Game-based assessment (GBA) refers to the use of games, both video games and other types of games, to assess learners’ various competencies—skills, knowledge, and dispositions. Practical implementations of GBA can vary from context to context. Different GBA models have varying levels of assessment capacity, and the foci depend on how much the role of assessment is emphasized in the process of designing the game. For example, if both assessment and learning designers are involved in the early design stages, it is more likely that assessment and game mechanics will be seamlessly integrated (i.e., stealth assessment). However, even if a game is developed without explicitly supporting assessment, educators can still use it to create GBA activities.

Well-designed GBA can (a) provide engaging and authentic contexts, (b) elicit evidence for the competency of interest, and (c) motivate learners to continuously adjust their actions, which can lead to learning. The underlying assumption of GBA is that when the learners attempt various problems in GBA, their interactions with the game provide evidence for underlying competency, and the gameplay simultaneously provides immediate feedback in response, motivating them to continuously modify their actions and strategies. This entry first discusses the advantages of GBA, then describes the use of evidence-centered design (ECD) for GBA. It concludes by discussing practical challenges faced in the use of GBA.

Advantages of GBA

The educational assessment community has recognized the needs for new kinds of assessment that (a) are based on modern theories of learning, (b) provide authentic real-world problems, (c) require application of multiple competencies, and (d) provide teachers and students with actionable information. Games, particularly video games, can be used as a vehicle for such assessments. Game and assessment design share similar principles of learning and employ compatible design processes. That is, game design focuses on creating mechanics that can continuously monitor and quantify players’ interactions with the game and provide feedback to the players or summarize their performance in relation to other players’ skills or resources. Similarly, educational assessment is the activity of observing what students say, do, or make, and of quantifying these observations in a meaningful way, to make more general inferences about their skills and knowledge.

Video games have great affordances for educational assessment for several reasons. First, playing video games is an integral part of daily life for many children and teens. A nationwide survey in 2015 found that 8- to 18-year-olds on average spend around 80 minutes each day playing video games, including games played on console and handheld video game players, computer games, and mobile games. Second, large amounts of data generated from gameplay can be rapidly collected without interrupting the learners’ engagement in video games, which means assessment can be seamlessly embedded in their daily activity. Third, this ability to extract data can yield rich, comprehensive student models, which can be used to diagnose students’ learning needs, provide formative feedback, and change gameplay to maximize learning according to the player’s ability level. Fourth, GBA employs challenging problems involving the types of complex situations necessary to evaluate the application of 21st-century competencies that are often underemphasized in conventional school education. Finally, when people are engaged and motivated with a given task, more accurate inferences can be made about them.

ECD for GBA
To create a well-designed GBA, game and assessment designers need a common language to align game and assessment mechanics. ECD, an assessment design framework, has been widely adopted by the educational game community to develop assessment models for GBA, which in turn support the balance between game design and assessment design. The central principle of ECD is that educational assessment is an evidentiary argument. ECD guides the design and implementation of assessment as a principled process by formalizing the assessment structure to systematically align students’ in-game actions with the specific competencies about which the assessors wish to make inferences. ECD is a process of addressing three questions that should be asked in any assessment design: what, where, and how are we measuring. This process leads to several design objects including competency, task, and evidence models, as follows:

A **competency model** (CM) directly reflects the types of claims that the assessor wishes to make about students at the end of the assessment. Typically, one CM is used for a given assessment, but ECD explicitly assumes multidimensionality of CMs. In GBA, CMs represent students’ skills, knowledge, and other traits for which a given game can provide evidence.

A **task model** (TM) involves tasks, which are individual units of activity attempted by the student. The student’s interactions with the task produce a *work product* that is then scored and used to update the assessor’s inferences about the student’s competency. A work product is an object that students produce as they respond to or interact with the assessment. A work product can be as simple as a response to a multiple-choice item in conventional assessment or as complex as a series of actions and choices in interactive environments such as video games. A TM is a collection of the task features (i.e., TM variables) that the assessment designer must consider when engineering the contexts necessary to elicit evidence of the targeted aspects of competency. Each TM must have different levels of evidentiary strength or focus. Therefore, each TM variable has a range of possible values and provides one or more functions that influence the argument structure of the assessment.

An **evidence model** bridges CMs and TMs by specifying the student work products and associated scoring rules and by using statistical models that send the collected information to the CM. An evidence model includes two processes. The first process is an evaluation component that considers the salient features and values of work products for an evaluative outcome. This process involves evidence rules, which are comparable to scoring rubrics. In GBA, evidence rules can be specifications of players’ observable behaviors in the game, and how these behaviors afford evidence with different levels of strength. The second process is the statistical component that analyzes how the obtained new evidence relates to CM variables in probabilistic terms.

**Practical Challenges**

Increasingly, educational game researchers and practitioners are emphasizing the importance of aligning students’ learning with what students do in games. However, it is often unclear how these researchers and practitioners leverage assessment to conceptualize game design around the competency of interest, even if they claim to use ECD. Therefore, there is a need for greater communication between design teams and the broader community. This communication must address the diverse methods and processes by which design teams, which often include learning scientists, subject-matter experts, and game designers, can seamlessly integrate design thinking and the formalization of assessment models. Some specific challenges that researchers and practitioners might face include the following:
How can assessment models be formalized?
How can formalized assessment models be translated into game design elements?
At what point(s) in the game design process does this translation occur most effectively?
How can CMs be transformed into interesting, engaging game mechanics?
How can psychometric qualities be ensured without being too prescriptive?

Furthermore, because GBA design requires the satisfaction of both psychometric and entertainment criteria, it is based on the assumption that GBA can offer a “sweet spot” that simultaneously meets these two different sets of criteria. However, little is known regarding how game and assessment designers can balance the design considerations of games versus assessment to maximize the effectiveness of GBA without losing game-like characteristics such as fun and engagement. Both researchers and practitioners need to develop an archive of design patterns and design principles that are specific to GBA.

See also Computer-Based Testing; Computerized Adaptive Testing; Evidence-Centered Design

Yoon Jeon Kim
http://dx.doi.org/10.4135/9781506326139.n281
10.4135/9781506326139.n281

Further Readings


