC, USA

ABSTRACT

In the past decade the serious games initiative has produced a number of games being used today as adjuncts to institutional training. Many of the games produced through the serious games initiative do not incorporate adaptive assessment to assure the acquisition of real-world skills from virtual training. Detailed performance assessment techniques are often lacking. Successful performance assessment in virtual training requires that students are placed into multiple simulated contexts and are challenged with different tasks to perform under various conditions, to specified standards. The aggregate of training situations must adequately cover the space of real-world situations. This article discusses important concepts related to virtual training performance assessment including critical tasks, performance criteria, forward recommendation, errors of commission, red screen alerts, reusable competency definitions, ill-structured domains, violent domains, and other game design templating. By integrating these concepts into serious game design, trainers can develop sufficiently varying simulated tasks to ensure that serious games meet real-world adaptive needs.

Keywords: Performance Assessment, Serious Games, Simulation Training, Situated Assessment, Skills Competencies

INTRODUCTION

The intent of this article is to provide strategies for a game development expert audience to improve how student performance is assessed in serious games. We adopt a perspective of simulation training. In simulation training, the goals are for the student to gain, practice, and demonstrate skills, and transfer skills to real-world situations, ensuring that there is ‘coverage’ over the range of situations over which to apply skills. Methods used during simulation training include a range of presentation techniques (everything from animated caricatures to true virtual reality), but our focus—hence the lessons learned in this article—has been to develop and present on desktop (non-gaming, non-immersive) systems.

Serious games are very different from simulation training. The point of simulation training is not for the user to have fun; the point is to engage the user in the learning situations. Consequently, users of training are not termed “players”; they are “students”. A game is defined as “a physical or mental contest, played according to specific rules, with the goal of amusing or rewarding the participant,” and with no specific intent to train (Zyda, 2005). Though simulations are commonly rendered via a game engine, and simulation developers use similar
tools as game developers, simulated training has a direct and specific outcome. Simulations may be enjoyable, and may need a storyline and game-like play, but those features are in essence secondary to the primary goal of learning.

Despite this distinction, many concepts involved in the study of game design can improve serious games. Such concepts include narrative (Riedl & Young, 2010; Rowe, Shores, Mott, & Lester, 2010); theme and thematics (e.g., background sounds that influence the ‘ambiance’ that could be critical to learning); use of all senses (Chalmers & Debattista, 2009); and better measures of engagement including increasingly common physiological measures (Liu, Agrawal, Sarkar, & Chen, 2009; Nasoz, Alvarez, Lisetti, & Finkelstein, 2003). It only makes sense for the simulation trainer to take advantage of existing gaming content, capacity, and experience in developing serious applications.

Just as our experience has been strengthened by the work of game designers, their work may be improved by our experience in simulation training development. With a specific focus on assessment, and even more specifically on performance assessment, we present lessons learned from our work in simulation training.

PERFORMANCE ASSESSMENT

Clinical procedures require accurate execution and timing to safely administer care. The human error that occurs in clinical treatment settings is referred to as medical error (Zhang, Patel, & Johnson, 2002), and can be reduced through targeted training (Dror, 2011). Performance assessment within simulation training for medical procedures can assist students in identifying weaknesses in their skills and provide the opportunity for corrective training.

As one example, we developed a simulator for clinicians to practice interacting with patients who may have been exposed to a bioterrorist agent (Kizakevich, Lux, Duncan, Guinn, & McCartney, 2003)—a rare but critical event. Through a systematic process, clinicians interact with the patient to elicit information about the present illness, past medical history, and lifestyle and medical risks, order diagnostic tests, make diagnostic hypotheses, and plan prescriptions, follow-up, and referrals. Along the way a clinician might fail to ask important questions or order laboratory tests that could contribute to a differential diagnosis.

Performance assessment involves the evaluation of students’ learned skills (Lampton, Bliss, & Morris, 2002). For the purposes of this article, students’ knowledge, including facts, concepts, rules, and policies, are measured only as residual artifacts of their ability to accomplish skills. The approach is to do so in a situated environment, putting the students into a situation and monitoring their activity. It is a kind of learning-by-doing (Aldrich, 2005; Frank et al., 2004). For assessment to be actionable, at least these questions need to be answered: What does the student know how to do? What can the student do in what context? Into what situation does the student go next?

In simulated situations, the idea is to push students to address their misconceptions and misperceptions that adversely influence performance of skills (Hubal & Frank, 2008). A situation is presented where something is faulty, the instructor knows what that “something” is, and the simulation can monitor the student’s activity to gauge if the student also identifies the fault. In the bioterrorism example, that something is medical symptoms with unknown cause, possibly bioterrorist, and the student’s subjective and objective assessment to form a diagnosis and plan. But these faults are not meant to be obvious to the student, often enticing the student to make mistakes while trying to perform the task. Dynamic measures are taken of student performance during the critical components of the task. Given the student’s performance, subsequent situations are presented with new faults.

In these situated assessments, the assessment moves away from the non-interactive (e.g., surveys, where most activity involves branching based on multiple choice responses) and the non-distributive (i.e., tasks requiring hands-on performance, which is very important to assess but not typically in a simulated environment).
Understanding Games Through Complexity Thinking Approach
[www.igi-global.com/article/understanding-games-through-complexity-thinking-approach/214860?camid=4v1a](www.igi-global.com/article/understanding-games-through-complexity-thinking-approach/214860?camid=4v1a)

Assessing Past, Present, and Future Interactions with Virtual Patients
[www.igi-global.com/article/assessing-past-present-future-interactions/74791?camid=4v1a](www.igi-global.com/article/assessing-past-present-future-interactions/74791?camid=4v1a)