Stealth Assessment and Digital Learning Game Design

Chapter 10

Authors: Ginny L. Smith, Valerie J. Shute, Seyedahmad Rahimi, Chi-Pu Dai, and Renata Kuba

Digital games, commonly referred to as video games, are hugely popular around the world. According to the Entertainment Software Association (ESA, 2021), about 211 million Americans of all ages play video games, which is dwarfed by the estimated 1.52 billion players in the Asia Pacific region. Such games are often well designed, highly engaging, and require the players to apply various competencies to succeed (e.g., problem solving skills, creativity, and relevant content knowledge). Digital games can induce a state of flow, where players only attend to their gameplay and all else fades into the background (Csikszentmihalyi, 1990). At the turn of the century, the Federation of American Scientists (2006) along with other academics, like Vogel and colleagues (2006) and Gee (2003), recognized the potential of harnessing the powerful engagement of video games to support learning. However, adding purposefully designed learning content, supports, and assessment into a medium known for entertainment is challenging—especially with the goal of not ruining the fun factor or disrupting flow. More than a decade ago, challenges like this prompted the development of stealth assessment (Shute, 2011), as well as the now burgeoning field of digital game-based learning design.

Design Features of Digital Learning Games

Researchers in the field of digital game-based learning have sought to discover the magical concoction that successfully blends learning and fun in gameplay. Creating a well-designed digital learning game involves careful consideration of a set of game design features (Ke et al., 2019). The first feature is the core mechanic or game mechanic (i.e., how the player interacts in the game; Fullerton, 2019). The core mechanic generates gameplay, defining the goals, challenges, and feedback of the game (Adams, 2014; Ke et al., 2019). Once determined, design decisions associated with other features can be made. The narrative structure or game-world comprising the user interface is equally important to design. While not all digital learning games include a storytelling component, the game world must coherently link player actions, game objects, and goals through the core mechanic (Adams, 2014, Ke et al., 2019). Inclusion of in-game supports can further enhance learning and performance from gameplay (Ke, 2016; Wouters & van Oostendorp, 2013). However, an examination of how designers craft such game-based learning systems to include stealth assessment is needed (Shute & Ventura, 2013). How does the design of the core mechanic of a digital learning game impact the design and development of a stealth assessment? Must the design and development of the game and stealth assessment occur at the same time, or can the game design come first? Much of the game-based learning literature does not include discussion of the design decisions regarding the connections between the core mechanics to the assessment and support of learning during gameplay (Ke, 2016). The following chapter examines two cases of digital learning games and stealth assessments, one where the core mechanic was designed alongside the stealth assessment and one where the core mechanic was designed before the stealth assessment. We seek to illustrate the design process and feasibility of incorporating stealth assessment through these cases, so that game designers chose to incorporate stealth assessment in future digital learning games.

Stealth Assessment

When a person engages in gameplay, they generate copious amounts of data (e.g., time spent on a game level, specific actions taken). Stealth assessment takes advantage of this stream of information, allowing students to be assessed on a set of competencies without interrupting their gameplay. Burying the assessment in the game makes it invisible, to the point where learning and assessment become blurred (Shute & Ventura, 2013). Moreover, while stealth assessment can be used to measure final outcomes, it is also intended to be formative, supporting the development of competencies during the learning process (i.e., gameplay). Consequently, games that incorporate a stealth assessment can offer real-time assessment of students' learning and progress, to both the students and instructor using gameplay data from log files (e.g., Shute, Rahimi et al., 2020). These real-time student models are personalized versions of a competency model developed for the game through evidence-centered design.

Evidence-centered design (ECD; Mislevy et al., 2003) provides the theoretical framework for building stealth assessment models embedded in digital games. The ECD process starts with developing the competency model. The competency model is the representation of the theoretical concepts being assessed. Developing the competency model involves identifying and structuring relevant variables into meaningful relationships. The competency variables (e.g., Newton's First Law in a physics understanding competency model) comprise the knowledge or skills to be measured by the assessment. The next step in the ECD process is to consider the evidence necessary to make valid claims about student competency. The evidence model establishes the specific relationships between the competency variables and their associated metrics (e.g., Newton's First Law and its related gameplay behaviors). The evidence model is the link between the competencies and the tasks students perform within gameplay. The task model defines the specific features of tasks that will elicit the necessary evidence (e.g., the specific actions of the gameplay behaviors). The use of evidence-centered design as the framework for building the stealth assessment ensures alignment between the learning activity and assessment by linking student actions to the measured competency variables (Mislevy et al., 2003).

While a stealth assessment must be closely connected with the gameplay and learning tasks, the literature is limited regarding the design process. This chapter will focus on the design process of stealth assessment in the context of two well-designed digital learning games. We will illustrate our points with two case studies highlighting work we have done in the past relative to designing stealth assessments (a) where the stealth assessment and game were developed concurrently, and (b) where the stealth assessment was developed after the game was developed. By presenting the two unique cases, we contribute to the discussion on the design process for stealth assessment in digital game-based learning. Before presenting the case studies, we first describe the general step-by-step process for developing stealth assessments. In addition to the use of ECD, using a more delineated process highlights other key features of stealth assessment such as providing real-time feedback and learning supports directly within the game. Explanations of the first nine iterative steps, summarized by Shute, Ke, and Wang (2017) and an additional step 10 are shown in Table 1. Each step can be revisited and revised based on experts' feedback and playtest results. The process involves a team of learning scientists, game developers, instructional designers, measurement experts, and content experts.

Table 1.

10-step process to accomplish stealth assessment.

Step	Action/Process
1	Develop the competency model (CM) of targeted knowledge, skills, or other attributes based on literature and expert reviews. Generating multiple versions of potential models (at least two) allows for more in-depth discourse with content experts through model comparison.
2	Determine (or develop) the game (or digital learning environment) where the stealth assessment will be embedded. Having access to the game source code is best for embedding stealth assessment seamlessly.
3	Delineate a full list of relevant gameplay actions/indicators that serve as evidence to inform the CM. This can be done through consultation with content experts, engaging in extensive gameplay, and/or watching videos of expert solutions. Knowing how to accurately identify different in-game behaviors provides links between gameplay and associated competencies.

4	Create new tasks in the game, if necessary (i.e., the task model, or TM). Specific tasks are designed to elicit evidence (i.e., in-game behavior) of targeted competencies.
5	Create a Q-matrix (Tatsuoka, 1983) to organize and align the tasks and competencies (e.g., game tasks in rows and competencies in columns). The matrix helps establish a balanced set of tasks in the CM. Including task difficulty and discrimination estimates in the matrix make it an augmented Q-matrix (Kim, Almond, & Shute, 2016).
6	Determine how to score indicators using classification into discrete categories (e.g., yes/no; poor, okay, good, very good – relative to quality of the actions). This becomes the automated scoring rules part of the evidence model (EM). See Shute and Wang (2016) for some examples of scoring indicators.
7	Establish statistical relationships between each indicator and associated levels of the CM variables. This is the statistical model part of the EM. Subject-matter experts can provide a priori input regarding relationships.
8	Pilot test and modify parameters. The pilot data can provide valuable empirical data for updating the evidence model for better reliability and alignment.
9	Validate the stealth assessment with external measures (e.g., employ a pretest/posttest of content knowledge/skills). Comparing the results of the stealth assessment and pretest/posttest data illuminates any flaws in the model and gives direction for model improvement (see Shute et al., 2020).
10	Use the current information about a player's competency states to provide adaptive learning support (e.g., targeted formative feedback, progressively harder levels relative to the player's abilities). This represents ongoing research. The goal is for the learning supports (cognitive and affective) to improve learning and performance, without disrupting flow.

An accurate stealth assessment can be designed, developed, and implemented via these ten steps to measure and support a variety of content area knowledge and skills. In this chapter, we will focus on the first five steps of the process when discussing the two cases. Steps six through ten are not concerned with the design process of the game and stealth assessment. They focus on building the internal mechanics, validating the assessment, and adapting learning supports and thus are out of the scope of this chapter. Please note, we also skipped Step 2, as the games in both cases were predetermined. For more information on developing the internal mechanics of a stealth assessment such as the scoring engines and statistical models (i.e., Steps 6 and 7), see chapter 12.

Case 1: Concurrent Design and Development of a Game and Stealth Assessment

Physics Playground (Shute et al., 2019) is a 2-dimensional, web-based game intended to assess and support 7th to 11th grade students' Newtonian physics understanding. The goal across the numerous game levels is always the same——move the green ball so that it hits the red balloon. The current version of *Physics Playground* is the result of an iterative process through which we designed and developed the game and the stealth assessment of physics understanding. We describe this process below. We start the discussion with an overview of the *Physics Playground* design and development team. We then examine steps one, three, four and five of the stealth assessment design process. We conclude with a discussion of the components of the current version of the game.

Physics Playground Team

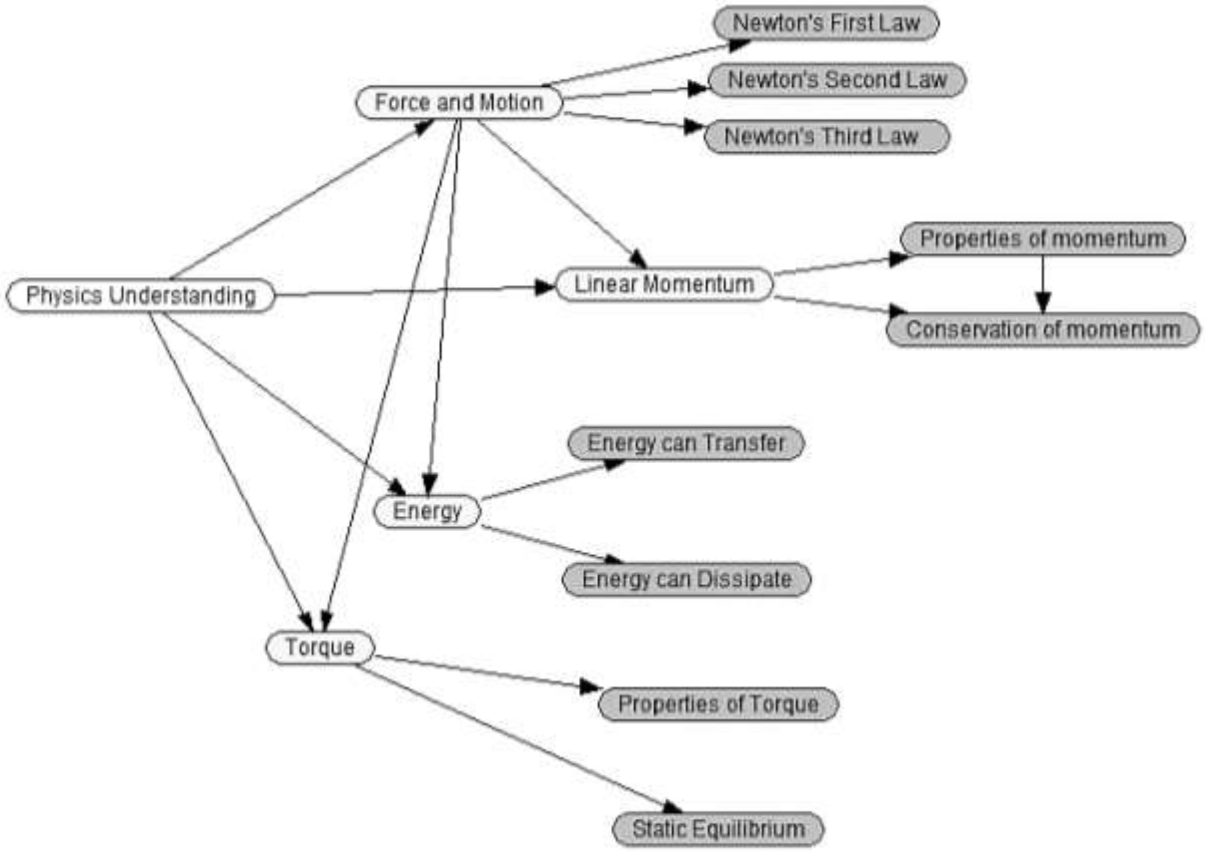
To design and develop the game and the stealth assessment of physics understanding within the game, the research team included experts in assessment, instructional design, learning game design, multimedia design, game development, and three physics education experts. The *Physics Playground* team was diverse including faculty and graduate students with various ethnicities, different areas of expertise, and rich experiences across the education landscape. The team met biweekly for over three years during the design and development project. The research team was also split into smaller teams to address specific design and development project components. For example, a learning support team was created to lead the design and development of the learning supports and the external test items of physics understanding. We had a programming team in charge of developing and preparing the game for studies. We also had a measurement team in charge of implementing and evaluating the stealth assessment machinery on the server. Each small team met weekly to focus on their specific parts of the project, and all small teams presented an update on their progress at the whole team bi-weekly meetings so team members could give and receive feedback, and the team could make decisions.

Step 1: Designing and Refining the Competency Model of Physics Understanding.

At the start of the project, *Physics Playground* had one type of game level (i.e., sketching levels) with two game mechanics that players could use to move the green ball: drawing objects on the screen and clicking on the ball to nudge it left or right. This version of the game started with an interactive video tutorial demonstrating how to use the nudge mechanic to move the green ball and how to use the drawing mechanic to draw four simple machines—ramps, levers, pendulums, and springboards—to move the green ball. While these game mechanics allowed for coverage of several key Newtonian physics competencies, one goal of the project was to expand the physics in the game to cover a more complete set of competencies. We planned to incorporate the following competencies into the game: (a) Newton's three laws of force and motion; (b) potential and kinetic energy; (c) torque and conservation of angular momentum; (d) collisions and conservation of linear momentum; and (e) energy and dissipative forces. Based on the Next Generation Science Standards (2013), the design and development team built several different physics competency models. Each model was discussed and analyzed for feasibility and coverage of competencies. After much debate and evaluation, the team selected the model shown in Figure 1 to implement within the game.

Figure 1

The physics understanding competency model for Physics Playground.



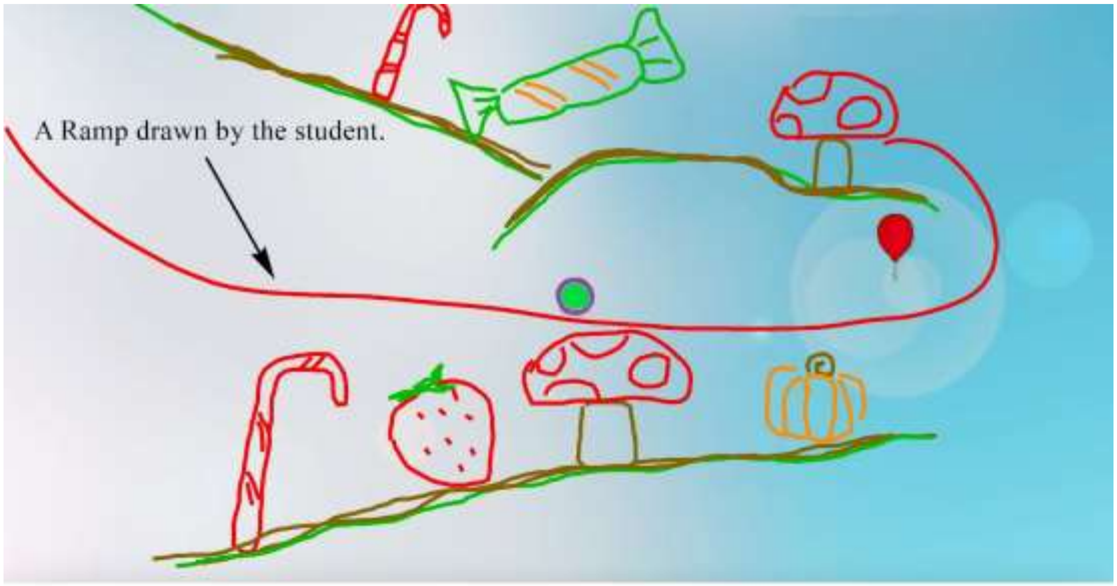
Step 3: Identify Game Mechanics and Behaviors Aligned with the Physics Competencies

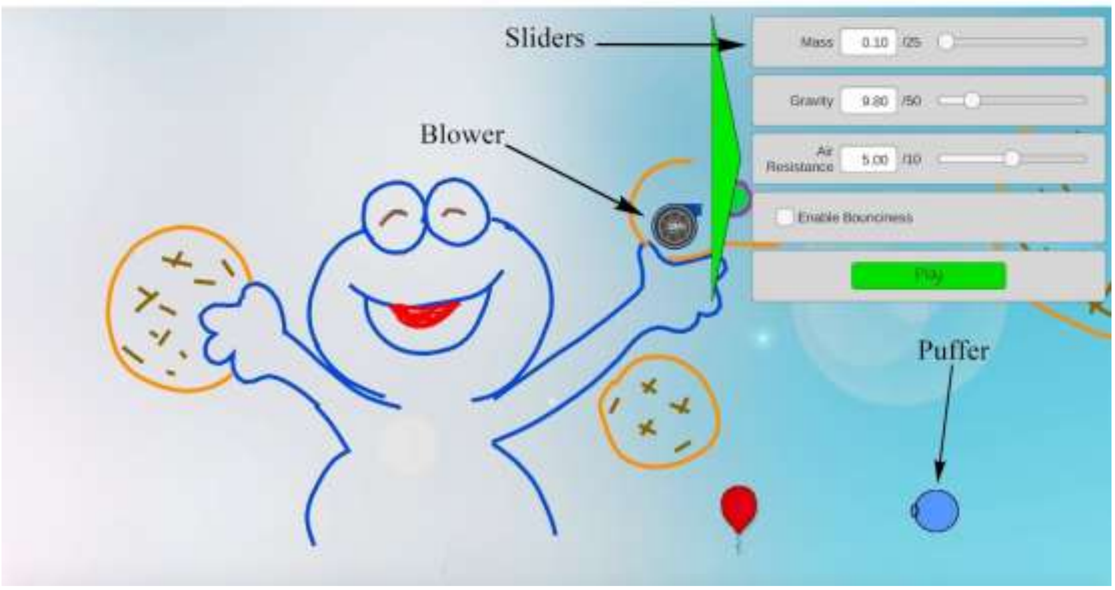
Upon examining the previously designed set of sketching levels in the game and the new competencies, we realized that some competencies (e.g., Newton's second and third laws of force and motion) were not adequately covered. Based on the ECD framework we needed to identify player behaviors that would support claims of students' mastery levels of all the competencies in the game. From the drawing mechanic, several behaviors had previously been identified, such as the number of objects drawn, and whether the player drew one of the simple machines discussed in the tutorial. However, these behaviors were not able to provide evidence for some of the expanded learning content in the competency model. The team spent many months discussing, testing, and evaluating different game mechanics that could be used to cover the expanded set of physics competencies. One idea that was ultimately discarded was using prediction-based levels, where players had to predict the outcome of a level based on force and/or momentum (e.g., predicting the location on the screen where the ball would land). After multiple brainstorming sessions, rapid prototyping of ideas, and internal testing, the team ultimately decided on using a slider mechanism where players adjust the physics parameters of the level to move the ball.

The sketching levels were all restricted to the real-world values for gravity, air resistance, and a standardized mass of the green ball. By loosening these restrictions, we could measure students' understanding of how these parameters affect an object's motion. We added a panel with three sliders (Figure 2 bottom image)—one to adjust the mass of the ball, one to adjust gravity, and one to adjust air resistance. The panel also had a button to turn on and off affecting the bounciness of the ball, available in levels related to Newton's Third Law. Additionally, we needed new game mechanics beyond nudging to move the ball, thus we added adjustable and static blowers (i.e., fans with and without adjustable sliders) and a puffer that would release one puff of air per click. The bottom image of Figure 2 is an example of a game level showing the new game mechanics (e.g., sliders, a static blower, and a puffer). The addition of these new game mechanics allowed for new types of player interactions that we could use to assess students' physics understanding. We could now collect slider movements, the use of bounciness, and the use of puffers and blowers as observations for assessing students' physics understanding.

Figure 2

Example of a sketching level (top image) and a manipulation level (bottom image) in Physics Playground





Step 4: Creating New Game Tasks

The new game mechanics helped us create a new type of game level, dubbed manipulation levels. The team's game developers created a web-based level editor that any team member could use to create a sketching or manipulation level for the game (more information about the level editor is discussed in a later section). Team members created game levels and presented the levels at our biweekly team meetings. The game levels were revised iteratively based on feedback from the team and physics experts until approved for use in the game. Once a game level was approved it was added to the new Q-matrix.

Step 5: Creating/Updating the Q-Matrix

As we added manipulation levels and a few newly created sketching levels to our Q-Matrix, we looked for balance across the number of game levels per competency. The Q-matrix spreadsheet had the nine lowest-level physics concepts (darkest shade in Figure 1) in columns and the game level names in rows. Every game level was assigned a primary and secondary physics competency by the team of physics experts. In the augmented Q-matrix we assigned a value of 1 for the level's primary competency and a value of 2 for the level's secondary competency (see Table 2 for an excerpt). Moreover, each game level was also assigned two difficulty levels. One difficulty level was based on the type of game mechanic (GM) used to complete the level. The second difficulty level was based on the physics understanding (PU) used to complete the level. The composite difficulty score was the sum of GM and PU. Both difficulty levels were computed using a rubric scaling from 1 to 5 (see Shute, Rahimi et al., 2020 for more details) by multiple raters to ensure accuracy. Using the augmented Q-Matrix the team could identify if we have enough game levels and enough diversity of difficulty within the game levels to adequately cover each competency and player ability level. If we discovered a competency in need of more coverage, we returned to step 4 and created new game levels until we achieved adequate coverage to accurately assess and support all nine physics concepts. Table 2 contains an excerpt from the augmented Q-matrix.

Table 2

Excerpt of the Augmented Q-Matrix for Physics Playground

	N1L	N2L	N3L	POM	COM	ECT	ECD	POT	Equil.	GM	PU	Comp.

L1	1	2	0	0	0	0	0	0	0	2	3	5
L2	0	0	0	0	0	1	0	2	0	3	5	8
…	…	…	…	…	…	…	…	…	…	…	…	…
Ln	1	2	0	0	0	0	0	0	0	2	2	4

Note. L1, L2, etc. = different game level names; N1L = Newton First Law; N2L = Newton Second Law; N3L = Newton Third Law; POM = Properties of Momentum; COM = Conservation of Momentum; ECT = Energy Can Transfer; ECD = Energy Can Dissipate; POT = Properties of Torque; Equil. = Static Equilibrium; GM = Game Mechanics difficulty; PU = Physics Understanding Difficulty; Comp. = Composite difficulty score.

Current Version of Physics Playground

The current version of *Physics Playground* includes the original set of sketching levels that can be solved by drawing ramps, levers, pendulums, and/or springboards on the screen and nudging the ball, along with manipulation levels that can be solved by enabling the ball's bounciness, adjusting the ball's mass, air resistance, or gravity sliders, or by changing the amount of force exerted on the ball by a puffer and/or adjusting a blower. As mentioned, *Physics Playground* includes a level editor whereby non-technical users (e.g., students, teachers, stealth assessment designers) can create their own levels by drawing objects (e.g., lines and shapes) on the screen. Starting with an empty stage, one can place the ball and balloon anywhere on the screen and draw any number of obstacles between them. Using this level editor, we have created more than two hundred game levels targeting various physics concepts. The current version of *Physics Playground* also includes a set of learning supports and an incentive system which are out of the scope of this chapter (for more information on the learning supports see Shute, Smith et al., 2020; for more information on the incentive system see Rahimi et al., 2021).

Case 2: Design and Development of a Stealth Assessment Post Game Development

Variant: Limits™ (v1.0.1; 2017) is a 3D digital learning game designed to complement introductory college calculus courses by enhancing conceptual understanding through gameplay. The game is linear and organized into four Zones, each comprised of graphical puzzles centered around specific learning objectives. Students play as an onscreen character who has been tasked with restoring power to the planet where they awakened at the start of the game. A narrator provides guidance, and a plotted path is available if requested to show students the correct direction to head. The calculus is embedded in the learning puzzles that students must solve to complete pathways for the power or the player to move. Students progress through the four game zones, building on previous skills as they incorporate new skills in their solutions. As with the previous case, we will discuss the design and development of the stealth assessment using the first five steps of the iterative process described in Table 1. However, unlike the previous case, here we examine the process of creating the stealth assessment when the learning game is already developed.

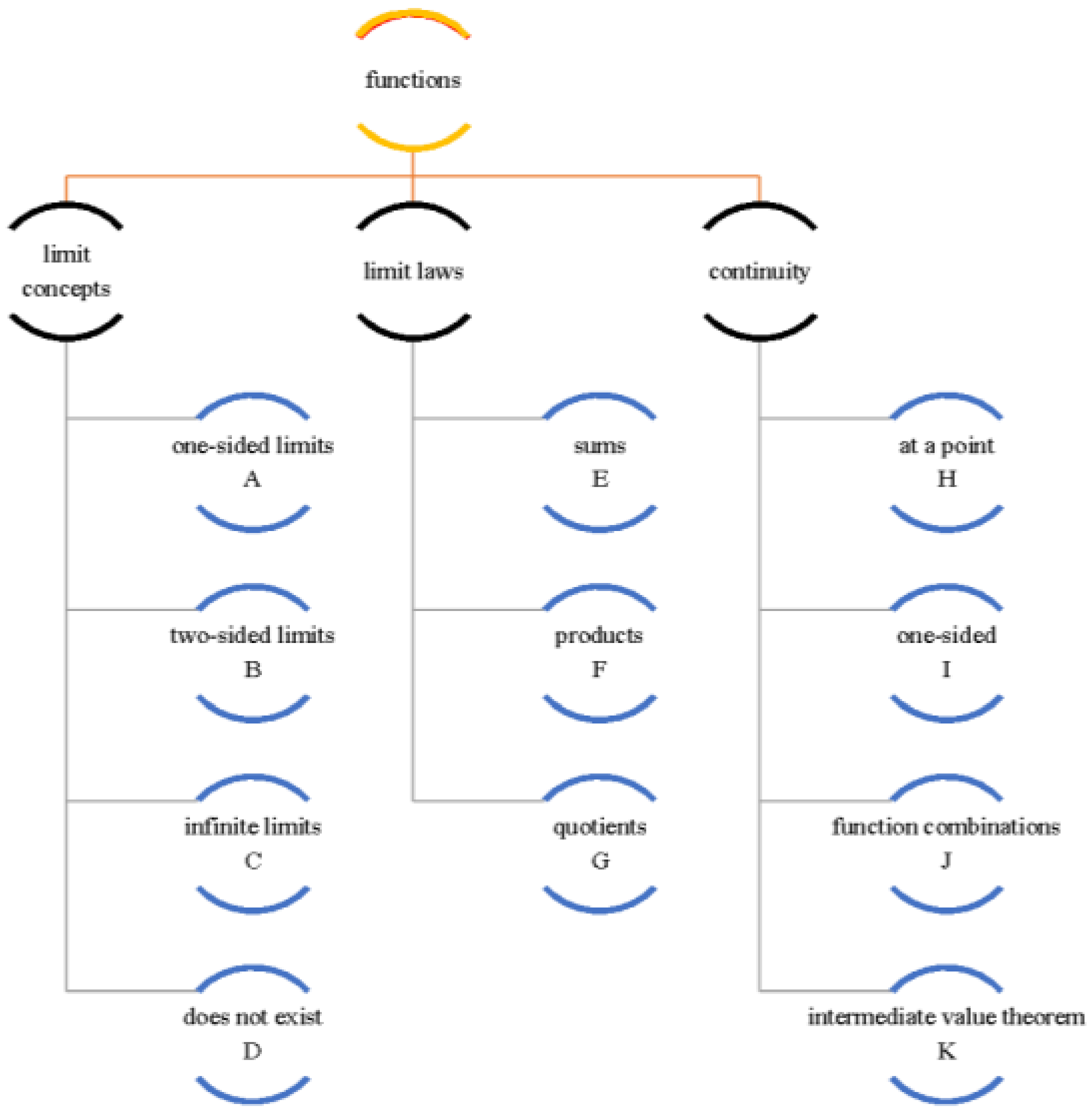
Step 1: Designing and Refining the Competency Model of Calculus Understanding

After developing *Variant: Limits*, the game design company, Triseum, contracted with our research lab to design and develop a stealth assessment measuring players' mastery of the targeted calculus topics and skills. The game was originally designed around a set of learning objectives, not a competency model. So, unlike the previous case, the competency model had to fit the existing game's content. To that end, we worked closely with an instructional designer from Triseum and the game's project manager. We were granted access to the game and its learning design documents to help discern the key calculus competencies. After dozens of hours of gameplay and reviewing the design materials, we created a prototype for the competency model.

We then worked with the company's calculus experts to iteratively refine the competency model. After three iterations, the team settled on the competency model illustrated in Figure 3 (originally published in Smith et al., 2019).

Figure 3

The calculus competency model for Variant: Limits



Step 3: Identifying Game Mechanics and Behaviors Aligned with Calculus Competencies

After finalizing the competency model, we examined each game puzzle to determine how the puzzle connected to the competency model, documenting the associated gameplay behaviors.

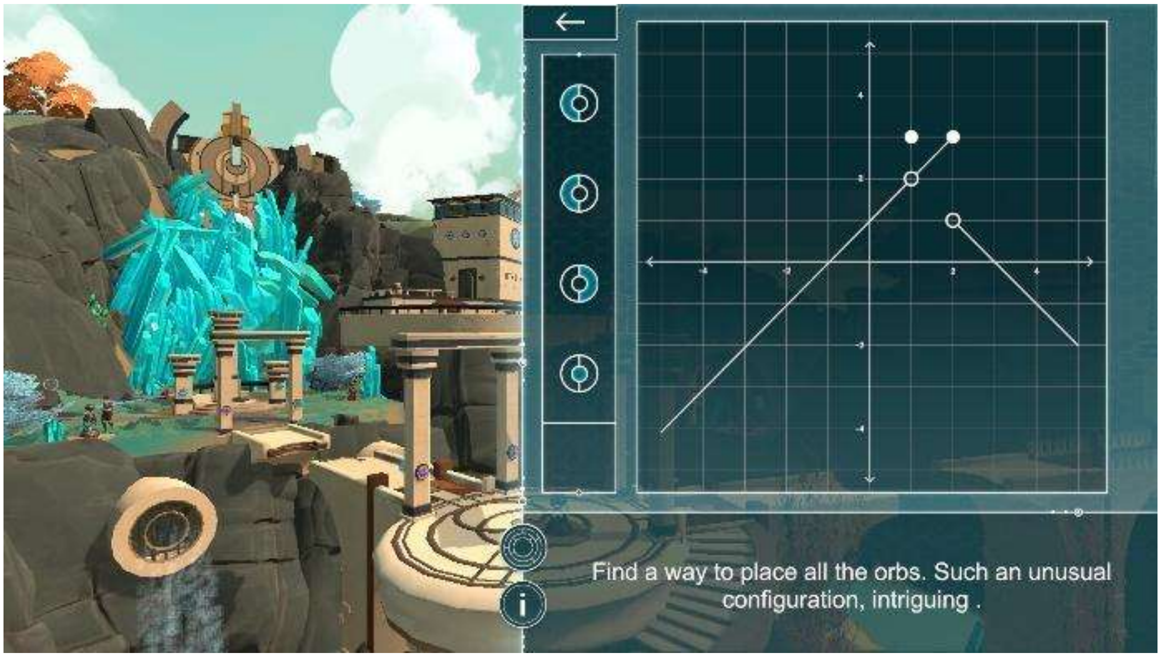
Each zone of *Variant: Limits* focuses on different calculus skills and introduces new puzzle types (i.e., game mechanics). There are five puzzle types used in the game: orb placement puzzles, orb editing puzzles, puzzles using input sliders, orb creation puzzles, and value placement puzzles. All puzzles are graphical in nature and appear on a coordinate plane.

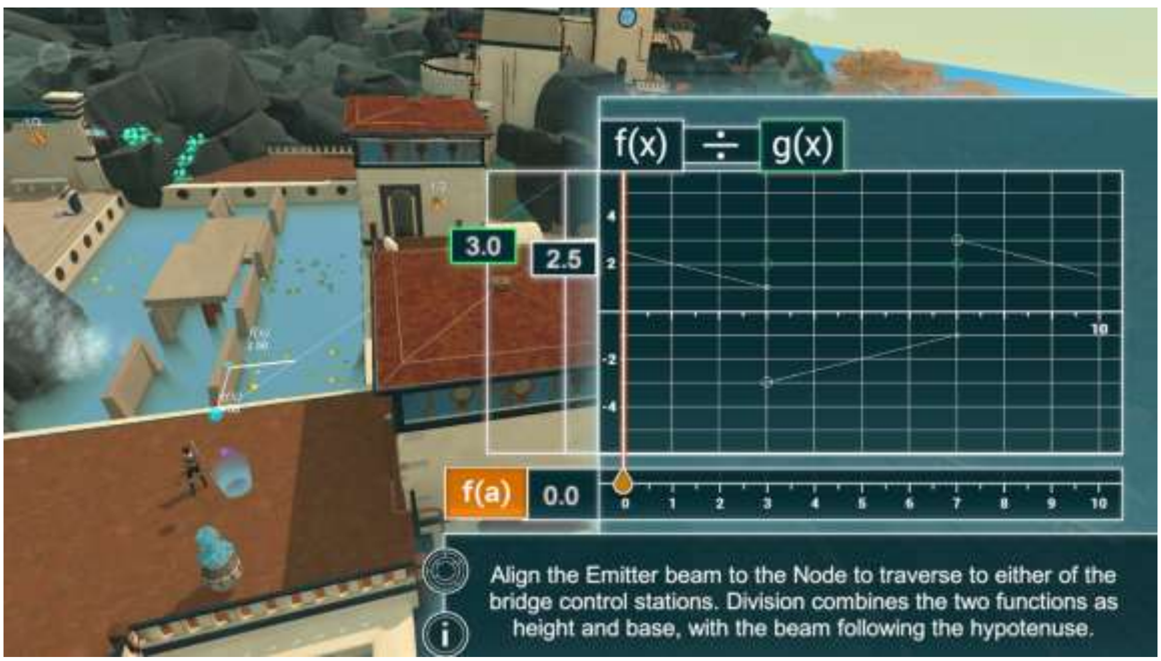
In the puzzles, “orbs” are points on a graph. Each orb is made up of three regions, an inner circle, an arc surrounding the left half of the inner circle, and an arc surrounding the right half of the inner circle. Based on the orb’s shading, the orb represents one of the possible function behaviors at a given point. If the left arc is shaded, that means the function is continuous to the left of that point. If the center circle is shaded, that means there is a function value at that point. And if the right arc is shaded, that means the function is continuous to the right of that point.

For orb placement, orb editing puzzles, and value placement puzzles the graphs are marked with points of interest and players must move the correct orb or value to each point. See Figure 4 (top image) for an example of an orb placement puzzle. For puzzles with input sliders, students must identify the correct function value (i.e., location on the horizontal axis) to solve the puzzle (Figure 4 bottom image). For orb creation puzzles, players must determine the type and number of orbs to create for a given input value in a function combination. Players are given unlimited attempts to solve the puzzles. Each puzzle also contains an information button that describes that puzzle’s game mechanic in case students had forgotten what to do. As the game is linear, players must complete each puzzle to move forward in the game. Like the previous case, the different game mechanics allow for coverage of different competencies within the game.

Figure 4

Example of an orb placement puzzle (top image) and a puzzle with input sliders (bottom image) in Variant: Limits





Once we identified the game mechanics, we organized and defined the gameplay behaviors connected to a player's mastery of the calculus competencies. The most common gameplay behavior was moving an orb or value to its correct location. The other main gameplay behavior was changing the slider values. For orb placement puzzles, the observed behavior was moving the correct orb(s) to the correct location(s). The observed behavior was the same for orb editing puzzles, but there was an added gameplay behavior of first creating the correct orb. For value placement puzzles the observed behavior was moving a value (i.e., number) not an orb. For puzzles involving sliders, the observed player behavior was the player's changing of the slider value. Finally for orb creation puzzles, the identified gameplay behavior was how accurately the player generated the orbs as well as the accuracy of the player's submitted solution.

Step 4: Revising Game Tasks

While the stealth assessment team did not have the ability to create new puzzles (i.e., game tasks), we were able to suggest revisions to existing puzzles to ensure adequate coverage of the 12 calculus competencies in the stealth assessment competency model. For example, we requested points of interest be moved or the shape of the function changed, so the solution required the use of different calculus knowledge and thus addressed a different competency. We also requested that the programmers add *hooks* (i.e., a marking mechanism to capture certain gameplay behaviors) in the logfile data recorded in the game. For example, in the orb editing puzzles, we requested that the logfile capture the number of attempts a player made when editing the orb as well as the number of attempts a player made when placing the orb in its correct location.

Step 5: Creating the Q-Matrix for Variant: Limits

The 12 calculus competencies and differing game mechanics were introduced in the game as students progressed. For example, the first Zone contained only orb placement puzzles and gameplay was linked to only four of the twelve calculus competencies in the model (i.e., Competencies A, B, H, and I, see Figure 3). Players must progress past certain points in the game to accumulate any evidence for later competencies. Therefore, the design of the first Q-matrix for

Variant: Limits was organized by Zone, and each puzzle within the Zone was matched to one or more of the calculus competencies in the competency model. After this first level of organization, an additional Q-matrix was created grouping the puzzles by their competency. Once adequate coverage of the competencies was confirmed, we used the Q-matrix grouped by competency to further delineate each puzzle into its different math instances (e.g., number of orbs in the puzzle). This spreadsheet was ultimately transformed into the scoring rubric created in steps 6 and 7 of the stealth assessment. For more information on the scoring and validation of the stealth assessment for *Variant: Limits* see Smith et al. (2019).

Discussion

The two cases discussed in this chapter illustrate the versatility of stealth assessment and digital game design. Each case required a different design and development process, with different tasks needed to complete the steps to accomplish a stealth assessment. However, in the first case with *Physics Playground*, the design and development effort occurred simultaneously, as the team made decisions about what types of game mechanics would allow for assessment and support of the expanded competency model. The possibilities were limited only by the creativity of the research team. Research team members spent hours brainstorming, prototyping, and debating the best gameplay interactions (i.e., game mechanics) to elicit observable evidence of players' understanding of the nine physics concepts. In the second case with *Variant: Limits*, the game mechanics were designed and developed *before* the stealth assessment. The major task for the stealth assessment team was identifying the competencies that could be measured by the existing game mechanics. The stealth assessment team spent hours playing the game, reading the learning design materials, and developing a competency model that could assess and support the existing calculus content in the game. However, while the specific design and development process differed between the two cases, the result of both cases was a successfully designed and implemented stealth assessment.

The differing design and development processes also required different focus and expertise within the team to accomplish the tasks. The research team for *Physics Playground* was large and diverse with multiple small teams (e.g.., learning support team, measurement team) working on different aspects of the project simultaneously. Expertise in learning game design, assessment, programming, instructional design, and physics education was needed to create the current version of *Physics Playground*. And, while the same amount of expertise was needed to create *Variant: Limits*, most of it was needed prior to the stealth assessment team joining the project. This meant that the *Variant: Limits* stealth assessment team was small, and the main expertise needed on the team was assessment design.

As the examination of these two stealth assessment designs illustrate, the design and development process differs when concurrently creating a stealth assessment and learning game compared to creating a stealth assessment for a preexisting game. However, one important similarity of these two design cases is that both teams had access to the game's source code. For both the *Physics Playground* stealth assessment of physics competencies and the *Variant: Limit* stealth assessment of calculus competencies, it was necessary to create new game mechanics or make changes to the game to adequately cover the competencies assessed and supported through gameplay. In the case of *Physics Playground*, we could accomplish these iterative steps because we developed *Physics Playground* ourselves. In the case of *Variant: Limits*, we had secondary access to the game's source code by requesting changes to the puzzles and additional data hooks. Lack of access to a game's source code can limit the design and development of a stealth assessment. Therefore, in both cases it was important that the game developers were either part of the team or accessible to the team.

Limitations

Although the cases we presented here shed light on the design process of digital learning games and stealth assessment, our examination contains some limitations. The chapter's focus is on the design process; therefore, it is mostly descriptive with references to related research studies provided for more information. Our goal was not to empirically rate the two design processes, but to make them visible to readers. Game-based learning literature often does not include the design and development process of the learning game but focuses instead on the finished game and its performance (Ke, 2016). However, future studies might examine digital game and stealth assessment design through a different lens such as design efficiency or resource use. Also, we examined the design process when the game and stealth assessment are designed together and when the stealth assessment is designed for a preexisting game. Further insights could be gained from examining the stealth assessment design process across different fields or different populations. Including more design cases would also provide more insights into stealth assessment design. The clearer the landscape of stealth assessment design, the more available stealth assessment becomes to other designers, educators, and researchers.

Implications

Based on the two design cases discussed in this chapter, it is clear that game design plays a large role in the process of creating stealth assessments for digital learning games. The advantages of designing the stealth assessment alongside the learning game is that the possibilities are endless on what can be created using ECD and the 10-step process described here. However, the same process can also be applied to preexisting games with success. As the field of designing digital learning games continues to grow, the use of stealth assessment can grow with it if the process continues to be made clear. Then students and instructors across content areas and populations can make use of the copious amount of data available through gameplay when designing or using digital learning games.

References

Adams, E. (2014). *Fundamentals of Game Design* (3rd ed.). Pearson Education Inc.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* Harper Perennial.

Entertainment Software Association. (2021). *2021 Essential facts about the video game industry*. https://www.theesa.com/resource/2021-essential-facts-about-the-video-game-industry/

Federation of American Scientists. (2006). *Summit on educational games: Harnessing the power of video games for learning.*

Fullerton, T. (2019). *Game design workshop: A playcentric approach to creating innovative games* (4th ed.). CRC Press.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.

Ke, F. (2016). Designing and integrating purposeful learning in game play: a systematic review. *Educational Technology Research and Development*, *64*(2). https://doi.org/10.1007/s11423-015-9418-1

Ke, F., Shute, V., Clark, K. M., & Erlebacher, G. (2019). *Interdisciplinary design of game-based learning platforms: A phenomenological examination of the integrative design of game, learning, and assessment*. Springer International Publishing. https://doi.org/10.1007/978-3-030-04339-1

Kim, Y. J., Almond, R. G., & Shute, V. J. (2016). Applying evidence-centered design for the development of game-based assessments in Physics Playground. International *Journal of Testing*, *16*(2), 142-163. doi:10.1080/15305058.2015.1108322

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the Structure of Educational Assessments. *Measurement: Interdisciplinary Research and Perspectives*, *1*(1), 3–62. https://doi.org/10.1207/S15366359MEA0101

Next Generation Science Standards Lead States (2013). *Next Generation Science Standards: For States, By States*. The National Academies Press. https://doi.org/10.17226/18290

Rahimi, S., Shute, V., Kuba, R., Dai, C.-P., Yang, X., Smith, G., & Alonso-Fernández, C. (2021). The use and effects of incentive systems on learning and performance in educational games. *Computers & Education*. 165, 1-17. https://doi.org/10.1016/j.compedu.2021.104135

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–523). Information Age Publishers.

Shute, V. J., Almond, R. G., & Rahimi, S. (2019). *Physics Playground*. [digital game]

Shute, V. J., Ke, F., & Wang, L. (2017). Assessment and adaptation in games. In *Instructional techniques to facilitate learning and motivation of serious games* (pp. 59–78). Springer International Publishing. https://doi.org/10.1007/978-3-319-39298-1_4

Shute, V. J., Rahimi, S., Smith, G., Ke, F., Almond, R., Dai, C.-P., Kuba, R., Liu, Z., Yang, X., & Sun, C. (2020). Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games. *Journal of Computer Assisted Learning*, *37*(1), 137-141. https://doi.org/10.1111/jcal.12473

Shute, V. J., Smith, G., Kuba, R., Dai, C. P., Rahimi, S., Liu, Z., & Almond, R. (2020). The design, development, and testing of learning supports for the Physics Playground game. *International Journal of Artificial Intelligence in Education*, 127–141. https://doi.org/10.1007/s40593-020-00196-1

Shute, V. J., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. The MIT Press.

Shute, V. J. & Wang, L. (2016). Assessing and supporting hard-to-measure constructs. In A. A. Rupp, & J. P. Leighton (Eds.), *The handbook of cognition and assessment: Frameworks, methodologies, and application,* (pp. 535-562). Hoboken, NJ: John Wiley & Sons, Inc.

Smith, G., Shute, V., & Muenzenberger, A. (2019). Designing and validating a stealth assessment for calculus competencies. *Journal of Applied Testing Technology*, *20*(S1), 1–8. www.jattjournal.com

Tatsuoka, K. K. (1983). Rule space: an approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, *20*(4), 345–354.

Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers and Education*, *60*(1), 412–425. https://doi.org/10.1016/j.compedu.2012.07.018

Variant: Limits (Version 1.0.1; 2017). [Computer Software]. Triseum.

Vogel, J. J., Vogel, D. S., Cannon-bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229–243.