What Do We Know About Assessment in Games?

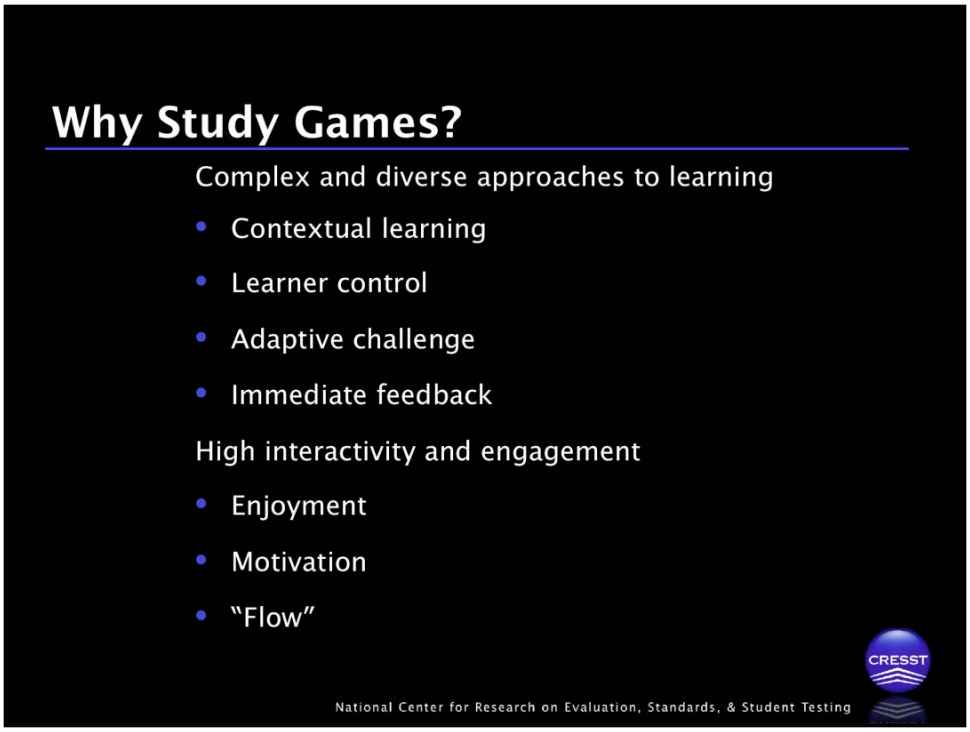
Eva L. Baker and Girlie C. Delacruz

American Educational Research Association
Annual Meeting

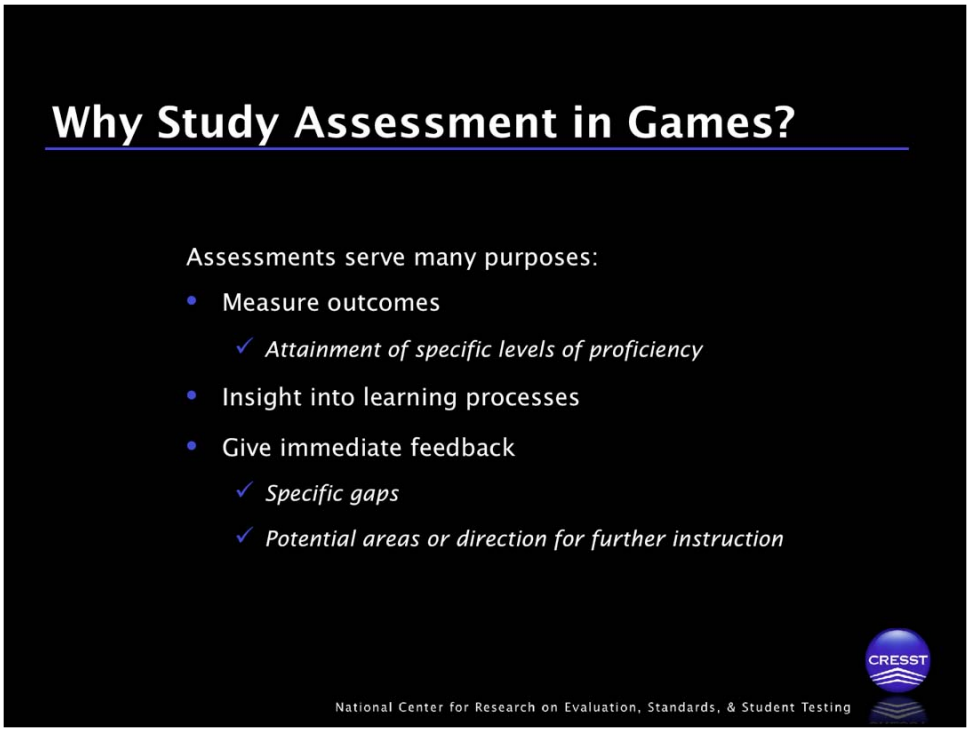
New York , NY - March 23, 2008

National Center for Research
on Evaluation, Standards, & Student Testing

UCLA | Graduate School of Education & Information Studies | cresst.org



As I am sure you will hear in other talks, there are many reasons to look at the use of games for learning purposes. Some of these include games' potential to create complex and diverse approaches to learning. Many games can contextualize the learning experience. They can give the learner control over how to explore their learning environments. Games can respond to the learner in terms of feedback and creating challenges based on what they do. It is clear that games are definitely engaging and motivating. I'm sure most of you have experienced that "flow" while playing a game where nothing else matters except accomplishing those challenges like beating your score or gaining more powers. Given these potential benefits, it just makes sense to look at games to leverage their positive aspects to improve learning environments,



Now why should we care about the role of assessment in game-based learning? Like all other learning environments, good assessment practices are critical for effective learning, What I want to talk to you about today is what Eva and I advocate are important considerations to make when we think about assessment and games. Assessments serve many purposes, not only to measure outcomes of learning as we have traditionally viewed them. But when we watch people as they engage in the game, this is also an opportunity to gain insight into the learning process. Assessments help us provide feedback to both the student and teacher, pointing to specific areas of difficulties people are having which may provide direction for further instruction..

How Do We Assess in Games?

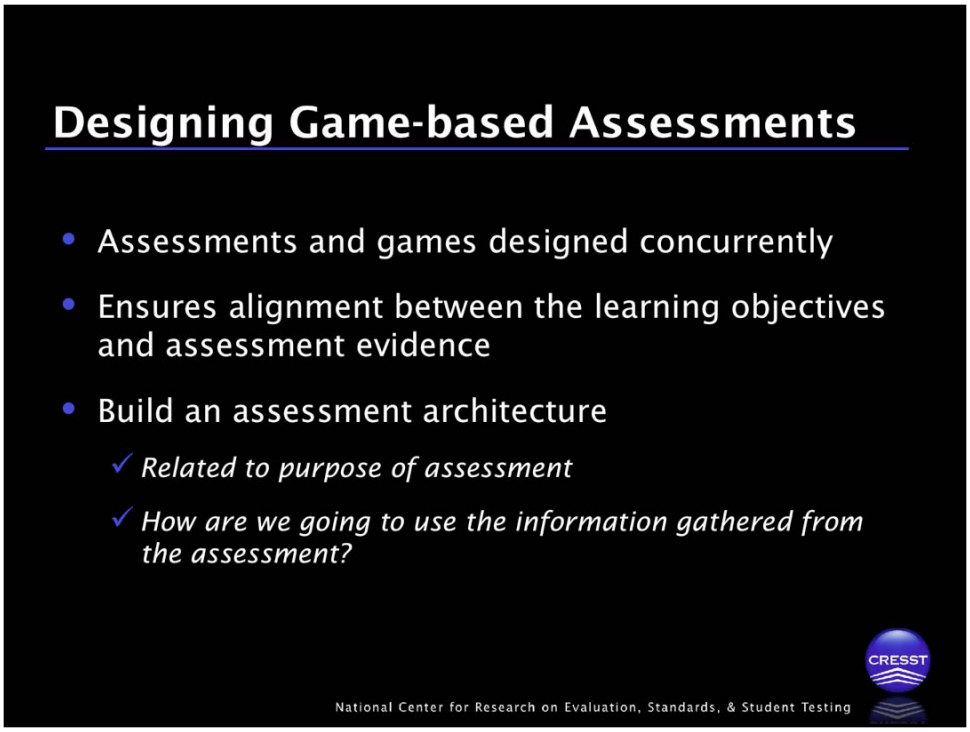
- Typical scoring mechanisms
  - ✓ *E.g., targets acquired or obstacles overcome*
  - ✓ *Related to learning objectives or are they motivational?*
- Wrap-around assessments
  - ✓ *Knowledge maps, external measures, interviews*
- Embedded assessments
  - ✓ *Ask relevant questions in situ*
  - ✓ *Analyze clickstreams*
  - ✓ *Facilitate modulating environments*

http://xbox360media.ign.com/xbox360/image/article/795/795776/americas-army-true-soldiers-20070611044817446.jpg

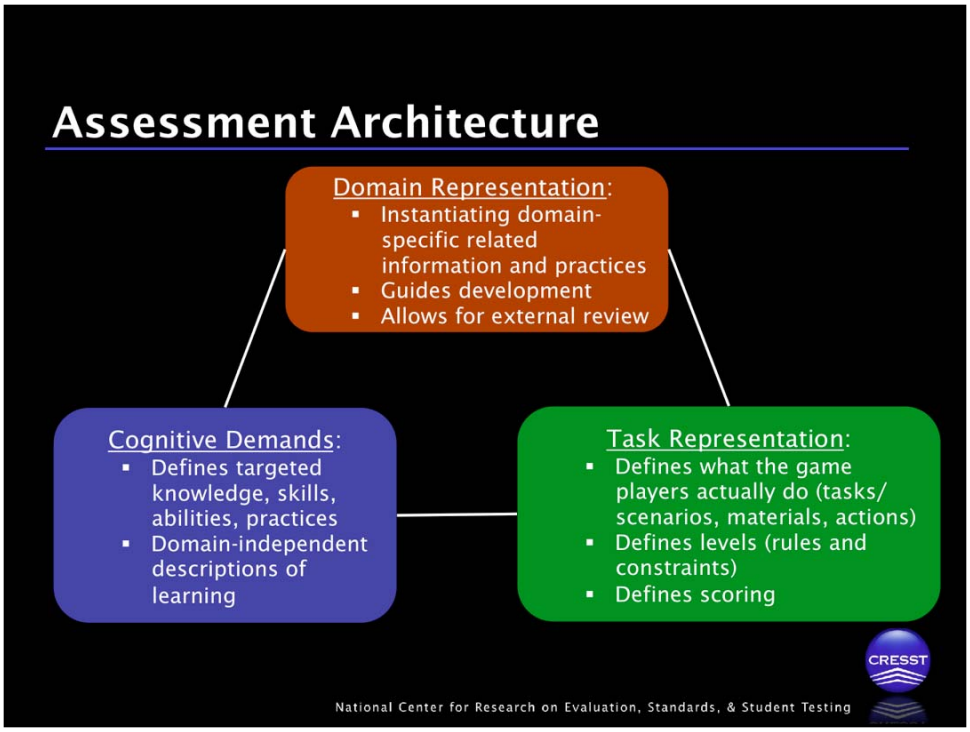
National Center for Research on Evaluation, Standards, & Student Testing

CRESST

So how do we assess in games? It is not a stretch to view typical scoring mechanisms built into games such as number of targets acquired or obstacles overcome as a form of assessment, This is not unlike when we tally up the number of wrong or right answers on a test. However, it is not always clear that what is counted relates to the learning objectives of the game. Often times, these metrics are developed for motivational purposes such as to provide just the right amount of challenge. So what else do we do? People studying games have used a variety of wrap-around assessments. For instance, Harry and his colleagues have used knowledge maps to measure conceptual understanding, Rich and Roxanna have used external measures of retention and conceptual understanding. Others such as David Schaffer and Kurt Squire interview game players on the related content. We a;so argue that the promise of game-based assessment is its potential for formative assessment. You can also embed assessments by incorporating into the game play the asking of relevant questions. Another way would be that the assessment would be experienced as part of the game through evaluation of students' online clickstream behavior which we at CRESST have used for many of our online simulations. This is what I meant earlier when I said that by watching game players' behaviors and actions, we can gain insight into their learning processes. Embedded assessments allow us to react to the learner or learners, modulating the game environment in response their activity. However, while embedded assessments are ideal, complex off the shelf games rarely make their source code available for seamless assessment capabilties for well known reasons.



So how should we approach design game-based assessments? First, we propose that when designing the game, the assessments should be designed concurrently rather than as an after-thought. By designing the assessments up front while we establish our learning objectives and the game design, we can ensure alignment of our measures as well as capture the right actions. In order to do this,you need to build into the game what we are calling an assessment architecture. What goes into the assessment architecture is related to the purpose of the assessment. That is, what we consider depends on how we are going to use the information that we will gather from the assessment.



The assessment architecture is made up of three parts: First, you must specify the cognitive demands you want to measure, which describes the knowledge, skills, abiltiites, and behaviors that are the target of the learning environment. Eva's model-based assessment approach calla for domain-independent descriptions of the relevant aspects to learning such as content understanding, problem solving, metacognition, communication, and teamwork and collaboration. We then instantiate these aspects of learning with the domain-specific information and practices that are related. This allows us to build in a domain representation which is an explicit representation of what is to be learned and assessed, as well as the behaviors or performances that reveal these constructs.. We at CRESST have been using software based ontologies as our domain representation to capture the necessary information and relationships and automate the assessment process. The domain representation serves many functions. It essentially defines the universe of what is to be learned and assessed, guiding assessment development by specifying what can be sampled from when we design our scenarios and tasks. Because it is explicit, it also enables external evaluation for validity purposes which I will get into later. Finally, the assessment architecture needs the task representation which defines what it is the game-player will actually do. The task representation portion of the assessment architecture is comprised of the materials or environment with which the learners are expected to interact, the game narrative or scenarios, the actions taken by the game players, and the scoring method. it is important to note that these elements of the assessment architecture are not to be considered independent of each other since elements of each informs the design of the others. As you can see, the assessment architecture of the game represents the assessment design with different levels of specificity, making modulations possible without having to start from scratch. Also, we want to stress that the assessment architecture is a socially constructed process, such that it must

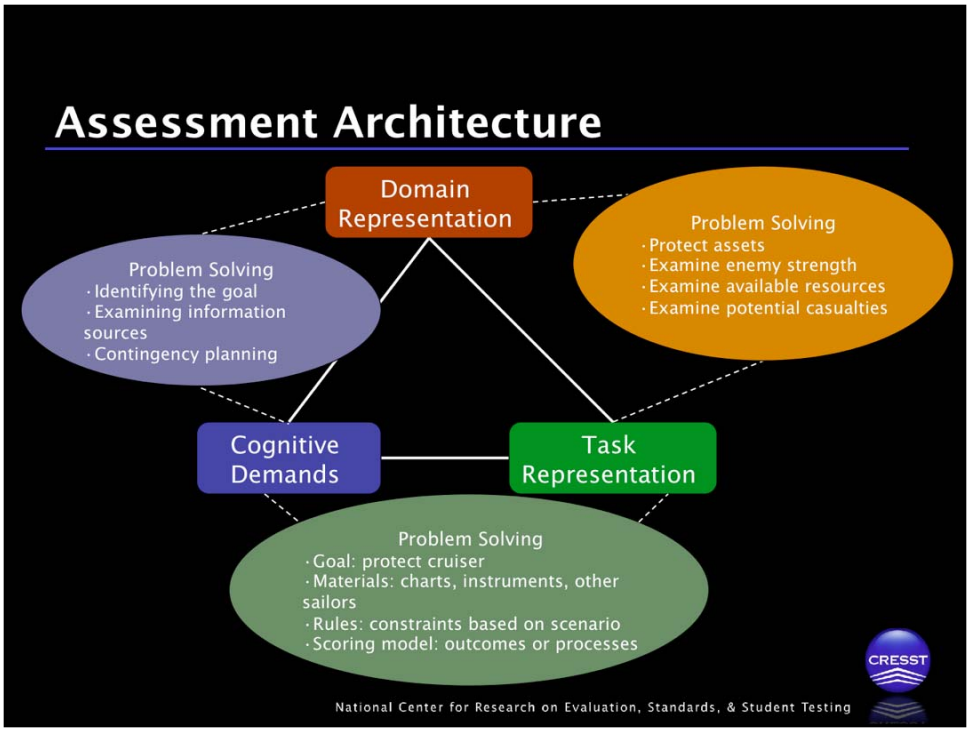
Assessment Architecture

http://blog.800hightech.com/wp-content/uploads/usn-naval-fleet-cruise.jpg

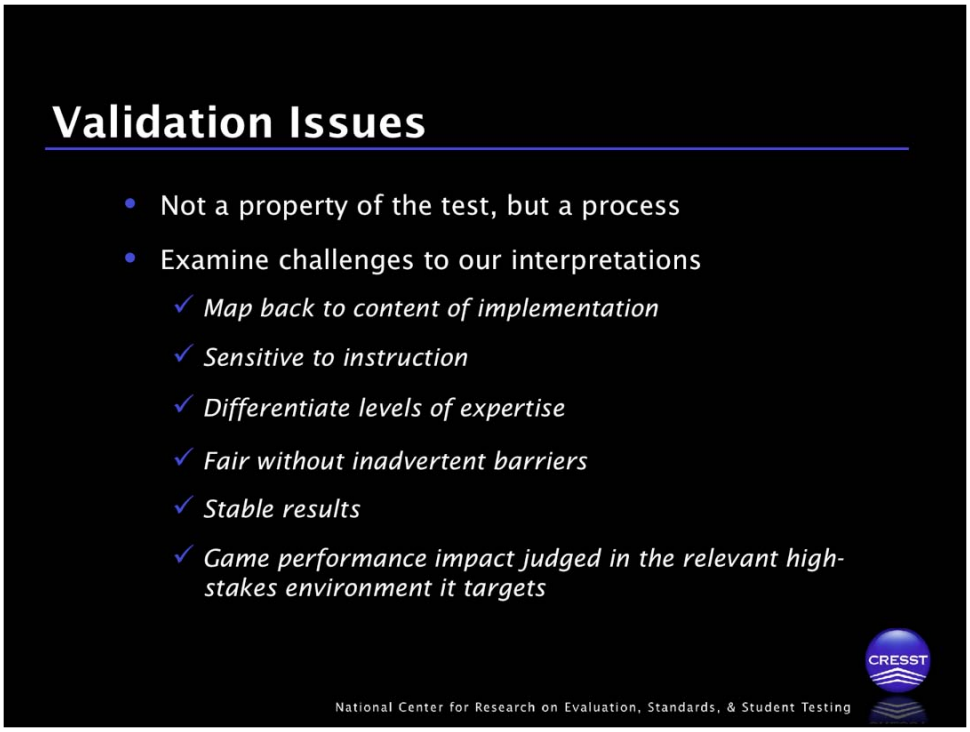
National Center for Research on Evaluation, Standards, & Student Testing

CRESST

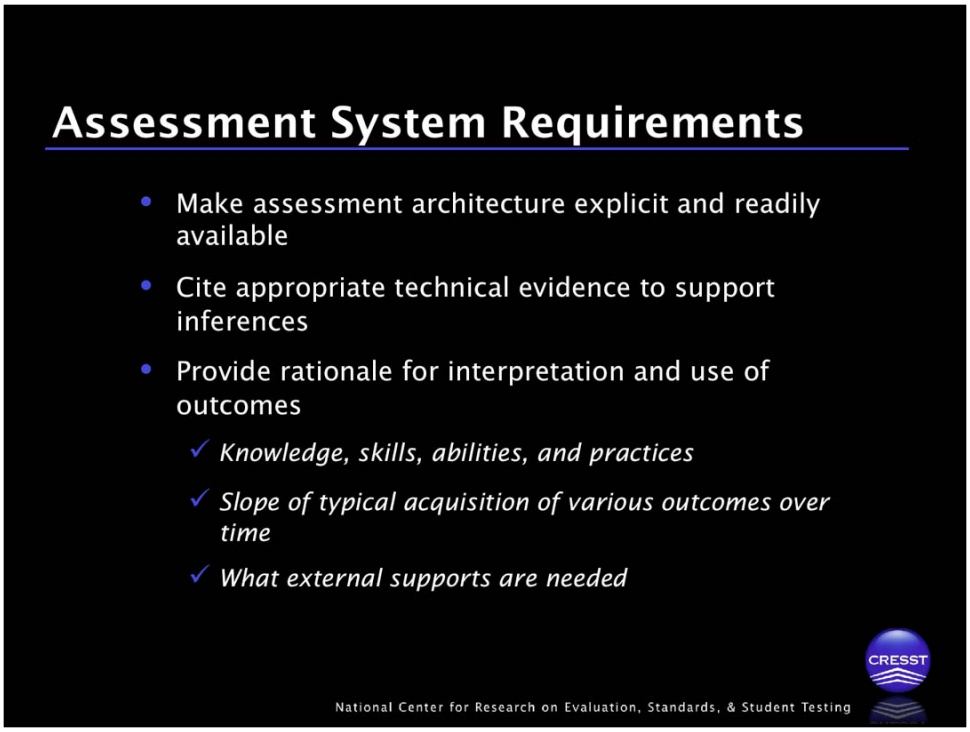
I know that was a lot to take in, so I want to go over what building an assessment architecture would theoretically look like. Let's consider the design of a game targeting the training of naval officers in setting up a basic air-defense plan.



Let's say we were interested in a game targeting the naval officer's problem solving skills. This may require identifying the goal of the various scenarios, examine the various information sources and be ready with contingency plans when they face the unexpected. Our domain representation would consist of things such specifying the various goals of the scenarios such as protecting assets like the cruiser or aircraft carrier. We would also identify the possible information sources that would be made available in the game such as the enemy's strength, the fleet's available resoures, the geography of the mission, and the potential casualties. From this, we would create the various tasks or scenarios, including specifying he materials needed such as the charts, instruments, or other participants on the ship. The task representation would allow us to define the various levels of the game, setting up the rules that correspond with the constraints of the scenario. For instance, during one task, you might need to to protect the cruiser while also keeping free your airplanes. However, as the game goes on, the scenarios could be fairly complex by incorporating realistic obstacles such as instrument failure. Finally, depending on the purpose of the assessment (was it to certify proficiency or look for places in the process that needs work), the task representation would capture the scoring model. For instance, if the purpose of the game was to certify proficiency, then one might look at simply the outcomes of the various scenarios. However, if one wanted to identify potential areas of remediation, one might look at the actions taken by the game player such as how one interacted with the charts or instruments, or the information accessed or possibly more importantly, did not access.



Even when we design them carefully, like all other forms of assessment, game-based assessments need to face the task of validating the interpretation of the outcomes. Assessment validity is not a property of the test, but rather a process that is related to how the information garnered from the assessment is to be used and interpreted. It requires consideration of a host of issues to to determine potential challenges to our interpretations of the outcomes based on the purposes our assessments aim to serve. This involves examination of its mapping to the content, are our assessments sensitive to instruction or can they differentiate levels of expertise? We must look at whether the game has been designed fairly without inadvertent barriers such as over-reliance on unnecessary language and and whether proper support has been given to different groups of learners? Are the results of the assessments stable? We must also consider the consequences of our interpretations by looking at game performance impact in the high-stakes environment it targets such as actual job performance.



For validity purposes, our assessment architecture must be made explicit and readily available like the system requirements of the game. The assessment system requirements would be able to cite the appropriate technical evidence that was gathered to provide the rationale for interpretation and use of the outcomes. This would include what we discussed in previous slides, such as the cognitive demands of the game, the knowledge, skills, abilities,and practices that were the target of the game as well as the slope of typical acquisition of various outcomes over time, and what external supports, if any, are needed to support the learning in the game.

In our endeavors to improve game-based learning through good assessment practices, we are working on these key elements, some of which I've touched upon today. We are improving our approaches to specifying the assessment architecture and capturing it explicit representations so that they are useful as assessment system requrements. We are working toward examining how best to capture process data for formative assessment purposes. For the cases where we do not have access to the game design proecss right from the beginning, we are working toward creating better criterion-based wrap-around assessments. We are also working on building in self-assessment measures especially because many of these games are unsupervised to promote self-regulated learning. Finally, we are particularly interested in transfer and generalization, that is where the oitucomes of the game are asessed under different conditions and different constraints. Together, we believe these elements are key to ensuring that assessments are utilized appropriately and effectively in games.


CRESST
National Center for Research
on Evaluation, Standards, & Student Testing
UCLA | Graduate School of Education & Information Studies
VISIT US ON THE WEB
cresst.org
baker@cse.ucla.edu and gdelacruz@cse.ucla.edu