



Howard Everson &lt;howard.everson@gmail.com&gt;

## FW: ADL Research Newsletter March 2011

Everson, Howard &lt;heverson@air.org&gt;

Mon, Mar 7, 2011 at 9:21 AM

To: "howard.everson@gmail.com" &lt;howard.everson@gmail.com&gt;

From: ADL Research Newsletter[SMTP:ALAN.WORKMAN.CTR@ADLNET.GOV]

Sent: Monday, March 07, 2011 9:21:37 AM

To: Everson, Howard

Subject: ADL Research Newsletter March 2011

Auto forwarded by a Rule

# ADL Newsletter for Educators and Educational Researchers



Advanced Distributed Learning for Educators and Educational Researchers

March 2011

### EDITOR'S CORNER

There is a lot of excitement among educators and trainers about computer games and people are highly motivated to play such games. Therefore, it is understandable that anyone concerned about instruction is interested in using games for instructional purposes. These concerns led to the development of a forthcoming book, due out this Spring, about the subject ("Computer Games and Instruction," edited by S. Tobias and J.D. Fletcher, see Everson's article at left for the complete reference). One issue raised in the literature is the use of games for evaluation, and we are devoting this issue of the Newsletter to an article about that issue.

Those interested in computer games may wish to attend a debate about the value of games for the delivery of instruction at the annual meeting of the American Educational Research Association. The debate is between Valerie Shute, Florida State University, whose work on stealth assessment is summarized in the article at left. Shute feels that games can contribute to instruction, and Richard Clark, University of Southern California, who is more pessimistic about their value. The session is on Sunday, April 10, at

### Video Games to Assess Learning

Howard T. Everson  
Center for Advanced Study in Education  
City University of New York

"Games, in the 21st century, will be a primary platform for enabling the future." -Jane McGonigal, Institute for the Future

It has been estimated that more than 180 million people in the United States alone are "active gamers" i.e., they report playing video games more than thirteen hours per week (McGonigal, 2011). One of the more popular games, World of Warcraft, has over 10 million subscribers, some playing the game a staggering four hours a day, seven days a week (Bohannon, 2010). Those avid gamers, according to contemporary game designers and educational researchers (Tobias & Fletcher, 2011), are in fact developing their higher-order cognitive skills, like deductive and inductive reasoning and problem solving skills; and they are refining and furthering their conceptual understanding of gaming's intellectual challenges by writing—blogging and contributing wiki articles, nearly a quarter million at last count, to the WoWWiki website (Clarke-Midura & Dede, 2010; McGonigal, 2011). By all measure, they have become an established, ongoing learning community.

The growth and popularity of *serious* games, a term first introduced by Charles Abt (Abt, 1970) when promoting the use of simulations and games to improve training programs in business and learning in schools, was documented recently by Sara Corbett in an article appearing in the *New York Times Magazine* (Corbett, 2010). Corbett, in her thorough and penetrating look at how video games are used in schools, describes the effort underway at *Quest to Learn*, a non-charter middle school in New York City. School-wide initiatives, according to Corbett, are supported increasingly by

8:15- 10:15am, at the New Orleans Marriott, Room La Galerie 3.

As always, please send me any comments, questions, or suggestions and they will be considered for inclusion in future issues of the [Newsletter](#)

Sig Tobias

[sig.tobias.ctr@adlnet.gov](mailto:sig.tobias.ctr@adlnet.gov)

If you are interested in subscribing to this newsletter, please click on the button below.

 Join Email List

#### Prior Articles:

[About CORDRA](#) (Dec. 2006).

[ADL Introduction](#) (Jan. 2006).

[ADL Instructional Objects for Educational Use](#) (April 2007).

(April 2007).

[ADL Object Registry and Repository Infrastructure](#) (Feb. 2008).

[Consensus?](#) (Oct. 2010).

[Constructivist & Explicit Instruction Debate Followup](#) (April 2007).

[Constructivist & Explicit Instruction Debate Postscript](#) (Sep. 2007).

[Constructivist Pedagogy Is Superior – It Is a Matter of Definition](#) (Oct. 2010).

[Effectiveness of Web Based Training](#) (April 2006).

[E Learning and ADL in Korea](#) (April 2006).

[Games for Learning and Weak Vs Strong Instructional Guidance](#) (Sep. 2006).

[Games, Learning, and Society Conference](#) (Sep. 2006).

[KERIS Introduction](#) (April 2006).

[Kirschner et al. Discussed by Rosenshine](#) (Sep. 2006).

[Kirschner, Sweller, Clark Paper Discussion](#) (Sep. 2006).

[Knowledge Economy, SCORM, and Design-Based Research](#) (Sep. 2007).

[Learning Education Training Systems Interoperability \(LETSI\)](#)

generous funding from foundations, U.S. Education Department, and the National Science Foundation. The MacArthur Foundation, for example, has supported *Quest to Learn* as a demonstration site for innovative technology-based instruction and assessment.

These initiatives make it clear that computer technology is now beginning to change the ways in which we teach. Simulations and *serious* games, many believe, will also further our understanding of the way people learn, and the ways in which educators measure and calibrate that learning. As David Shaffer, a games for learning researcher at the University of Wisconsin tells it, educators need to look seriously at video games “because they create new social and cultural worlds—worlds that help people learn by integrating thinking, social interaction, and technology, all in service of doing things they care about” ([Gee & Shaffer](#), p. 105). Recognizing the power of video games, the science and mathematics education community, with substantial support recently from the National Science Foundation, is taking up the challenge of using video games to improve teaching and learning in key science, technology, engineering, and mathematics (STEM) domains. Sophisticated video games designed to teach numerical methods, statistics, and biology, to name only a few, are finding their way into science and math classes in many high schools and colleges ([Mayo, 2009](#)). Video game players, like learners in other more traditional academic learning environments such as classrooms and laboratories, are using their knowledge, conceptual understanding, and reasoning abilities to solve problems encountered as they move through increasingly complex and challenging levels of a serious game, ones designed explicitly to promote learning ([Clarke-Midura & Dede, 2010](#)).

The growing interest in computer-based video games as a tool for learning suggests that investments in this technology are worthwhile and hold promise. Yet it remains unclear that we can use these technologies effectively to advance our understanding of what and how much is learned. Does their power to create rich scenarios and present complex tasks and challenges offer the potential to transform educational assessment? Can games be designed to change the way we test students? Many learning scientists and assessment experts are suggesting the answer is a resounding yes ([Gee & Shaffer, 2010](#)).

Youngsters (and adults) playing popular video games such as *Halo* or *Tetris*, according James Gee, a scholar who studies video games and learning at Arizona State University, are not only developing their learning skills, they are also working on what may be the next generation of educational assessments. For Gee, “a video game is nothing but a series of tests” ([Rothman, 2010](#), p.1). Gee's view of video games is becoming more wide-spread among psychometric experts, too. Advances in computer technology, assessment design, and statistics hold promise for adapting video games to function as educational assessments.

#### Games as Educational Assessments

In their seminal book *Knowing What Students Know : The Science and Design of Educational Assessment*, Pellegrino, Chudowsky, and Glaser ([2001](#)) suggest that educational assessments need to be designed not only to tell us what students know and can do, but also ought to be more closely integrated with instruction. For learning scientists like Pellegrino and his colleagues at the University of Illinois-Chicago's *Learning Sciences Research Institute*, educational assessment design can and should be “part of the quest for improved instruction” ([Quellmalz & Pellegrino, 2009](#), p. 81).

(Feb. 2009).

[Learning Education Training Systems Interoperability \(LETSI\)](#) (Feb. 2009).

[Minimally Guided Instruction Effectiveness](#) (Sep. 2006).

[More Research Needed](#) (Oct. 2010).

[Multi Media Lab in Taiwanesea](#) (April 2006).

[Newsletter Purpose](#) (Jan. 2006).

[Paradigms, Experts, Novices, and Lessons](#) (Oct. 2010).

[Report on the "Games for Change" Meeting](#) (Sep. 2009).

[Report on the Joint ADL Co-Lab Implementation Fest 2008](#) (Oct. 2008).

[Responses to SCORM, LETSI, and Learning from Instructions](#) (Oct. 2008).

[SCORM, LETSI, and Learning from Instruction](#) (Oct. 2008).

[Search and Discovery of Instructional Objects](#) (Feb. 2008).

[Search and Discovery of Instructional Objects](#) (Feb. 2008).

[Semantic Net](#) (June 2009).

[Sweller & Tobias Agree](#) (Oct. 2010).

[Tamkang University's MINE Lab Introduction](#) (April 2006).

[Training for Adaptable Performance: A Workshop Report](#) (Sep. 2009).

[Training Evaluation Information on the ADL Website](#) (Feb. 2008).

[Web 2.0 and ADL](#) (June 2008).

Newsletter archives, as well as the current issue, are available in the [archive](#).

More recently, James Gee and his colleague David Shaffer extended this view by arguing persuasively that what stands in the way of using computer technology and video games for improving teaching and learning are standardized tests, and educators' over reliance on paper-and-pencil forms of testing. Gee and Shaffer suggest that current conceptions and uses of standardized tests maintain and foster an unfortunate separation between learning and assessment. Like many in education today, Gee and Shaffer see contemporary forms of standardized testing as the "tail that wags the dog of learning" ([Gee & Shaffer, 2010](#), p.6). Gee and his colleagues challenge us, therefore, to move from designing games for learning to the design of games for testing—a call for a more seamless integration of teaching, learning and assessment.

As noted earlier, there is a small but growing community of cognitive scientists, learning theorists, and educational measurement specialists (i.e., psychometricians) working to integrate assessments into game designs. Their challenge, however, is formidable largely because contemporary psychometrics has been developed to deal with tests comprised of multiple-choice and short constructed response tasks—test items that can be administered relatively easily and cheaply in paper-and-pencil format and assessed using rather simple binary (correct/incorrect) scoring methods.

Methods and models of test score analysis have evolved in response to the popular and predominant educational testing protocol. During the last half of the 20th century, for example, psychometrics progressed from classical test theory, which relied on statistical methods based largely on classical or "frequentist" statistics to estimate the correlations between and among an examinee's item responses, to latent variable models that incorporate Bayesian statistics and item-response theoretic methods to model the underlying cognitive constructs (e.g., verbal or quantitative reasoning, or problem solving) that drive performance on educational tests ([Embretson & Reise, 2000](#)). The complex tasks and simulations we see in most video games for learning, no doubt, will push contemporary psychometrics to develop statistical techniques and measurement models that work for learning environments that are rich and textured, and that produce multidimensional and dynamic data. Video games for learning and testing will require new and very different psychometric models, particularly if the goal is to integrate robust, valid assessments into game designs and, perhaps more important, if educators and teachers are to appreciate fully the value of games for learning.

Moreover, the potential to track students' keystrokes, mouse clicks and other actions and communications, and then map them onto frameworks of higher-order cognitive skills and abilities is what makes games for assessment appealing to some in the educational measurement field. But digging for insights into how people learn by mining terabytes of data—a few thousand gigabytes—will require that psychometricians develop and extend data mining tools that take advantage of advances in computational methods and statistics to make sense of the very large amounts of assessment related data generated by video games and simulations ([Ye, 2003](#)). By design, simulations and video game scenarios are rich and multidimensional—many different combinations of knowledge, skills, and abilities are tapped in these simulations and scenario-based games. Probability-based models, like Bayesian belief networks, artificial neural networks, or newly emerging social network analysis may be more suited to assessing learning in scenario-based, dynamic games for learning ([Mislevy &](#)

**Academic ADL Co-Lab**  
 222 W. Washington Ave.  
 Suite 470  
 Madison, WI 53703  
 Phone: 608.263.6657  
 Fax: 608.265.0070  
[www.academiccolab.org](http://www.academiccolab.org)

**DISCLAIMER:**

This newsletter is provided as a service to those in the academic community and does not necessarily represent the opinions of or an endorsement by the Advanced Distributed Learning (ADL) Initiative.

[Levy, 2007](#); [Quellmalz, Timms, & Schneider, 2009](#)). To highlight these issues and explore the possibilities of using video games as educational tests, samples of a few of the more promising assessment approaches are presented next.

**WestEd's SimScientists**

The *SimScientists* project ([www.simsScientists.org/home/index.php](http://www.simsScientists.org/home/index.php)) is part of WestEd's research and development portfolio exploring the role that simulations can play to enrich science learning and assessment. The simulation-based assessments are designed to present dynamic, engaging interactive tasks testing complex science knowledge and inquiry skills that move beyond what is tested with paper-and-pencil measures. The WestEd team's approach uses evidence-centered design mapped to multi-user environments to "create and link *student models* (content and inquiry targets) to *task models* that will elicit evidence of the targeted knowledge and inquiry, to *evidence models* of the scoring and reporting that will be used to describe student progress and achievement on the targets" ([Quellmalz, Davenport, Timms, & Buckley, 2009](#), p. 2).

In addition to using evidence-centered assessment design, the team at WestEd is exploring the use of Bayesian belief networks, model tracing methods developed originally for use with cognitive tutors at Carnegie Mellon University ([Anderson, Corbett, Koedinger, & Pelletier, 1995](#)), and rule-based decision trees to focus on students' misconceptions of the scientific concepts and theories embedded in *SimScientists*' assessment system. The interested reader can find more complete descriptions of the *SimScientists* project, along with a number of papers describing the technical and psychometric details, at [www.simsScientists.org/publications/index.php](http://www.simsScientists.org/publications/index.php).

**Epistemic Network Analysis**

David Shaffer and his colleagues at the University of Wisconsin-Madison, along with collaborators from the University of Maryland, are attempting to use both evidence-centered design and epistemic network analysis (ENA) to design assessments for digital learning environments ([Shaffer, Hatfield, Svarovsky, Nash, Nulty, Bagley, Frank, Rupp, & Mislevy, 2009](#)). These researchers view games for learning as being *epistemic games*—a situated game world guided by professionals mentoring novices. Distinct from traditional direct instruction, the framework for epistemic games uses more authentic professional training methods derived and developed as part of a community of practice with its own cultural grammar. This culture of grammar, accordingly, is composed of skills, knowledge, identities, values, and epistemology (Shaffer and his colleagues refer to the framework as SKIVE). This perspective provides researchers with the connective tissue that forms the network of relationships—conceptual, practical, moral, personal, and epistemological. An example of an epistemic game, *Urban Science*, can be found at <http://epistemicgames.org/eg/category/games/urban-planning/>.

Epistemic network analysis, like other forms of social network analysis, permits the development of network graphs that capture data about gamers' learning at different time points in the game. The elements of the learner's epistemic frame are represented as nodes in the graph. Dynamic network graphs then provide ways to assess the connections between and among novices and experts in the community of practice instantiated in the game. Epistemic network analytic methods are then used to assess the



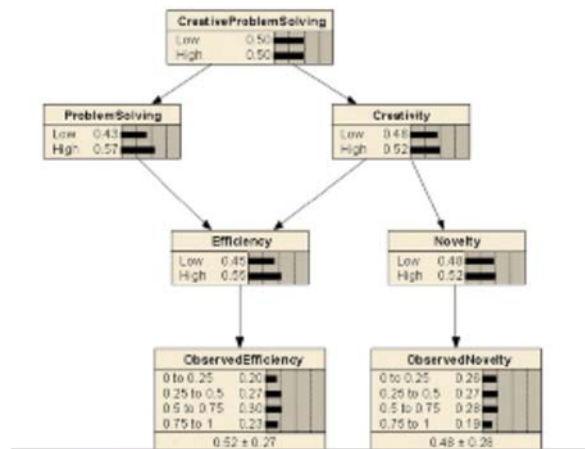
development of the epistemic frames and, in theory, model the development of complex thinking and reasoning skills as a player moves through the game. The mathematics (or psychometrics) of ENA are complex and much of the work is still in the early stages of development. A more comprehensive overview of epistemic games and epistemic network analysis can be found at <http://epistemicgames.org/eg/category/projects/dynamic-stem-assessment-through-epistemic-network-analysis/>. For those interested in learning more about evidence-centered design, an accessible set of papers can be found at [www.education.umd.edu/EDMS/mislevy/papers/ECD overview.html](http://www.education.umd.edu/EDMS/mislevy/papers/ECD%20overview.html).

### ***Stealth Assessments***

Valerie Shute and her colleagues ([Kaya, 2010](#)) have developed the idea of “stealth assessments”, i.e., assessments that are seamlessly integrated into the very fiber, the software code, of the video game’s instructional framework in an effort to support learning. Like the work on epistemic network analysis described earlier, Shute’s approach also includes evidence-centered assessment design, but in her *stealth* framework the assessments play a more formative role, thereby allowing for a closer link to instruction and the delivery of content more closely tailored to the learner’s needs.

The main assumptions underlying stealth assessment research are that: (a) learning by doing (required in game play) improves learning processes and outcomes; (b) different types of learning and learner attributes may be verified and measured during game play, (c) strengths and weaknesses of the learner may be capitalized on and bolstered, respectively to improve learning, and (d) formative feedback can be used to further support student learning ([Shute, 2011](#)).

The measurement approach underlying stealth assessments relies on instantiating the learner’s competency model, i.e., the collection of knowledge and skills to be assessed, using Bayesian belief networks. This approach uses probabilistic graphical models that represent a set of random variables and their conditional dependencies via a directed acyclic graphs. The approach relies on efficient algorithms that allow the system to make inferences and “learn” from the data generated by playing the educational game. More generally, Bayesian belief networks can represent and solve decision problems under uncertainty. These networks can capture a player’s actions in real-time, and use the associated influence diagrams to provide evidence of whether the learner is acquiring the desired competencies. Again, the evidence-centered design framework provides the assessment system’s evidence model. Figure 1, below, adapted from Shute’s current work represents a fragment of the competency model used inside a game designed to assess creative problem solving called *Oblivion*.



**Figure 1.** A Bayesian network model instantiating the competencies estimated via an assessment embedded within a game designed to measure creative problem solving.

According to Shute, the upper five nodes of the network provide evidence that connects the learner's action to her competencies using estimates based on Bayesian probability distributions. Each node has two or more discrete states, and the marginal probabilities are estimated for each state. These Bayesian belief networks can be used within the context of the game to "infer" varying levels of the learners' competencies as they navigate through the game. In turn, they hold promise for informing teaching and learning as the student, the game player, works through the sequential scenarios of the game. In this sense, with embedded "stealth" assessments, the learning game can be designed to adapt to the learners' response patterns, making inferences about what the learner knows at particular points in the game's sequence.

This Bayesian technology is at work in many commercial applications, like when Amazon.com makes suggestions about what books you might like based on the patterns it "learns" from your earlier purchases. Like the other examples of this new generation of educational measurement and psychometrics, Shute and her colleagues are at the very early stages of development. For more detailed descriptions of this line of work you can visit Valerie Shute's website: [myweb.fsu.edu/vshute/pdf/shute\\_pres\\_h.pdf](http://myweb.fsu.edu/vshute/pdf/shute_pres_h.pdf). You can also read her chapter in the [Tobias and Fletcher](#) book (2011).

### Conclusion

As Mayo suggests in her piece on video games and STEM learning ([Mayo, 2009](#)), the ability to bring computer-based video games to scale and sustain their use in education will likely rest on the development of a profitable business model. By viewing games as assessments, we may find a workable business model that will make investments in video games for learning and assessment economically viable. Schools and school districts, in the future, may be able to convert some of their current expenditures on textbooks and traditional standardized tests into more innovative and productive video game designs—learning and assessment enabled video games.

This new generation of computer-based games, games for testing as Gee and Shaffer refer to the them, offer the potential for transforming the how, what, when, where, and why of educational testing. As Quellmalz and Pellegrino noted in their perspective on technology and testing, these 21st

century technologies facilitate “the design of complex, interactive tasks that extend the range of knowledge, skills, and cognitive processes that can be assessed” ([Quellmalz & Pellegino, 2009](#), p. 81).

Obviously, it is difficult (if not downright foolish) to predict which of these psychometric approaches, or others yet to be developed, will win out. Nevertheless, it is likely, given the current research and development efforts underway, that they will lead to more coherent, integrated video game designs that will align learning, assessment and instruction in more real-world, scenario-based games for learning and testing.

## References

Abt, C.C. (1970). *Serious Games*. NY: Viking Press.

Anderson, J.R., Corbett, A.T., Koedinger, K.R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences*, 4(2), 167-207.

Bohannon, J. (2010). Game-miners grapple with massive data, *Science*, Vol. 330, pp. 30-31.

Clarke-Midura, J. & Dede, C. (2010). Assessment, technology, and change. *Journal of Research on Technology in Education*, 42 (3), pp. 309-328.

Corbett, S. (2010). Game theory. *New York Times Magazine*, 19 September, pp. 54-70.

Embretson, S.E. & Reise, S. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Erlbaum Associates.

Gee, J.P. & Shaffer, D.W. (2010). *Looking Where The Light Is Bad: Video Games and the Future of Assessment*. Epistemic Game Group Working Paper No. 2010-02. Madison, WI: University of Wisconsin-Madison.

Kaya, T. (2010). A ‘stealth assessment’ turns to video games to measure thinking skills. *The Chronicle of Higher Education*, 13 February, 2010.

Mayo, M.J. (2009). Video games: A route to large-scale STEM education? *Science*, 323, 79-82.

McGonigal, J. (2011). *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. New York, NY: Penguin Press.

Mislevy, R.J. & R. Levy, R. (2007). Bayesian Psychometric Modeling From An Evidence-Centered Design Perspective. In C.R. Rao & S. Sinharay (Eds.) *Handbook of Statistics 26: Psychometric*, pp. 839-866. NY: Elsevier Press.

Pellegrino, J., Chudowsky, N. & Glaser, R. (Eds). (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academic Press.

Quellmalz, E.S., Davenport, J., Timms, M., & Buckley, B. (2009). *Quality Science Simulations for Formative and Summative Assessment* (available at <http://www.simsScientists.org/publications/index.php>).

Quellmalz, E.S. & Pellegrino, J.W. (2009). Technology and testing. *Science*, 323, p. 75-79.

Quellmalz, E.S., Timms, M., & Buckley, B. (2010). *Assessment of Student Learning in Science Simulations and Games* (available at [http://www7.nationalacademies.org/bose/Gaming\\_Sims\\_Commissioned\\_Papers.html](http://www7.nationalacademies.org/bose/Gaming_Sims_Commissioned_Papers.html)).

Rothman, R. (2010). Video games take testing to the next level. *Harvard Education Letter*, 26(6) pp. 1-3.

Shaffer, D.W., Hatfield, D., G.A. Svarovsky, G.A., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A.A., & Mislevy, R.J. (2009). Epistemic Network Analysis: A Prototype for 21st Century Assessment of Learning. *International Journal of Learning and Media*, 1(2), pp. 1-21.

Shute, V.J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J.D. Fletcher (Eds). *Computer Games and Instruction*. Charlotte, NC: Information Age Publishers.

Tobias, S. & Fletcher, J.D. (Eds.) (2011). *Computer Games and Instruction*. Charlotte, NC: Information Age Publishers.

Ye, N. (Ed.) (2003). *The Handbook of Data Mining*. Mahwah, NJ: Erlbaum Associates.

---

[Forward this email to a friend](#)

[Interesting? Share this email on social networks](#)

---

[Change email address / Leave mailing list](#)  
[Powered by YMLP](#)

---