Chapter 8

INFORMATION TRAILS: IN-PROCESS ASSESSMENT OF GAME-BASED LEARNING

Christian Sebastian Loh, Ph.D.
Virtual Environment Lab (V-Lab)
Dept. of Curriculum & Instruction
Southern Illinois University Carbondale, USA / csloh@siu.edu

Abstract: Assessment is an issue that is important to educators because without it, there is no way of telling if learners have arrived at the intended destination. Conducting assessments with game-based learning requires new tools and research methodologies because traditional face-to-face techniques do not transfer well into the multi-user virtual environments found in game worlds. Formative assessment is more useful to instructors and learners because it provides multi-point feedback to help them self-reflect and improve on what they are doing. This chapter describes an 'all-rounded' assessment system for game-based learning, which take into consideration the needs of multiple consumers (of information), namely, the Administrators, the Trainers or Educators, and the Learners. The assessment system made use of *Information Trails*© – a viable assessment methodology to collect user-generated action data as the game-based learning is occurring (hence, an *in-process* assessment). The data collected are then remotely transmitted to an external data storage using telemetry, and displayed in real-time via a data visualization application known as, *Performance Tracing Report Assistant*© (or, *PeTRA*). The final online assessment report can be tailored differently according to individual needs of the Learners, Trainers, and Administrators.

Key words: Formative, In-process Assessment, Telemetry, Data Visualization, Virtual Environments, Information Trails

1. INTRODUCTION

Imagine the following scenario: A large game development company contacted a local high school about the opportunity to beta-test a new digital game pertaining to leadership training (21^{st} century skills) for the seniors. The school administrator, the teachers, and the seniors are all excited about the possibilities offered by the game and are interested to know if game-based learning is indeed as effective as hyped. The seniors were asked to put in about 30 hours of game play in order to give the game enough time to 'work' its magic. Eager to see game-based learning in action, the teachers and administrator agreed that a third of that time should take place in the school computer lab under the teachers' supervision. The Non-Disclosure Agreement was signed and all went well.

After three months or so (10-12 weeks), the project was concluded amidst much fanfare, but many had questions about the outcomes. Besides feeling great, is it possible for the students to evaluate their own success objectively? How can the teachers ascertain if 30 hours of gameplay (as recommended by the game company) is adequate to acquire the skills taught in the game?

Extracurricular activities, such as the Future Business Leaders of America (FBLA) and National Honor Society, have been the venue for student leadership training; how well would the game compare to these traditional approaches? Some teachers were wondering if there is a way to know which of the classes performed better and considered contacting the game company for a breakdown of the records. Should they even bother? Are such records being kept at all? How would the school administrators document the effectiveness of the game in a report for next month's Parent-Teacher Association (PTA) meeting?

1.1 Who is asking the question?

"How do we assess the effectiveness of game-based learning?" is obviously the big question that is begging to be answered. But before we proceed to discuss the implications of that question (as will be dealt with in the rest of the chapter), let us consider first, who is asking the question?

While it is natural to focus on the play-learners (the high school seniors, in our scenario) as the target audience in a discussion about assessments for game-based learning, we need to recognize that these learners are not necessarily the 'customers' of the serious games. We counted at least three different user-groups of game-based learning, and each came with their own agenda. In fact, out of the three user-groups, the learners are probably the ones with the least interest about assessment of game-based learning. We will examine who these user-groups are and what added values game-based learning will bring for them:

1. The first user-group consists of the *Learners*, who are the primary target of game-based learning. They are the ones who will have firsthand experience with the game and are supposed to benefit most from its usage. As such, the learners need to have a sense of what goals they have achieved (over time spent) throughout the learning process. Information that is useful to this group of users includes keeping scores on the number of outstanding and completed learning goals, time taken to complete certain levels of learning, total time spent in the learning environment per day/week/semester and bottlenecks (where they may be 'stuck' or killed in the game). Such information needs to be made available to the learner in the form of a simple report for self-evaluation purposes.

   While the Learners are indeed the primary 'consumers' of game, they are not necessarily the 'customers' of game-based learning applications; meaning, the Learners are not the selectors and purchasers of these resources. These learning applications are "often chosen or paid for indirectly by program sponsors, not the participants themselves" (Aldrich, 2009, p. 15). Whereas, in the digital games for entertainment market, the purchasers are the one who tend to be using the games.
2. The second user-group is made up of the *Trainers*. They are the assessors of the game-based learning and have immediate supervisory role over the learners. These are school teachers in our scenario, but could easily be instructors or supervisors in the business training industries, or sergeants in the military. Because of their responsibilities over the learners, they need to keep track of what is happening in the virtual game environments and monitor the learners' activities to ensure the learners are on-task. As assessors, they require a means to easily visualize the learners' data, both individually and en masse.

   Some kind of software-based reporting is necessary for the assessors to monitor the learning progress of the learners, track the number of objectives met, identify mistakes made by the trainees, and allow for appropriate remediation to be prescribed in a timely manner. This means that the report should ideally reflect real-time data and not an 'after action report' made available only after the game is completed (three months later according to our scenario). Data visualization functions are very important to the Trainers group because outlier(s) – i.e., learner(s) who are behaving differently from the expected norm – must be spotted as early as possible. A real-time report would empower the Trainers to take action early enough in the training cycle to alert the learner(s) of their situation, evaluate said action (that is out of the norm), and correct that action via remediation (or not), before the mistakes become entrenched.

3. The *Administrators* made up the third user-group. They may be sponsor(s) of the game-based learning or the reporting officer(s) situated above the Trainers in the organization chart, or both. In very large organizations, there may be more than one level of administrators. In a military context, for example, the administrators could be the commanders of a large scale joint-exercise. In our scenario, the PTA and district superintendent may also be included in this user-group.

   This group of users is usually less interested in individual performance reports about the learners. Instead, as sponsors, they are most concerned with the Benefit Cost Ratio (BCR) of the game-based learning. In other words, from an investment point of view, the Administrators are the purchasers, and they want to make sure that the game-based learning products actually ‘deliver.’ Sometimes, performance data of Trainers may also be of interest to the Administrators. Trainers’ data that is associated with Learner achievements can be used to show additional efforts put in by the trainers, and to determine which Learning Center is out-performing others. Let’s say a certain school was found to have the best achievement score among others in the same district after a certain multiplayer online game for learning was implemented. The superintendent may be interested to find out if this particular school had used a different approach to raise achievement scores. All this information should be presented in some kind of intelligent online assessment report, capable of highlighting the weaknesses, strengths, accomplishments, potentials for improvement, and may need to be sortable by trainers, learning centers, and other filters.

While all three user-groups, Learners, Trainers and Administrators, benefit from the addition of a powerful assessment reporting system in game-based learning, no such assessment system exists (to the best of the author’s knowledge) at the time of the writing. Obviously, as long as the needs of the customers of game-based learning are not being satisfied, the demand for assessment of game-based learning will continue to grow.

2. ASSESSMENTS AND GAME-BASED LEARNING

In education, assessment is regarded as an important and integral part of the learning process. If learning is likened to a journey, then textbooks, classroom teaching, e-learning, games and simulations are the vehicles that deliver the learners from starting point A to end point B. From an Administrator’s point of view, assessment is the quality assurance protocol that ensures the learners have indeed arrived at the correct destination – i.e., achieving the stipulated learning goals based on the benefit/cost negotiated. Learning activities without an assessment component are informal and similar to the endeavors of hobbyists, at best.

Proponents of game-based learning have asserted this to be a highly suitable medium to impart 21st century skills to the gamer generation (see Aldrich, 2009; Gee, 2007; Gibson, Aldrich, Prensky, 2006; van Eck, 2006). This lead some people to perceive game-based learning as a 21st century approach to learning brought on by digital technology. However, Botturi and Loh (2008) found many ancient ties between game playing and learning and suggested that game-based learning is just a new approach to revive an ancient tradition. Some parallels between game playing and learning persist even today: e.g., school principals were regarded as 'game masters' of the arena by the ancients.

Unfortunately, digital games are not all created equal and are, therefore, not all suitable for learning. As Chen and Michael (2005) noted, the inclusion of assessment components appears to be the main difference distinguishing the more 'serious' games from the rest that were created for entertainment. Sans the requirement for learners to demonstrate the 'abilities' they have acquired from the course of instruction (Joosten-ten Brinke, Gorrisen, Latour, 2005), there are no means of knowing if the learners have indeed 'arrived' at the learning destinations.

Outside education and research communities, game-based learning has also received acclaim from the business industries and training sectors (e.g., Aldrich, 2009; Kapp and O'Driscoll, 2010). However, the appeal of serious games and game-based learning to these industries is not so much in the ability to automate training tasks (as do other computer-based instructions), but to co-locate massive numbers of trainees simultaneously to mitigate the high costs typically associated with training (e.g., Duffy, 1997; Wilson *et al.*, 2008). As the military, large corporations and institutions of higher learning implement large-scale virtual environments for training and e-learning, the demand for formalized assessments with game-based learning is sure to increase.

2.1 Two types of assessments

Newcomers to the games and simulations research will probably be overwhelmed by the massive body of literature (see Hays, 2005; O'Neil & Robertson, 1992; Tobias & Fletcher, 2007) covering all sorts of issues from design, graphics, mode of delivery, narratives, theories, and philosophies to their potential uses for learning. Although some researchers are currently working to address the need for assessment in game-based learning (e.g., Rupp, Gushta, Mislevy, & Shaffer, 2010), more effort is needed to fill the gap.

Educator-researchers refer to two different kinds of assessment: summative (assessment for learning) and formative (assessment of learning). Summative assessment is typically conducted towards the end of a course of instruction because it is designed to test a person's understanding, retention, or mastery of the subject after a course completion. The After Action Reports (AAR) used by the military are a prime example of summative assessment. Formative assessment *of* learning, on the other hand, is designed to measure the amount of learning that is still taking place while the course of instruction is on-going, and that assessment can occur as many times as deemed necessary by the trainer or instructor.

When taken as a conterminous process, assessment of learning is actually more useful to educators than summative assessment because it helps them fine tune the instructional and learning processes. Feedback, often cited in formative assessment research, has been found to be the single most powerful influence in learning improvement (see Black & William 2009, 1998; Hattie, 1987). Likewise, in peer- and self-assessment (used as formative assessment) among students of online collaborative learning environments, the feedback received at multiple points over the learning process has been shown to provide students with self-reflection on the learning process, help them identify areas for improvement, and take ownership of their own learning (Lee, 2008). It is clear that an effective formative assessment component would benefit not only the instructors, but also the learners, in an interactive online learning environment such as game-based learning (see Figure 1).

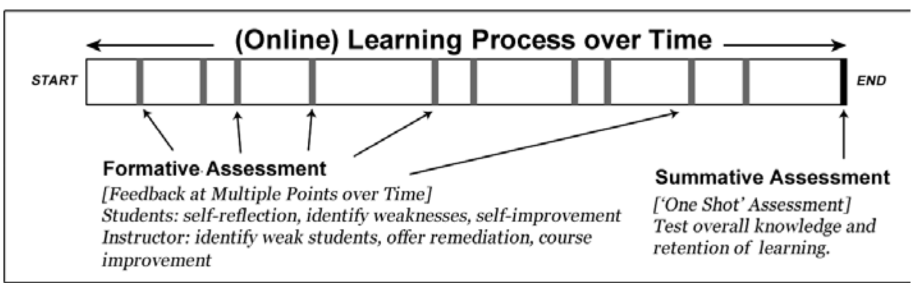

Figure 1: Overview of formative and summative assessments in a learning environment

2.2 Assessing game-based learning: The issues

When both instructor and learners are face-to-face, an instructor can directly observe the learners' physical behaviors as evidence of learning and participation (Harrington, Meisels, McMahon, Dichtelmiller & Jablon, 1997). Traditional assessment metrics such as test scores, classroom participation, and time-on-task were originally crafted to take advantage of the simultaneous presence of both the trainers and trainees at one physical location.

The situation changes dramatically when trainers are no longer able to 'see' learners face-to-face. (Some online learning applications attempt to overcome this problem by allowing trainers to 'see' the learners using web-cams and video streaming technology.) Until there is a safe way to put probes into the minds of learners to directly measure the amount of learning that occurs, trainers must rely on external measures for assessment. Although some online learning environments allow students to virtually 'raise their hands' to ask questions during online lessons, other direct observational measurements of human actions, behaviors, and expressions still prove to be difficult. This means that researchers in the field must create new tools to collect better data. This area of research is obviously still in its infancy as current literature is equivocal about how best to conduct assessment with game-based learning. There also appear to be more problems than available solutions at this juncture. For instance:

a) Without properly designed games, there will be nothing to assess with. Should educators create new games from the ground up with commercial game engines (i.e., the industry model), or modify existing commercial games using development kits (i.e., the 'grass root' model)? Current game development models used by the game industry tend to exclude teachers' inputs. Cheaper and easy-to-use game development tools are in order, as are game development models that are suitable for use by educators (see Younis, & Loh, 2010).
b) Because many of the known traditional assessment methodologies are not directly useable within virtual environments, researchers may need to search for new, effective, and meaningful ways to conduct assessments with game-based learning. Traditional statistical methods are not as effective compared to educational data mining (EDM) in dealing with massive amounts of data obtainable from online learning environments. New assessment and data analysis methods are both in demand.
c) Based on the criticism that combining assessment with games can severely interrupt 'flow' (Csikszentmihalyi, 1991) and render the game 'not fun to play' (Prensky, 2001), some researchers have proposed

workarounds through ‘stealth assessment’ (e.g., Shute & Spector, 2008; Shute *et al*., 2010). However, others have reported that the effects appeared negligible (e.g., Reese, 2010). More research is needed in this area.

d) Some researchers see game-based learning as an extension of e-learning and suggest that the assessment component should be integrated into a Learning Management System (LMS) that is compliant with SCORM (i.e., Sharable Content Object Reference Model) (Moreno-Ger *et al*., 2008). Others see this type of learning to be digital games with instructional intent, and they should therefore, have the feel of ‘real games’ approximating commercial production quality (van Eck, 2006). If so, then the assessment component ought to be integrated into the game engine (e.g., Loh, Anantachai, Byun, & Lenox, 2007), and not reside within an LMS. Is there a third, or even a fourth, approach to resolving this issue?
e) Current understanding of game-based learning is built upon summative assessment studies conducted after training has been completed. Researchers need to move out of their comfort zones and begin looking into the development of formative assessments that take place throughout game-based learning (e.g., Loh & Byun, 2009; Reese, 2010).

The list of issues goes on.

2.3 Measuring performance in virtual environment

In today’s workplace, be it virtual or physical, performance improvement has much to do with waste reduction and output increase. While many work incidents could indeed contribute to ‘waste’ and require reduction, one of the worst types of waste is ‘habitual man-made mistakes’ because it costs the company twice as much to re-train workers to unlearn their mistakes. Moreover, as is the case of a recent study by the National Transportation Safety Board, flaws in flight simulators used to train airline pilots have been linked to more than half of 522 fatalities in U.S. airline accidents since year 2000 (Levin, 2010). Such flaws and mistakes – even when it was not directly the fault of the workers – result in losses for the company, both in terms of legal compensations and reputation. It is important, therefore, for trainers and trainees to strive to recognize human errors in tandem during training and rectify these mistakes before they have a chance to become entrenched and turned into costly errors.

Whereas physical training games such as basketball, javelin throwing, and sprinting, build up real muscles in the body and improve psychomotor skills; training with digital games is more suited for the building up of ‘brain muscles’ and cognitive thinking skills. Hence, advocates are calling for the development of more game-like environments that teach 21st century skills, which include leadership, project management, and negotiation skills (Aldrich, 2009; Prensky, 2006). Apparently, 'brain muscle' training in game-like environments is not unlike physical (muscle) training, as the core features in many serious games consisted of numerous "trials and errors and repetition of steps" (Saridaki, Mourlas, Gouscos, & Meimaris, 2007). Evidently, both physical and cognitive training games utilize regular practice and just-in-time feedback to 'strengthen' relevant muscle groups in the learners as they progress towards the learning goals.

Due to the amount of repetitive training and the number of trainees involved in some multi-user online (training) games, monitoring all the events that are happening would easily lead to trainer fatigue. Since it is deemed more cost effective to co-locate trainers with trainees in a one-to-many ratio, trainers will necessitate appropriate supports to better monitor trainees' actions en masse; especially if they are expected to detect deviations in the trainees' behaviors that could lead to habitual errors.

Bearing in mind that some game-based training may last as long as 20-40 hours (spread over several weeks), unchecked errors have the good possibility of becoming entrenched through reinforcement. The greater the potentials of an online multiplayer training millions of trainees simultaneously, the greater the risk; as even one small error can quickly accelerate to reach critical mass. Therefore, besides presenting appropriately designed contents for learning, a good, game-based training must also support formative assessments that are targeted at both the instructors as well as the trainees, for all the reasons and the learning supports mentioned in earlier sections.

3. GATHERING EMPIRICAL DATA

In commercial game development, once a game is completed, it is quickly turned into profit. Very few developers would actually be interested in '*in-process*' data collection unless it somehow contributed to the usability of their games (which might, in turn, affect overall profits). In this chapter, the term 'in-process assessment' is used specifically to refer to an ongoing formative assessment conducted throughout the game-based learning while the game session is ongoing.

Adding an assessment components to serious games (mentioned by Chen & Michaels, 2005) would constitute additional work for the programmers, who must be paid. Game developers see assessment components in games as an additional cost overhead that undercut their profit margins. Unless developers knew beforehand about how to recuperate the costs, they would be reluctant to invest in the creation of an assessment component, much less to integrate one into a game engine. This might explain why there have been very few games created with assessment components, despite high interest among the game-based learning community for them. Fortunately, the tide began to turn after Georg Zoeller presented on ‘developer-facing telemetry for games’ at the Game Developer Conference (GDC) 2010. [More information can be found at http://gdc.gulbsoft.org/talk]

3.1 Telemetry

The American Heritage Dictionary defines ‘telemetry’ (n.d.) as “The science and technology of automatic measurement and transmission of data by radio or other means from remote sources to receiving stations for recording and analysis.” In simpler terms, telemetry is a technological process that allows remote data collection and information retrieval. Since telemetry’s origin in the 19^{th} century, it has been used by many industries, including the medical field, law enforcement, wildlife research, space exploration, motor racing, and traffic control. In many cases, the objects of interest were tagged with technological devices that allowed remote tracking and the data collected by these devices were compiled into metrics, which were then remotely sent back to the researcher for recording and analysis. A ‘developer-facing telemetry’ suggested that the results of the analysis were meant for developers’ (and not gamers’) consumption. Based on our discussions, assessments for game-based learning can be said to comprised of ‘learner-,’ ‘trainer-,’ and ‘administrator-facing’ telemetries.

In his presentation, Zoeller (2010), a Lead Technical Designer of Bioware, disclosed how he had made use of a data collection server during the development of Dragon Age: Origins (2009) to track and reward developers’ activities and to collect in-process beta testers’ data for game balancing and design improvement. (The same telemetry is also employed in *Dragon Age 2* (2011), evidence of this can be found in the config.ini file.) The most difficult part of the telemetry to him was the ‘data visualization’ process: to convert the raw data into a humanly understandable format, to afford him a better understanding of the information, and to use it to steer game improvement. Traditional graphs are not useful for this type of analysis because of the unconventional data collected. New ways of data visualization are required and must be ‘invented’ accordingly on a scenario-by-scenario basis.

Some months later, Chris Pruett (2010), an independent game developer, published an article detailing how he had created a small-scale metrics reporting system for telemetry and used it to improve a mobile phone game he had developed. Once the data were obtained through telemetry, Pruett used the ‘heatmap’ technique (see Figure 2) to help him visualize the ‘bottlenecks’ in his mobile game and subsequently improve its game play via ‘balancing.’ Bottlenecks, in this case, meant areas that were too difficult for players (i.e., they died), and ‘balancing’ meant tweaking the game to provide players with better weapons, weaker bosses, more health potions, etc. (to help them overcome the bottlenecks). The overall intent was to provide gamers with a challenging, but enjoyable time instead of making them feel frustrated to the point of giving up.

We should recognize that game developers and academic researchers will both benefit from enhanced game engines, imbued with telemetric capabilities to track and report player and game events remotely. On one hand, the data analysis process provides insights to the game developers on how to improve the usability and design of their games. On the other hand, the game metrics are of value to researchers for the *in-process* assessment of game-based learning. The creation of game development tools integrated with telemetry could throw open the flood gates and help make assessments for game-based learning a reality in the future.

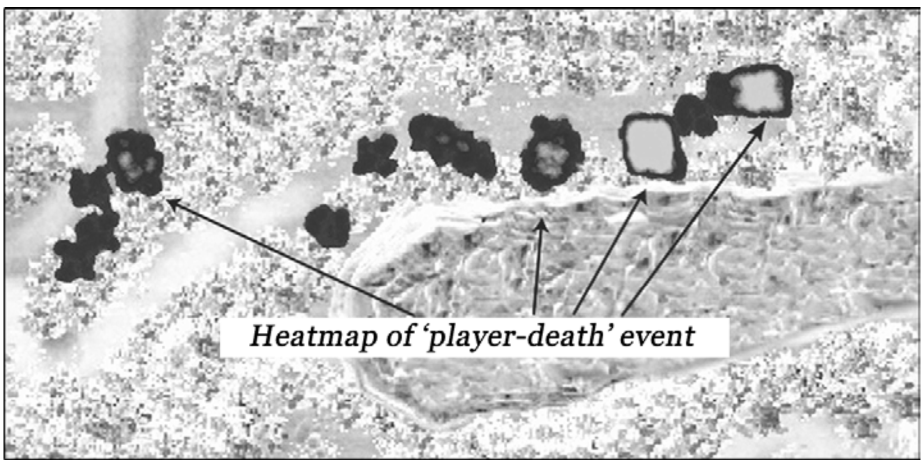

Figure 2: Example of a ‘heatmap’ showing the zones where ‘player-death’ event occurs *most frequently.* [Partial screenshot of *PeTRA*, used with permission.]

3.2 Psychophysiological Measurement

Researchers in the fields of psychology, cognition science, usability testing, and human-computer interface have had a long history in using automatic event loggings to study human (and animal) behaviors within interactive systems (e.g., Skinner, 1938). Since digital games are interactive systems, such psychometric methodologies have also proven useful for researchers in the field of 'gameplay experience' research. In these studies, players' reactions during gameplay are meticulously recorded and matched to game contents, using a combination of quantitative and qualitative approaches including: video recordings of gameplay sessions, interviews of attitudes, self-reports, and psychophysiological measurements (which graph emotional responses and states of arousal of players during gameplay).

Video game 'user experience' (UX) researchers believe that the combinations of data are indicative of the levels of 'flow' (Csikszentmihalyi, 1990) and engagement in the players. As such, the research findings can reveal how players perceive the game contents (as boring, engaging, fun, etc.). Such information is useful to the game publishers who can then decide to take advantage of the information (or not) to adjust and improve their products. Some of the psychophysiological measurement includes:

a) Measurement of skin electrical conductivity indicative of fear and excitement using Galvanic Skin Response (GSR) and Electro-Dermal Activity (EDA),
b) Measurement of brain wave patterns of players during game play using Electroencephalograms (EEG),
c) Measurement of cardiovascular activities (e.g., heart rate variability, beat per minutes) of players under different levels of excitement and fear using Electrocardiograms (ECG),
d) Measurement of facial muscle activities (e.g., smile, frown, etc.) during game play using Electromyography (EMG),
e) Measurement of pupil diameter under different emotional and arousal influence through pupillometry and eye-movement, and
f) Measurement and analysis of gaze directions upon the computer screen during game play using Electro-Oculograms (EOG).

However, as digital games grew in complexity, researchers have begun to voice the need for an integrated logging framework that would afford automatic psychometric data collection and make gameplay research easier (see Nacke, Lindley & Stellmach, 2008; Sasse, 2008). As expected, game publishers are slow to comply with the request; none have so far. While the inaction could be due to the additional cost incurred, or the failure to see a quick profit turn-around; it is also possible that the call is once again being perceived as another 'academic advice' (as mentioned in the previous section).

Since a person's motivation and engagement level can greatly impact learning, psychophysiological data can indeed be useful for the assessment of players' 'affective performance' in game-based learning. At the very least, psychophysiological data should be usable in conjunction with other assessment methods to triangulate research outcomes. In this early stage, only findings with first-person-shooters have been reported (e.g., Nacke, Grimshaw, Lindley, 2010; Nacke & Lindley, 2010; Nacke, Lindley & Stellmach, 2008). Would the psychophysiological measurement prove to be equally informative for other game genres, including: role playing games, strategy games, and massively multiplayer online games? The large gap in literature indicates that this field has a lot of potential for growth in the future.

3.3 The "black box" effect

In order to improve the process of learning and instruction, educators must constantly experiment with new methods of instruction and assess their effectiveness. Pretest-posttest experimentation is a common research method employed by educators in traditional classrooms to ascertain the effectiveness of untested instructional processes. On the first look, the pretest-posttest methodology may appear to be useful for the assessment of game-based learning (e.g., Kebritchi, 2008). Typically, two identical tests are administered, one before (pretest) and one after (posttest) a certain experimental method of instruction (i.e., intervention). Keeping other variables constant, the difference in achievement scores (Δ), i.e., posttest minus pretest ($t_2 - t_1$), may then be attributed to the improvement brought about by the intervention itself.

Even though the pretest-posttest method of inquiry can indeed demonstrate positive effects for game-based learning, it cannot fully explain which chain of events or sequence of actions performed by learners (in the game) actually contributed to those positive effects. In this sense, game-based learning remains an impenetrable "black box" because no one knows for sure how or why the intervention works (even if it does). Unless we educators quarantine learners individually, prevent them from speaking with one another, and restrict access of external learning materials, how can we be sure that the change in achievement scores (Δ) truly reveals the amount of learning gained?

Moreover, the "black box" effect renders the intervention vulnerable to external threats because it is impossible to identify if any external factor has entered the system and has affected the data collected. For example, there is no way to tell if trainees are trying to "game the system" – i.e., exploiting properties of the system to succeed in the environment rather than learning the materials as intended by the system designer (see Baker, Corbett, Roll, Koedinger, 2008). Naturally, the Administrator group could not allow the existence of a loophole as big as this within the system.

From the perspectives of the Trainers, the over-reliance on posttest results is unsettling also. By the time the (overall) effects of game-based learning can be determined via the *posttest*, it may be too late and too costly to re-train the Learners. While this problem is not immediately apparent in ‘clinical’ research studies that subject learners to only 1-2 hours of game play, the effect is amplified in commercial off-the-shelf (COTS) games that require much longer (20-40 hours) to complete. The inclusion of telemetry into serious games would be the first step in the right direction towards true assessments for game-based learning.

This does not mean that researchers should not use other methods to assess the effectiveness of game-based learning. For instance, qualitative analysis remains an important research methodology when we are looking for rich data involving small group of learners, or in case studies. Because the data collection processes for qualitative analysis often require long hours of video-recording, record-keeping, and meticulous audio transcriptions, it is not practical for the assessment of game-based learning involving a large population of learners. For every hour of game play which requires three times that amount of time to analyze, it would take far too long to analyze thousands of learners who have each accrued 20-40 hours of game-based learning. (For a longer treatise on this topic, see Loh, 2009.)

4. INTRODUCING INFORMATION TRAILS©

The research leading to *Information Trails* began with one supposition: if a person’s actions and behaviors are determined by his/her decision-making process, is it possible to break down (or, reverse engineer) the decision-making process based on the person’s actions and behaviors? The logic behind the supposition is very similar to a crime scene investigation in which a CSI agent tries to determine how a crime was committed based on the evidence found at the crime scene. If decisions are the products of a person’s knowledge schema, then it should be possible to express the effects of learners’ actions (e.g., speed, accuracy and strategy) in a learning environment as a function of their understanding of the learning problems versus their problem solving skills or abilities. Go down that road and substitute multiuser virtual environments (MUVE) for ‘learning environment,’ and the path will eventually lead to *Information Trails* (Loh, 2007). The only obstacle remaining is the ‘back box’ of game-based learning, which can be cracked open using telemetry.

Conceptually, *Information Trails* is a series of event markers deposited within any information ecology at certain intervals over a period of time. The event markers can later be retrieved from the information ecology for storage and data analysis. In practice, streams of user actions are automatically tracked and recorded at intervals, triggered by ‘event calls’ issued from the game engines over the entire course of game-based learning (see Figure 3). The detailed data collected can be used to visualize the most common paths taken by learners to reach certain learning goals, and may be used to compare a learner's problem-solving strategy against that of an expert's. Deviations from the normal route could either mean unusual approaches to reach learning goals or be indicative of misguided decisions leading to man-made errors.

With large amounts of data collected in massively multiplayer environments, hidden patterns of learner behaviors can be uncovered through EDM. It is then up to the trainers to decide what course of action should be taken to remediate or to correct the deviation. The framework has been successfully developed into several working prototypes through a series of funded research. Not surprisingly, the military was the first party to show some interest in the project. This explains why player-movement was the first feature to be investigated (and implemented). Besides military and business training, preliminary data suggest that *Information Trails* can also be used to trace learning within online virtual learning environments (VLE) for medical simulation/training, and virtual worlds.

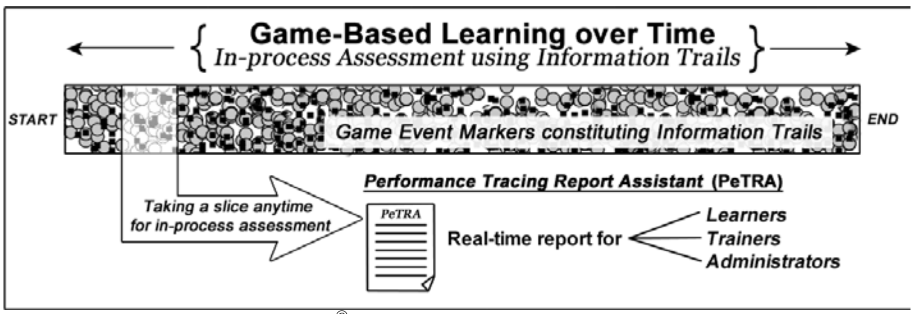

Figure 3: Information Trails©: In-process assessments for game-based learning.

4.1 From games to *Information Trails*

Debuted in 2002 and 2006, the commercial off-the-shelf (COTS) *NWN* and *NWN2* (produced by Bioware, and Obsidian Entertainment, respectively) were part of a series of third-person role-playing games published by Atari. The game has its origin as a pen-and-paper *Dungeons & Dragons* game set in a fictional world called *Faerun*, where men and other fantastic creatures (e.g., dwarves, elves, dragons, giants) inhabited the land. One unique feature that separates *NWNs* from many other COTS games is the included game development kit (GDK). With the GDK, gamers are given the authoring tool to create their own game modules/stories for sharing. This social game *mod*-ification practice was later named 'game modding' by the gamer community.

Although the default language of the game is English, it is fairly easy to modify the game's user interface into other languages, including Chinese. As a role-playing game, *NWNs* have great potential for use in the teaching of a foreign language. Educators who used *SecondLife* (*SL*) to create virtual environments for the teaching of foreign languages will, no doubt, find many similarities between modding in *NWNs* and rezzing in *SL* (Kaplan-Rakowski & Loh, 2010). Despite the medieval settings of the game environment in *NWNs*, it did not deter the U.S. military and NATO from adapting the game for training (Weil, *et al*, 2005) and research (Warren & Sutton, 2008).

Over the years, the modifiable game has steadily garnered a large group of followers; among them are many educators and researchers who have learned to 'mod' the game according to their needs. For example, some were created for scientific research (Gorniak & Roy, 2005), while others have been used to teach classroom learning subjects, ranging from journalism (Berger, 2006), to story writing (Robertson & Good, 2005) and mathematics skills (BBC News, 2007). Reader should note that *all* of the 'game modules' produced in this manner are standalone games, and are therefore, not directly assessable.

In order to create an *Information Trails* empowered game with assessment capability, *in-process* data collection (while the game session is ongoing) is a necessary step. User-generated data must first be retrieved from the game engine (as the game is being played) and then be stored apart from the game, in order to facilitate retrieval for data analysis (independent of the game). An event listener, *NWNX*, is employed to achieve the 'handshake' between the game engine and the remote/external database server (*MySQL*). The *NWNX* was originally created by *NWN* gamers (Stieger, 2008) to transform the standalone *NWN* into a server running online 'persistent worlds,' which are very early forms of massively multiplayer online games (MMOGs).

Once a communication channel between the game engine and database server is established, it is finally possible to transmit data for telemetry. The final step would be to create the online reporting interface according to the needs of the clients – be it Learners, Trainers, or Administrators. Since it would yield far too much data if we set out to capture every available event in the game, we have chosen to capture only a selected list of game events, using an Objective Hierarchy Map that ranked the events by importance according to game story development, and relevance to learning/training goals.

The game events currently being recorded include: conversations between players, players' death, players' spawn, players entering and exiting the game, items gained or lost, experience points gained by players, enemies killed, and learning goal(s) achieved. Movements of the players (as x-, y-coordinates) were recorded at regular intervals using the 'heartbeat script.'

As the name suggests, a heartbeat script is a script that is auto-fired by the game engine (just like a regular heartbeat).

4.2 An integrated assessment system for GBL

By leveraging the knowledge base amassed by the community of *NWN* 'game modders' (i.e., gamers who modify existing games for personal enjoyment), we were able to create the first working *Information Trails* system and showcased it at the 2008 international conference for Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational & Serious Games (CGAMES). Since then, we have continued to improve upon the user interface (UI) and the database engine of *Information Trails*, refining the workflow into a viable process for the tracking of user-generated action data in game-based learning using telemetry.

As Zoeller mentioned, the most difficult piece of the telemetry was data visualization. When we first began working on the data visualization of *Information Trails*, we had no idea what it would eventually look like, except that it should show the movement of the player's avatar graphically. After several iterations of product design and development, we believe we have succeeded in creating an online assessment system for game-based learning, comprised of a front-end for user-facing data collection (tracking), and a back-end for trainer-facing data analysis (reporting).

As shown in Figure 4, the *Information Trails* system is made up of several integrated components, including:

1. An online game with user authentication (to facilitate tracking of individual learners),
2. An event listener or a trigger for the data collection processes,
3. A database server to facilitate data collection and record keeping,
4. A component to visualize the data as useful information (in this case, *Performance Tracing Report Assistant*, or *PeTRA*), and
5. An optional game engine for in-house creation of game-based learning modules. (The making of the game may be out-sourced to commercial game development companies).

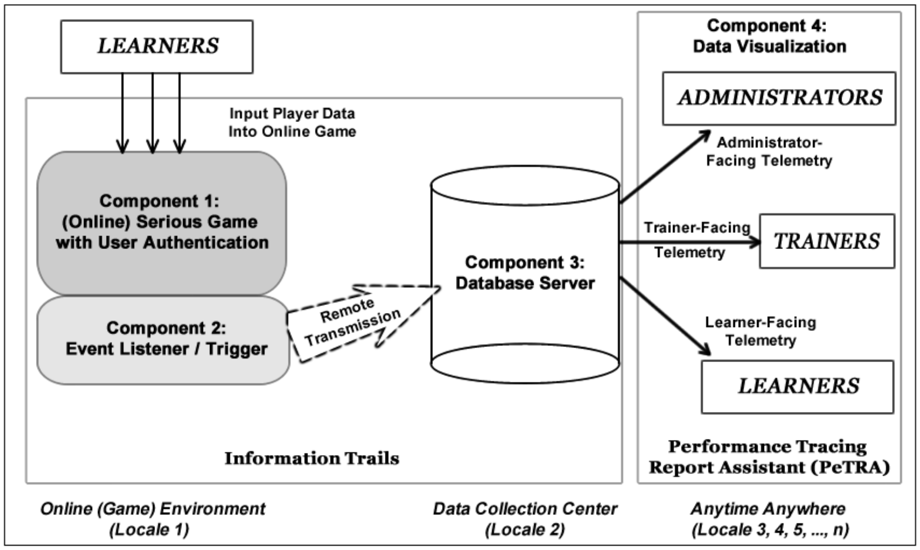

Figure 4: Components of *Information Trails* and their relationships with *Performance Tracing Report Assistant* (*PeTRA*)

4.3 Finding new ways to visualize data

After several attempts to visualize the collected user actions data, we finally settled on a bird's eye view of the area map. We used the game map to show positions of the learners within the game world over time as a series of connected dots. In the later versions of the report, we were able to overlay the path traversed on top of the area map, unlike in earlier versions (see Figure 5). The inclusion of the full-color area map was important to the trainers because the visual cues (i.e., the geographical layout) enabled them to understand the decisions behind the learners' actions (movements).

As soon as game-based learning begins, user-generated action data becomes available through *PeTRA*. An automated data recording, analysis, and visualization process is important to *Information Trails* (and possibly, assessment of game-based learning in general) because not all trainers are versed in handling vast amounts of data, or in interpreting what they mean. Trainers will appreciate not having to deal with the raw data in order to make sense of the information contained therein. The report also allows for the replay of users' actions in a step-by-step fashion for debriefing purposes. Time taken by learners to meet various learning objectives is reported and compared against that of the experts. In this case, as practice time increases, the time taken to complete a particular learning task is expected to decrease as the learners move towards mastery. In addition, since *PeTRA* is fully interactive, performing a ‘mouse-over’ above the dots will reveal the user actions that took place at that particular time and coordinates. Such real-time interactive features are highly suitable for software-based assessment (as in game-based learning) and are simply not possible with paper-based reports.

At this time, *PeTRA* is used mainly in debriefing, for Trainers to review and evaluate a player’s action in a game-based learning session. Future plans for *PeTRA* include creating a customizable interface to suit different user-groups, as well as looking into new ways to visualize the data collected as ‘useable information.’ For example, a learner should be able to access *PeTRA* at any time to review individual in-game actions and performance data for self-improvement. A trainer, on the other hand, may need to visualize the performance of multiple Learners’ who are under their charge. An administrator may only be interested in the performance index or the overall ROI of the game-based learning application.

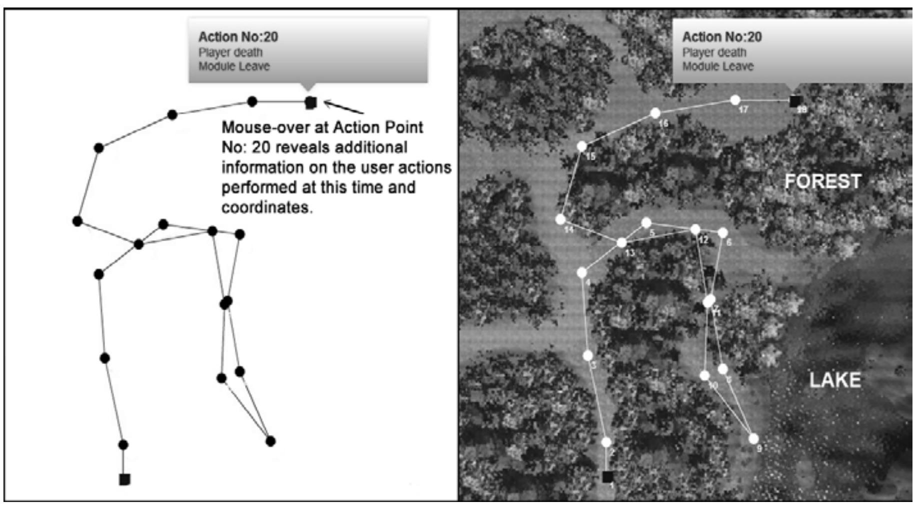

Figure 5: Data visualization of user actions. (Left: Early JAVA version without area map, Right: Later version with area map)
[Partial screenshot of *PeTRA*, used with permission.]

4.4 Current limitations

In an ideal situation, the telemetry for *Information Trails* should have been integrated into the game engine, with internal function calls available for remote data retrieval and transmission. However, since there was no such game engine available before *Information Trails*, we had to create the telemetry magic through much scripting. As a result, some of the *NWN2* game functions were too simplistic and limited for highly detailed behavior analysis. For example, Bioware’s developers used just two functions, namely *item_gained* and *item_lost*, to cover all events involving the adding or removal of items from a player's inventory. For obvious reasons, players could gain items in more than one way:

- obtained treasure chest
- bought from merchant
- stole from a non-player character (NPC)
- looted from a fallen enemy
- made by combining items (crafting) in the player's inventory
- created by a special spell
- given by an NPC, or another player in a persistent world

Since only one event call, *item_gained*, was available, it was impossible to truly tell how the item was 'gained' or obtained; a similar problem also existed for the event call, *item_lost*. We rectified the problem through the implementation of an 'Add Remark' function to allow us to easily annotate game events when needed.

As mentioned before, the economy for game development is very different from that of academic research. From the point of view of the game developer, all seven possible methods of gaining items (i.e., obtain, bought, stole, looted, crafted, created, given) were mere semantic differences that could easily be represented using one function: *item_gained*. Writing seven functions to represent each semantic possibility is viewed as inefficiency by programmers, regardless of the values they might hold for academic researchers. Game telemetry has the potential to change all that. Since items gained are often connected to the narrative of the story, understanding how players obtained certain items in the game may help improve the story and make a better game. As developers add more detailed user actions to their games, the data obtained by *Information Trails* will also become richer.

4.5 Future development

Even though *PeTRA* is already functional, our intention is to expand it into a full-fledge research system for in-process assessment of game-based learning, by standardizing the framework for users' action data and meta-data collection through game telemetry. That is, a complete learning design system from the development of Information Trails powered games, to *PeTRA*-powered online assessment reports.

The *Information Trails* assessment system requires the addition of several 'missing links' to make *in-process* assessment for game-based learning possible. The interdependent relationships among various components, which include: game engine, event listener, external database server, actionable learning and game objectives, and the in-process reporting tool, *PeTRA* is shown in Figure 6. (It should be obvious that without the assessment components, a standalone GBL engine will only produce more games that cannot be assessed.)

Given that *NWN2* is a 4-year old product, there is a need to expand *Information Trails* to other newer game engines – hopefully, one with integrated telemetry. As more game engines gain telemetry in the future, we hope to work with researchers around the world to standardize the list of user-generated action data, as well as the database structure to allow for open collaboration across other learning domains.

A standardized open database is also necessary for the development of new report assistants that will benefit user-groups from other industries. The separation of the report assistant from the game-based learning application is a necessity, because administrators and trainers who are not using the serious games can still gain access to the data visualization report, anytime, anywhere.

Future development of *PeTRA* will likely include a mobile or tablet version, which will provide the trainers and administrators access to the report while they are in the field. As cloud computing and ultra-portable mobile devices (such as iPAD2 and Android Tablets) gain popularity in the future, training organizations will begin looking into means to conduct 'distributed briefing' with these devices by directly obtaining data from the 'clouds.' As such, a *Report Assistant* for game-based learning will need to be cloud-friendly and be accessible through a browser from any of these mobile devices.

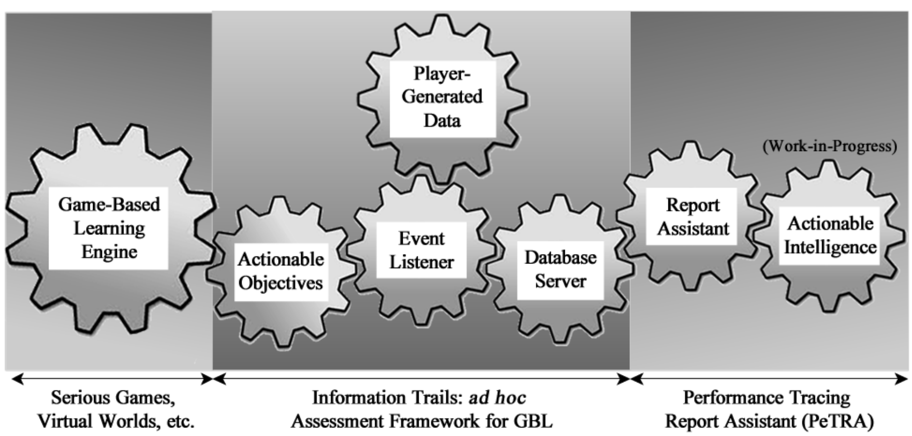

Figure 6: Relationships among various components of GBL with formative assessment capability

5. CONCLUSIONS

Assessment is a very important issue for game-based learning because without it there is no way to know if the learners have indeed achieved the proposed learning goals. Fortunately, the issue has begun to draw the attention of educator-researchers (as evidenced by this book), as well as game developers (in form of usability testing). Telemetry has been used in many areas in our lives although few have connected telemetry with game or assessment until now.

Even though game-based learning has the potential to revolutionize the way people learn, ineffective assessment methodologies will only muddy the waters and result in conflicting reports that will diminish the value of game-based learning. As it is, many technologies have been criticized as “useless,” “ineffective,” and showing “no significant difference” in improving education (c.f. Clark, 2007; Cuban, 2001). It is all the more important for researchers to focus their efforts in creating the right tools and finding the best assessment methodologies for the job.

Designing game-based learning is very different from designing entertainment games because the former requires the designer to take into consideration the many elements of learning assessment (such as learning objectives, instructional activities, etc.) and the latter has no need to do so. Linda G. Roberts, ex-Director of Education Technology for the U.S. Department of Education, once said, “I believed that researchers could improve the design and collection of data. Just as new technology created new opportunities for learning, it created ways to invent new tools for research and evaluation, particularly ways to track and monitor what, how, and when learning occurred” (2003, p. viii).

In the next few years, telemetry will gain importance as developers turn towards it to improve the usability and design of their games. As game engines with telemetry capabilities become available, assessment for game-based learning will become a reality. Data visualization will become the most challenging step in the assessment process as researchers and trainers struggle to make sense of the massive amount of data obtained from the online game-based learning environments. Instead of reinventing the assessment wheel at every turn, researchers should work together to solve common problems for the advancement of the field. This book (Ifenthaler, Eseryel & Ge, 2011) will become the cornerstone of that endeavor.

REFERENCES

Aldrich, C. (2009). *The complete guide to simulations & serious games* (p. 576). San Francisco, CA: Pfeiffer.

Baker, R. S. J. d, Corbett, A. T., Roll, I., & Koedinger, K. R. (2008). Developing a generalizable detector of when students game the system. *User Modeling and User-Adapted Interaction*, *18*(3), 287-314.

BBC News. (2007). *Computer game to boost key skills* Retrieved January 22, 2011 from http://news.bbc.co.uk/2/hi/uk_news/education/6254989.stm

Berger, A. (2006). “Neverwinter Nights” in the classroom. UMN News. Retrieved Sep 9, 2011 from http://www1.umn.edu/news/features/2006/UR_83484_REGION1.html

Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*(1), 1-73.

Black, P. J., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *1*(1), 5-31.

Botturi, L. & Loh, C. S. (2008). Once upon a game: Rediscovering the roots of games in education. In C. T. Miller (Ed.), *Games: Purpose and potential in education.* (pp. 1-22). New York, NY: Springer Science+Business Media, LLC.

Chen, S., & Michael, D. (2005). Proof of learning: Assessment in serious games. *Gamasutra*. Retrieved January 12, 2011 from http://www.gamasutra.com/features/20051019/chen_01.shtml

Clark, R. E. (2007, May-June). Learning from serious games? Arguments, evidence and research suggestions. *Educational Technology*, *47*, 56-59.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimum Experience*. New York, NY: Harper & Row.

Cuban, L. (2001). *Oversold and underused: Computers in the classroom.* Cambridge, MA: Harvard University Press.

Dragon Age: Origins (2009). [Computer Software]. Redwood City, CA: EA Games.

Duffy, J. (1997). *The role of virtual reality simulation in full scale joint military exercises* (p. 17). Newport, RI: Naval War College.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy* (2nd ed., p. 256). New York, NY: Palgrave Macmillan.

Gibson, D., Aldrich, C., & Prensky, M. (Eds.). (2006). *Games and simulations in online learning: Research and development frameworks* (p. 402). Hershey, PA: Information Science Publishing.

Gorniak, P., & Roy, D. (2005). Speaking with your Sidekick: Understanding situated speech in computer role playing games. In R. Michael Young, & John E. Laird (Eds.). Proceedings of the First Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE 2005). (Jun 1-5, 2005). Marina del Rey, CA: AAAI Press.

Harrington, H. L., Meisels, S. J., McMahon, P., Dichtelmiller, M. L., & Jablon, J. R. (1997). *Observing, documenting and assessing learning*. Ann Arbor, MI: Rebus, Inc.

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion* (p. 63). Orlando, FL: Naval Air Warfare Center.

Hattie, J. (1987). Identifying the salient facets of a model of student learning: A synthesis of meta analyses. *International Journal of Educational Research*, *11*(2), 187-212.

Hopson, J. (2006). We’re not listening: An open letter to academic game researchers. *Gamasutra*. Retrieved March 12, 2011 from http://www.gamasutra.com/view/feature/1783/were_not_listening_an_open_.php.

Ifenthaler, D., Eseryel, D., & Ge, X. (Eds.). (2011). *Assessment in game-based learning: Foundations, innovations, and perspectives*. New York, NY: Springer.

Joosten-ten Brinke, D., Gorissen, P., & Latour, I. (2005). Integrating assessment into e-learning courses. In R. Koper & C. Tattersall (Eds.), *Learning design: A handbook on modelling and delivering networked education and training*. (pp. 185-202). Berlin/ Heidelberg, Germany: Springer-Verlag.

Kaplan-Rakowski, R. and Loh, C. S. (2010). Modding and rezzing in games and virtual environments for education. In Y.K. Baek (Ed.), *Gaming for classroom-based learning: Digital role playing as a motivator of study*. (pp. 205-219). Hershey, PA: IGI-Global.

Kapp, K. M., & O'Driscoll, T. (2010). *Learning in 3D: Adding a new dimension to enterprise learning and collaboration* (p. 416). San Francisco, CA: Pfeiffer.

Lee, H. (2008). *Students' perceptions of peer and self assessment in a higher education online collaborative learning environment*. Unpublished Ph.D. dissertation. Retrieved February 8, 2011 from http://www.lib.utexas.edu/etd/d/2008/leeh55399/leeh55399.pdf.

Levin, A. (2010). Simulator training flaws tied to airline crashes. *USA Today*. Retrieved March 1, 2011 from http://travel.usatoday.com/flights/2010-08-31-1Acockpits31_ST_N.htm.

Loh, C. S. (2006). Tracking an avatar: Designing data collection into online games. Paper presented at the Annual Conference of the Association for Educational Communications and Technology (AECT 2006). Dallas, TX.

Loh, C. S. (2006). Designing online games assessment as "Information Trails". In D. Gibson, C. Aldrich, & M. Prensky (Eds.), *Games and simulation in online learning: Research and development frameworks*. (pp. 323-348). Hershey, PA: Idea Group, Inc.

Loh, C. S. (2009). Researching and developing serious games as interactive learning instructions. *International Journal of Gaming and Computer Mediated Simulations*, *1*(4), 1-19.

Loh, C. S., Anantachai, A., Byun, J. H., & Lenox, J. (2007). Assessing what players learned in serious games: *In situ* data collection, Information Trails, and quantitative analysis. In Q. Mehdi (Ed.), Proceedings of the Computer Games: AI, Animation, Mobile, Educational & Serious Games Conference (CGAMES 2007). Wolverhampton, UK: University of Wolverhampton.

Loh, C. S., & Byun, J. H. (2009). Modding *Neverwinter Nights* into serious game. In D. Gibson & Y. K. Baek (Eds.), *Digital Simulations for Improving Education: Learning Through Artificial Teaching Environments*. (pp. 408-426). Hershey, PA: IGI-Global.

Loh, C. S., & Li, I. H. (2010). Reducing re-training cost through on-demand, *ad hoc* assessment. Proceedings of the MODSIM World Conference and Expo 2010 - 21st Century Decision-Making: The Art of Modeling & Simulation. (Oct 10-13, 2010) Hampton, VA: MODSIM World Conference & Exposition.

Luthans, F. (2002). The need for and meaning of positive organizational behavior. *Journal of Organizational Behavior*, *23*(6), 695-706.

Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., & Fernández-Manjón, B. (2008). Educational game design for online education. *Computers in Human Behavior*, *24*, 2530-2540.

Nacke, L., Grimshaw, M., & Lindley, C. (2010). More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers, 22*(5), 336-343. Elsevier B.V. doi: 10.1016/j.intcom.2010.04.005.

Nacke, L., Lindley, C., & Stellmach, S. (2008). Log who's playing: Psychophysiological game analysis made easy through event logging. In P. Markopoulos *et al*. (Eds). *Fun and Games 2008, Lecture Notes of Computer Science*. *5294*, 150-157. Heidelberg, Germany: Springer-Verlag Berlin.

Nacke, L., & Lindley, C. (2010). Affective ludology, flow and immersion in a first-person shooter: Measurement of player experience. *Loading, 3*(5). Retrieved March 12, 2011 from http://journals.sfu.ca/loading/index.php/loading/article/view/72/71

Neverwinter Nights (2002). [Computer Software]. Lyon, France: Infogrames.

Neverwinter Nights 2 (2006).[Computer Software]. New York, NY: Atari, Inc.

O'Neil, H. F., & Robertson, M. M. (1992). Simulations: Occupationally oriented. In M. C. Alkin (Ed.), *Encyclopedia of Educational Research* (6th ed., pp. 1216-1222). New York, NY: Macmillan.

Prensky, M. (2001). *Digital game-based learning.* New York, NY: McGraw-Hill.

Prensky, M. (2006). *Don't bother me mom - I'm learning!* St. Paul, MN: Paragon House.

Pruett, C. (Sep 2010). Hot failure: Tuning gameplay with simple player metrics. *Game Developer Magazine. 19*(9). San Francisco, CA: Think Services. Retrieved March 7, 2011 from http://gamasutra.com/view/feature/6155/hot_failure_tuning_gameplay_with_.php

Reese, D. D. (2010). Introducing Flowometer: A CyGaMEs assessment suite tool. In R. van Eck (Ed.), *Gaming and cognition: Theories and practice from the learning sciences.* (pp. 227-254). Hershey, PA: IGI-Global.

Robert, L. G. (2003). Forewords. In G. D. Haertel & B. Means (Eds.), *Evaluating educational technology: Effective research designs for improving learning.* (pp. 290). New York, NY: Teachers College Press.

Robertson, J., & Good, J. (2005). Story creation in virtual game worlds. *Communications of the ACM*, *48*(1): 61-65.

Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *The Journal of Technology Learning and Assessment*, *8*(4), 3-41.

Saridaki, M., Mourlas, C., Gouscos, D., & Meimaris, M. (2007). *Digital games as a learning tool for children with cognitive disabilities: Literature review and some preliminary methodological and experiment-al results*. Paper presented at the European Conference on Games Based Learning, Scotland, UK.

Sasse, D. (2008). A framework for psychophysiological data acquisition in digital games. Unpublished Master thesis. Retrieved March 12, 2011 from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.140.5284&rep=rep1&type=pdf

Shute, V. J., Masduki, I., Donmez, O., Dennen, V. P., Kim, Y.-J., Jeong, A. C., *et al*. (2010). Modeling, assessing, and supporting key competencies within game environments. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge (Part 4).* (pp. 281-309). Boston, MA: Springer US. Retrieved March 5, 2011, from http://www.springerlink.com/content/t5261055733n18h2/

Shute, V. J., & Spector, J. M. (2008). SCORM 2.0 White paper: Stealth assessment in virtual worlds. Retrieved March 12, 2011, from http://www.adlnet.gov/technologies/evaluation/library/additional resources/letsi white papers/shute - stealth assessment in virtual worlds.pdf.

Skinner, B.F. (1938). The behavior of organisms: An experimental analysis. Cambridge, Massachusetts: B.F. Skinner Foundation

Stieger, I. (2008). Neverwinter Nights eXtender 4 (version 1.09) [software]. Starnberg, Germany: Stieger Hardware - und Softwareentwicklung. Retrieved from http://www.nwnx.org.

Telemetry. (n.d.). The American Heritage Stedman's Medical Dictionary. Retrieved March 9, 2011, from http://dictionary.reference.com/browse/telemetry

Tobias, S., & Fletcher, J. D. (2007). What research has to say about designing computer games for learning. *Educational Technology*, *47*(5), 20-29. Englewood Cliffs, NJ: Educational Technology Publications.

van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless… *Educause Review*, *41*(2), 16-30.

Weil, S. A., Hussain, T. S., Brunyé, T. T., Ferguson, W., Sidman, J. G., Spahr, L.S., & Roberts, B. (2005). Assessing the potential of massive multi-player games as tools for military training. In Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, (I/ITSEC 2005). Orlando, FL: NTSA.

Warren, R., & Sutton, J. (2008). Using a computer game for research on culture and team adaptability: Lessons learned from a NATO experiment. In the Proceedings of the NATO Research and Technology Organization (RTO) Human Factors and Medicine Panel HFM{142 Symposium on Adaptability in Coalition Teamwork. Copenhagen, Denmark (April 2008).

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., *et al*. (2008). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming*, *40*(2), 217-266.

Younis, B., & Loh, C. S. (2010). Integrating serious games in higher education programs. Proceedings of the Academic Colloquium 2010: Building Partnership in Teaching Excellence. Ramallah, Palestine: AMIDEAST.

Zoeller, G. (2010). Development telemetry in video games projects. Paper presented at the *Game Developer Conference 2010*. San Francisco, CA. Retrieved March 1, 2011 from http://gdc.gulbsoft.org/talk.

ACKNOWLEDGEMENTS

The research work in this article is made possible in part through funding from the Defence University Research Instrumentation Program (DURIP) grant from the U.S. Army Research Office (ARO). Screenshots of *PeTRA* (Figure 2 & 5) are used with permission from the author.

AUTHOR INFORMATION

Include the following information for each contributing author:

Complete name: Christian Sebastian Loh, Ph.D.

Institutional affiliation: Virtual Environment Lab (V-Lab)
Department of Curriculum & Instruction
Southern Illinois University Carbondale

Institutional address: 625 Wham Drive, Mailcode 4610
Carbondale, IL 62901, USA

Complete mailing address: Christian Sebastian Loh, Ph.D.
Dept. of Curriculum & Instruction
Southern Illinois University Carbondale
625 Wham Drive, Mailcode 4610
Carbondale, IL 62901, USA

Telephone number: (618) 4534206

Fax number (optional):

Email address: csloh@siu.edu

Website (optional): http://www.csloh.com, http://www.informationtrails.com

Short biographical sketch (200 words maximum):

Christian S. Loh (Ph.D.) is an Associate Professor of Learning Systems Design and Technology, and the Director of the Virtual Environment Lab (V-Lab) of the Department of Curriculum & Instruction, at the Southern Illinois University Carbondale.
Dr. Loh's research interests include assessment of learning within multi-user virtual environments and digital games with *Information Trails*© – an assessment framework he pioneered for game-based learning. He serves as associate editor for two international journals: the International Journal for Games and Computer Mediated Simulations (IJGCMS), and the International Journal for Game-Based Learning (IJGBL). He is a Past President of the Multimedia Production Division (MPD) of the Association for Educational Communications and Technology (AECT) and a regular judge (academic) for the annual Serious Games Showcase & Challenge

(SGS&C) competition hosted by the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC).