CHAPTER 10

ASSESSING LEARNING IN VIDEO GAMES

Valerie Shute, Matthew Ventura, Yoon Jeon Kim, & Lubin Wang

Video games are bad for you? That's what they said about rock and roll. Shigeru Miyamoto

Video games are hugely popular. For instance, revenues for the video game industry reached 7.2 billion in 2007, and overall, 72% of the population in the U.S. plays video games.[1] The amount of time spent playing games also continues to increase.[2] Considering the fact that many college students are avid game players, it is not surprising that institutions of higher education have begun slowly to accept video games.[3] That is, to meet the needs of a new generation of students entering college, educators and researchers have begun searching for ways to creatively incorporate video games into their courses.

Besides being a popular activity, playing video games has been shown to be positively related to a variety of cognitive skills (e.g., visual-spatial abilities; attention), personality types, academic performance, and civic engagement.[4] Video games can also motivate students to learn valuable academic content and skills, within and outside of the game.[5] Finally, studies have shown that playing video games can promote prosocial and civic behavior.[6]

In this chapter, we describe how well-designed video games can be used as vehicles to assess and support learning across a variety of knowledge and skills. We also present a framework for designing such embedded assessments into video games, and illustrate the approach with examples from a physics game. We conclude with our thoughts on future research in this area. Throughout the chapter, we use the term *video* *games* to refer to games that are played on gaming consoles, computers, and mobile devices.

Educational Benefits of Video Games

Before discussing ways to use games to assess learning, we first need to define what we mean by learning. Our conception of learning is that it is a lifelong process of accessing, interpreting, and evaluating information and experiences, and then transforming that into knowledge, skills, conceptualizations, values, and dispositions. This is a broader and more fluid view of learning compared to more conventional (e.g., cognitive, behavioral) perspectives which tend to ask: "Did the student remember X and Y on the test?"

Learning also represents a change from one point in time to another in terms of knowing, doing, believing, and feeling. Learning is not necessarily linear; that is, knowledge can start off very shallow, and then quickly explode into a rich knowledge base over a relatively short period of time. For example, in language learning, people start by learning a few words, but in a span of a year, that number can increase to hundreds of words and phrases, in addition to grammatical knowledge.[7]

The learning theories that best suit educational game design include socio-constructivism and situated learning.[8] Based on these theories, the learner is active in the learning process, where "doing" is more important than listening, and the learner determines the pace of learning. Moreover, learning in many cases is the result of interactions with a problem context where learners actively construct meaning in the process of solving problems, large and small.

How can video games foster learning? Video games can be seen as vehicles for exposing players to intellectual activities. Much like taking a course in college or playing a sport, video games engage players in activities that require intellectual effort. People who want to excel at something—from athletes to dancers to surgeons to computer programmers—spend countless hours practicing their craft. By continually refining techniques and developing new maneuvers to enhance their skills, they manifest the belief that practice is critical to improvement. There is considerable support in the literature, going back more than 100 years, for the idea that "practice makes perfect," or in its less extreme form, that "practice makes better".[9] The common conclusion across all of this work is that people become more accurate and faster the more often they perform a task. Content learning or skill acquisition thus represents a change in a person that occurs as a function of experience or practice. But practice can be boring and frustrating, causing some learners to abandon their practice and hence learning. This is where the principle of game design comes in – good games can provide an engaging environment designed to keep practice interesting.

Can important skills, like problem solving, really be improved by playing video games? Polya has argued that problem solving is not an innate skill, but rather something that can be developed, "Solving problems is a practical skill, let us say, like swimming… Trying to solve problems, you have to observe and imitate what other people do when solving problems; and, finally, you learn to solve problems by doing them."[10] Students are not born with problem solving skills. Instead, these skills are cultivated when students have opportunities to solve problems proportionate to their knowledge. Additionally, cognitive complexity theory predicts that video games should lead to learning because they simultaneously engage players' affective and cognitive processes.[11] These affective processes can be seen as dependent on how engaging a video game is, where engagement is a function of the core principles of good game design working in concert.[12] Some of the features of good games include: adaptive challenges, goals and rules, interactive problem solving, control (of learning and the game environment), ongoing feedback, and sensory stimuli.

Given the feature of presenting adaptive challenges to players, video games can actually cause a state of frustration (or "pleasant frustration").[13] In good games, obstacles, constraints, and generally wicked problems become something that we want to resolve because reaching for goals and ultimately succeeding is highly rewarding. McGonigal has referred to this as a positive kind of stress, called eustress, which is actually good for us, providing us with a sense of motivation and desire to succeed.[14] We see this pleasant frustration (or eustress) as a positive aspect of video games because it shows that students are being pushed to their limits, a requirement for teaching in the zone of proximal development.[15] This repeated frustration can also prepare students for tolerating frustration, a common emotional response in higher education.

Consider, for example, the popular game *Portal 2*. In *Portal 2,* the player has to navigate through a 3D environment with a "portal gun" that allows the player to teleport by shooting "portals" into special walls. In the game, the player must first determine the spatial environment and then use various tools and the portal gun to open a door. Frustration can arise quickly from: (a) not knowing the spatial environment, (b) not understanding the problem, (c) not knowing how to the use various tools to open the door, and (d) a recurring character in the game that taunts the player with insults. In many cases this frustration can be overwhelming after repeated failures. However, this type of failure and frustration is important to experience (and overcome) because it helps students prepare for the challenges of higher education, and life in general.

Thus, video games have the power to help students cope with frustration by repeated exposure to challenging intellectual activities. Additionally, educators who wish to take advantage of the potential of video games to support their students' learning may not know how to employ assessment in (or with) the game to capture information concerning what, how, and to what extent players are learning from the games. The next section introduces a systematic framework that can be used to create valid assessments to be embedded within video games.

Developing Good Educational Video Games

Video games typically require a player to apply various competencies to succeed in the game (e.g., creativity, problem solving, persistence, and collaboration). The competencies required to succeed in many games also happen to be the same ones that are needed to succeed in higher education, and that companies are looking for in today's highly competitive economy.[16] But for video games to gain a footing in education there needs to be more collaboration among educators, researchers, and game designers. Having a shared understanding of these currently rather isolated specialty areas is critical to move forward with the design of engaging educational games.

Collaborative research and development should include the right balance of educators and researchers – with expertise in assessment, learning, and content – as well as game designers to optimize the development of well-designed educational games. These games would allow for the assessment and support of learning in an engaging way within a rich and authentic context. Knowing what (and how well) students are learning in a video game is a function of sound assessment practices. That is, the primary purpose of an assessment is to collect information that will enable the assessor to make inferences about a person's competencies – what they know, believe, can do, and to what degree. Evidence Centered Design (ECD) is one such approach that is suitable for building valid and reliable assessments which ultimately may be embedded in video games to monitor and support learning.[17]

ECD is an assessment design framework that consists of three main models that work in concert: (a) competency model, (b) evidence model, and (c) task model. A good assessment (which could be a video game) elicits behavior that bears evidence about key competencies, and it must also provide principled interpretations of that evidence in terms that suit the purpose of the assessment.[18] To build an ECD assessment, the competency model is first defined. This represents a set of psychological constructs on which inferences are based.[19] These constructs can be knowledge, skills, dispositions, beliefs, or whatever you want to assess.

The evidence model attempts to answer the question about what behaviors or performances serve as evidence for variables within the competency model. Task models describe features of situations that will be used to elicit performance. The main purpose of tasks, such as levels or quests in a game, is to elicit dynamic evidence (which is observable and empirical) about competencies (which are unobservable and theoretical). Results from the task model in a traditional assessment consist of a set of items or problems; but in the context of video game-based assessments, task modeling produces a collection of problems designed to capture particular types of performance data that would then inform the targeted learning goals or competencies. Since individuals learn in action, assessment should be situated within this learning process.[20] The ECD framework helps operationalize what learners do in complex contexts and how it relates to constructs of interest. Additionally, ECD helps us link specific learner actions within games to constructs, without interrupting what learners are doing or thinking.[21]

In summary, the ECD framework is based on the assumption that assessment is, at its core, an evidentiary argument. Its strength resides in the development of performance-based assessments where what is being assessed is latent or not apparent.[22] The ECD framework begins by determining what we want to assess (i.e., the claims we want to make about learners), and clarifying the intended goals, processes, and outcomes of learning. This information about the student can be used to support learning. That is, it can serve as the basis for (a) delivering timely and targeted feedback to the student/player, as well as (b) presenting a new task or quest that is right at the cusp of the student’s skill level, in line with flow theory[23] and Vygotsky's zone of proximal development.[24] The next section introduces a technology for ECD based-assessment in video games.

Stealth Assessment in Video Games

In addition to the ECD methodology for assessment development, new assessment technologies are needed to capture the vast amounts of data that can come from video gameplay. One technology we developed and are using is called *stealth assessment*. Stealth assessments are performance-based assessments embedded within games to dynamically, unobtrusively, accurately, and transparently measure how players are progressing relative to targeted competencies.[25] Embedding performance-based assessments within games provide a way to monitor a player's current level on valued competencies, and then use that information as the basis for support, such as adjusting the difficulty level of challenges.

How does stealth assessment work? During gameplay, students naturally produce rich sequences of actions while performing complex tasks, drawing on a variety of competencies. Evidence needed to assess the competencies is thus provided by the players' interactions with the game itself (i.e., the processes of play), which can be contrasted with the end product(s) of an activity—the norm in most educational environments. Thus stealth assessment is built within the game (without changing fundamental game mechanics) and statistically aligns the game problems and players' interactions in the game with the underlying competencies of interest.

Currently, we are exploring the validity of stealth assessments in the video game *Newton's Playground*. In this research we are evaluating the degree to which our stealth assessments yield valid and reliable measures of the targeted competencies (e.g., creativity, persistence, and conceptual physics). The next section describes *Newton's Playground* with its stealth assessments to illustrate our performance-based approach.

Newton's Playground

Newton's Playground is a computer game that emphasizes two-dimensional physics simulations, including gravity, mass, potential and kinetic energy, and transfer of momentum. The objective of each problem in the game is to guide a green ball from a predetermined starting point to a red balloon. All movement obeys the rules of physics relating to Newton's three laws of motion. The primary game mechanic is drawing physical objects on the screen that "come to life" once the object is drawn. For example, in Figure 1, the player must draw a golf club on a pin (i.e., little circle on the cloud) to make it swing down to hit the ball. In the depicted solution, the player also drew a ramp to prevent the ball from falling down a pit.

<FIGURE 1 ABOUT HERE>

The speed of (and importantly, the impulse delivered by) the swinging golf club is dependent on the mass distribution of the club and the angle from which it was dropped to swing. The ball will then fly at a certain speed and trajectory. If drawn properly, the ball will hit the balloon. The various problems in Newton's Playground require the player to create and use pendulums, levers, springboards, and so forth to move the ball. All solutions are drawn with color pens using the mouse.

Agents of force and motion

Newton's Playground requires players to create and use the following devices to help the ball reach the balloon:

1. *Ramp*: A ramp can be employed to change the direction of the motion of the ball (or another object). In some cases, other tools (like a pendulum or nudge), are needed to get the ball to start moving.
2. *Lever*: A seesaw or lever involves net torque. A lever rotates around a fixed point usually called a fulcrum or pivot point. An object residing on a lever gains potential energy as it is raised.
3. *Pendulum*: A swinging pendulum directs an impulse tangent to its direction of motion. The idealized pendulum is a specialized case of the physical pendulum for which the mass distribution helps determine the frequency. One

can draw a physical pendulum in Newton's Playground, and the motion will be determined by the mass distribution.

4. *Springboard*: A springboard (or diving board) stores elastic potential energy provided by a falling weight. Elastic potential energy becomes kinetic as the weight is released.
5. *Pin*: A pin allows the position of one body to be fixed in space. Like a nail, it supplies a force large enough to resist motion of the point it is attached to. Two pins hold a body immobile against a background.
6. *Rope*: Ropes generally transmit tension between objects. If a rope is draped over a pulley with masses attached at both ends and the masses are equal, their weights are equal and the net force on each will be the difference between the tension pulling up on the mass and the force of gravity pulling down. Ropes can also acts like trampolines, generating forces on objects by stretching the rope and then removing the force (by deleting objects) to produce upward momentum on the ball.
7. *Nudge*: An arrow in Newton's Playground allows the user to poke/nudge an object into motion.

Newton's Playground is a game that we developed to support stealth assessment of focal competencies (i.e., conceptual physics understanding, conscientiousness, and creativity). This game was inspired by and modeled after a popular physics game called *Crayon Physics Deluxe* by Petri Purho. Newton's Playground has the same basic game mechanics as Crayon Physics Deluxe (e.g., draw objects that serve to move the ball to an end point), and our game uses the same physics engine (i.e., Box2D).

Developing stealth assessments in Newton's Playground required a comprehensive delineation of player actions that would count as varying levels success in the game (i.e., defining the evidence model). It also entailed creating new problems in the game in line with the task model to meet specific needs of the assessments. We now illustrate how we are currently implementing stealth assessments for two competencies (conceptual physics understanding and persistence) in the following sections.

Conceptual Physics

Over the past several decades it has become very clear that many students, who have gotten acceptable grades in one or more physics courses, actually have very limited practical understanding of the physics involved. Numerous studies have shown that a passing grade does not mean that a student has an appreciation of physical principles.[26] For instance, Halloun and Hestenes found that only 15% of their 478 college physics students showed an accurate understanding of the relationship between unbalanced forces and acceleration (i.e., Newton's 2nd law: $F = ma$).[27] This has led to widespread adoption of the text Conceptual Physics by Paul Hewitt,[28] and the development of two instruments, the Force Concept Inventory[29] and the Mechanics Baseline Test[30] now widely used to compare student mastery of the concepts of mechanics. Recognition of the problem has also led to a renewed interest in the mechanisms by which physics students make the transition from naive or folk physics to Newtonian physics[31] and to the possibility of video game playing assisting in the process.[32]

Physics engines are becoming pervasive in gaming environments, providing a sense of realism in a game (e.g., Havok engine). Within these gaming environments, players can experiment with principles of physics such as impulse, inertia, vector addition, elastic collision, gravity, velocity, acceleration, free-fall, mass, force, and projectile motion. The degree that players apply these principles correctly in Newton's Playground comprises evidence for conceptual understanding of physics.

Based on Hewitt's excellent textbook on foundational conceptual physics,[33] we interpret competency in conceptual physics to involve the following:

1. *Conceptual understanding of Newton's three laws of motion.* Newton's three laws of motion provide a conceptual understanding of how objects interact in the environment. The first law tells us that an object in rest stays in rest in the absence of any forces, and an object in motion stays in motion in the absence of any forces. The second law (F = ma) tells us how the motion of the particle (object) evolves when it experiences a nonzero net force. Here F is the net force applied (i.e., the vector sum of all the forces acting on the object), m is the mass of the object, and a is the object's acceleration. Thus, the net force applied to an object produces a proportional acceleration. That is, if an object is accelerating, then there is a net nonzero force on it. Any mass that is gained or lost by the system will cause a change in momentum that is not the result of an external force. In simple terms, it takes less force to accelerate an object that has less mass compared to one with more mass. The third law states for every action there is an equal and opposite reaction. This is commonly described by hitting a tree with a baseball bat. The force exerted on the tree by the swinging bat is equal to the force exerted back on the person swinging the bat.
2. *Conceptual understanding of Potential and Kinetic energy*. Potential energy exists when a force acts on an object to restore the object to its resting point (or "lower energy configuration"). For example, when a springboard (like in Newton's

Playground) is bent downward, it exerts an upward force to return to its un-bent position. The action of bending the springboard down requires energy, and the force acting on the springboard to return it to its resting point is potential energy. When the bent springboard is released, the stored energy will be converted into kinetic energy.

3. *Conceptual understanding of conservation of angular momentum or torque*. The angular momentum of a system of objects about any point of reference can be computed from the position and momentum of each of the objects. A useful mental image is that of a figure skater or gymnast. Figure skaters will begin an elegant spin with arms outstretched. Once they start spinning, they typically draw their hands inward so that they can spin more rapidly. The sum of the mass of each object making up the skater times the square of the (perpendicular) distance to the axis of rotation is called the skater's *moment of inertia*. For a rotating object, the angular momentum is the product of the moment of inertia and the angular velocity. With negligible friction between skater and ice, decreasing the moment of inertia by moving the arms inward increases the rotational velocity. Similar considerations apply to a gymnast doing somersaults while dismounting, or a diver on the way down to the water. A torque with a short moment arm can counterbalance the torque exerted by a much smaller force with a larger moment arm and vice versa. Consider an ordinary lever. The force of support at the fulcrum is not directly given, but the relation between torque and angular acceleration can easily be exploited by measuring torques from the fulcrum.

Newton's three laws is a parent principle that is pervasive in almost all problems in Newton's Playground. Successful use of each agent of force and motion is an indicator for a particular physics principle. Additionally, there are micro-indicators that inform each agent and principle as well. Table 1 displays our current set of micro-indicators for each conceptual physics principle.

<TABLE 1 ABOUT HERE>

All of the problems that are used in our stealth assessments require the player to create and use one or more agents of force and motion in the solution. A successful solution thus informs one or more of the competencies that we hope to develop in the student. As an illustration, consider the problem called *ballistic pendulum*, shown in Figure 2.

<FIGURE 2 ABOUT HERE>

This problem was inspired by an actual experiment often conducted in introductory physics courses in college to teach physics concepts. It requires the student to create a pendulum positioned so it hits the ball into a trajectory that ultimately hits the balloon. Successfully solving this problem suggests that the student has an intuitive concept of torque, linear, and angular momentum.

Consider another problem shown in Figure 3. This problem specifically requires the player to use a springboard solution. The springboard is a variant of the lever in which one flat board rests on an object that is pinned in place, but hangs over one edge. As shown in Figure 3, a weight is affixed onto the free end of the springboard. The edge acts as an instantaneous axis of rotation and the board experiences an angular acceleration which can be used to launch objects up into space. This requires knowledge of potential and kinetic energy, and conservation of angular momentum. As Table 2 indicates, there are a number of micro-indicators that we look for in a springboard problem. For example, understanding potential energy entails maximizing the “spring” of the springboard. This requires adding sufficient weight to exert a downward force on the springboard and can be accomplished either by dropping an object onto the springboard or attaching an object to the springboard and then deleting the object once the springboard has bent down to a sufficient angle. In the latter case, knowing when to delete the object from the springboard requires knowledge of angular momentum in order to maximize the upward force on the ball.

<FIGURE 3 ABOUT HERE>

Conscientiousness and Persistence

Persistence is the disposition to try hard, particularly in the face of failure.[34] It is one of the main facets of conscientiousness, which has emerged as an important competency in predicting academic performance[35] as well as positive life outcomes.[36] Conscientiousness is a multi-faceted competency that commonly includes tendencies related to being attentive, hard-working, careful, detail-minded, reliable, organized, productive, and persistent.[37]

A number of studies and meta-analyses have shown the importance of self-report measures of conscientiousness in predicting a variety of important outcomes while controlling for cognitive ability. Conscientiousness has consistently been found to predict academic achievement from preschool to high school, to the postsecondary level and adulthood.[38] Meta-analyses have linked conscientiousness with grades between r = .21 and .27, and the relationship is independent of intelligence.[39]

Assessing persistence in a game such as Newton's Playground is primarily based on seeing how long players spend trying to solve difficult problems. The challenge in this assessment design is that different ability levels can preclude a player from solving a problem. To address this issue, we created difficulty indices for all of the Newton's Playground problems. This allows us to incrementally increase the difficulty of problems to ensure that students will eventually get to problems they will have trouble solving. Difficulty indices include the following:

1. *Relative location of ball to balloon*. If the balloon is positioned above the ball in a problem, this forces the player to use a lever, pulley, springboard, or pendulum to solve the problem (0-1 point).
2. *Obstacles*. This refers to the pathway between the ball and balloon. If the pathway is obstructed, this requires the player to project the ball in a very specific trajectory to hit the balloon (0-2 points).
3. *Distinct agents of force/motion* (see previous section on Agents of Force and Motion). A problem may require one or two agents to get the ball to the balloon (0-1 point).
4. *Novelty*. This addresses whether a problem is novel relative to other problems played. Problem solution is not easily determined from experience with other problems (0-2 points).

Each problem was evaluated under all of the rubrics to yield a total difficulty score (i.e., ranging from 0-6). Consider a problem called *Cave Story* shown in Figure 4. This problem gets a difficulty score of 5 as the star is above the ball (1), there's one obstacle which is a narrow pathway (1), two agents are typically needed to solve it (1) and there's no other problem like that in the game (2). Thus the problem in Figure 4 would be a good problem to assess persistence as it will likely be unsolvable by many students.

<FIGURE 4 ABOUT HERE>

Conclusion

As described in this chapter, games often require a player to apply various competencies to succeed. Many of these competencies are considered to be valuable for college graduates to succeed in 21st century workforce.[40] For instance, in a strategy-based game like *Starcraft*, players must engage in causal reasoning and systems thinking. That is, they need to consider the ramifications of their actions in the game—not only on aspects of building their own interstellar galaxy, but also in relation to other players' interstellar galaxies (e.g., how to strategically collect resources that will grow faster than those of competing players). In addition, many games require divergent thinking to solve hard problems (e.g., *Crayon Physics Deluxe*), encourage players to thoroughly explore a space before moving on (e.g., *L.A. Noire*), and call for players to work in teams to solve complex problems (e.g., *World of Warcraft*).

We encourage researchers and educators to look at how video games can be used to assess key skills needed in the 21st century. These skills are important in today's society and video games have the power to assess and to help improve these skills through providing an engaging medium to practice the skills over extended periods of time.

Good game design coupled with a robust assessment approach should be the starting point for any research project focused on building a video game for educational purposes. That is, such research should combine game design with assessment methodologies such as ECD at the outset of the game design process, rather than considering assessment as an afterthought. These assessments should be grounded in theory, and should start with defining what competencies are important and how a video game can be used to assess and improve these competencies. Finally, more attention should be given to figuring out specifically how video games can help improve important new competencies. Since good video games hold such an engagement value, they are useful (and fun) tools for players to practice skills over extended amounts of time, especially for today's college students who grew up playing such games.

Notes

1. Tracy Fullerton, *Game Design Workshop: A Playcentric Approach to Creating Innovative Games* (Burlington, MA: Elsevier Inc, 2008); *Essential Facts About the Computer and Video Game Industry* (Entertainment Software Association, 2011).

2. Soledad L Escobar-Chaves and Craig A. Anderson, "Media and Risky Behaviors," *The Future of Children* 1, no. 18 (2008): 147-180.

3. Steve Jones, *Let the Games Begin: Gaming Technology and Entertainment Among College Students* (Pew Research Center, 2003).

4. Shawn C. Green and Daphne Bavlier, "Action-video-game Experience Alters the Spatial Resolution of Vision," *Psychological Science* 1, no. 18 (2007): 88-94; Rebecca Shaw, Andy Grayson and Vicky Lewis, "Inhibition, ADHD, and Computer Games: The Inhibitory Performance of Children with ADHD on Computerized Tasks and Games," *Journal of Attention Disorders* 4, no. 8 (2005): 160-168; Rebecca M Chory and Alan K. Goodboy, "Is Basic Personality Related to Violent and Non-violent Video Game Play and Preferences?" *Cyberpsychology, Behavior and Social Networking* 4, no. 14 (2011): 191-198; Matthew Ventura, Valerie J. Shute and Yoon J. Kim, "Video Gameplay, Personality and Academic Performance," *Computers & Education*, no. 58 (2012): 1260-1266; Edward A. Witt, Adam J. Massman and Linda A. Jackson, "Trends in Youth's Videogame Playing, Overall Computer Use, and Communication Technology Use: The Impact of Self-Esteem and the Big Five Personality Factors," *Computers in Human Behavior* 2, no. 27 (2011): 763-769; Marko M. Skoric, Linda L. Teo and Rachel L. Neo, "Children and Video Games: Addiction, Engagement, and Scholastic Achievement," *CyberPsychology & Behavior* 5, no. 12 (2009): 567-572; Christopher J. Ferguson and

Adolfo Garza, "Call of (Civic) Duty: Action Games and Civic Behaviors in a Large Sample of Youth," *Computers in Human Behavior*, no. 27 (2011): 770-775.

5. Sasha Barab and others. "Pedagogical Dramas and Transformational Play: Narratively Rich Games for Learning," *Mind, Culture, and Activity* 3, no. 17 (2010): 235-264; Renee E. DeRouin-Jessen, "Game on: The Impact of Game Features in Computer-based Training" (dissertation, University of Central Florida, 2008).

6. Ferguson and Garza 2011, 770-775.

7 . Geoff P. Smith, *Growing Up with Tok Pisin: Contact, Creolization, and Change in Papua New Guinea's National Language* (London: Battlebridge Publications, 2002).

8. Jean Lave and Etienne Wenger, *Situated Learning: Legitimate Peripheral Participation* (Cambridge: University of Cambridge Press, 1991); Lev Vygotsky, *Mind in Society: The Development of Higher Psychological Processes* (Cambridge: Harvard University Press, 1978).

9. William L. Bryan and Noble Harter, "Studies on the Telegraphic Language: The Acquisition of a Hierarchy of Habits." *Psychological Review* 4, no. 6 (1899): 345-375; Allen Newell and Paul S. Rosenbloom, "Mechanism of Skill Acquisition and the Law of Practice," In *Cognitive Skills and Their Acquisition*, edited by Jone R. Anderson, 1-51. (Hillsdale: Lawrence Erlbaum Associates, Inc., 1981); Walter Schneider and Richard M. Shiffrin, "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention." *Psychological Review* 1, no. 84 (1977): 1-66; Richard M. Shiffrin and Walter Schneider, "Controlled and Automatic Human Information

Processing: II. perceptual Learning, Automatic Attending and a General Theory," *Psychological Review* 2, no. 84 (1977): 127-190; Valerie J. Shute, Lisa A. Gawlick and Kevin A. Gluck, "Effects of Practice and Learner Control on Short- and Long-term Gain and Efficiency," *Human Factors* 2, no. 40 (1998): 296-310; Edward L. Thorndike, "Animal Intelligence: An Experimental Study of the Associative Processes in Animals," *Psychological Review* (Macmillan), no. 8 (1898); Jennifer Underdahl, Joycelyn Palacio-Cayetano and Ron Stevens, "Practice Makes Perfect: Assessing and Enhancing Knowledge and Problem-solving Skills with IMMEX Software," *Learning and Leading with Technology* 7, no. 28 (2001): 26-31.

10. George Polya, *How to Solve it* (Princeton, NJ: Princeton University Press, 1945).

11. Robert D. Tennyson and Jorczak L. Robert, "A Conceptual Framework for the Empirical Study of Instructional Games." In *Computer Games and Team and Individual Learning*, edited by H. F. O'Neil and R. S. Perez, 3-20 (Amsterdam: Elsevier, 2008).

12. Fullerton, 2008; Valerie J. Shute, Lloyd Rieber, and Richard Van Eck. “"Games...and...Learning." In *Trends and Issues in Instructional Design and Technology*, edited by R. Reiser and J. Dempsey, 321-332. (Upper Saddle River, NJ: Pearson Education Inc., 2011).

13. James P Gee, *Good Video Games + Good Learning: Collected Essays on Video Games, Learning and Literacy* (New York: Peter Lang, 2007).

14. Jane McGonigal, *Reality is Broken: Why Games Make us Better and How They Can Change the World* (New York City, NY: Penguin Press, 2011).

15. Vygotsky, 1978.

16. James P. Gee, Glynda Hull, and Colin Landshear, *The New Work Order: Behind the Language of the New Captitalism* (Boulder, CO: Westview, 1996).

17. Valerie J. Shute, "Stealth Assessment in Computer-based Games to Support Learning," In *Computer Games and Instruction*, edited by S. Tobias and J. D. Fletcher, 503-524 (Charlotte, NC: Information Age Publishers, 2011); Valerie J. Shute, Matthew Ventura, Malcolm I. Bauer and Diego Zapata-Rivera, "Melding the Power of Serious Games and Embedded Assessment to Monitor and Foster Learning: Flow and Grow," In *Serious Games: Mechanisms and Effects*, edited by U. Ritterfeld, M. Cody and P. Vorderer, 295-321 (Mahwah, NJ: Taylor and Francis, 2009).

18. Robert J. Mislevy, Russel G. Almond and Janice F. Lukas, *A Brief Introduction to Evidence-Centered Design* (Los Angeles: Center for Research on Evaluation, Standards, and Student Testing, 2004); Robert J. Mislevy, Linda S. Steinberg and Russel J. Almond, *On the Structure of Educational Assessment* (Los Angeles: Center for Research on Evaluation, Standards, and Student Testing, 2003).

19. Russel G. Almond and Robert J. Mislevy, "Graphical Models and Computerized Adaptive Testing," *Applied Psychological Measurement*, no. 23 (1999): 223-237.

20. James P. Gee, *What Video Games Have to Teach us About Learning and Literacy* (New York: Palgrave/Macmillan, 2003); Katie Salen and Eric Zimmerman, "Game Design and Meaningful Play," In *Handbook of Computer Game Studies*, edited by Joost Raessens and Jeffrey Goldstein, 59-80 (Cambridge, MA: MIT Press, 2005).

21. Shute, 2011.

22. André A. Rupp, Matthew Gushta, Robert J. Mislevy and David W. Shaffer, "Evidence-centered Design of Epistemic Games: Measurement Principles for Complex Learning Environment," *Journal of Technology, Learning and Assessment* 4, no. 8 (2010).

23. Mihály Csikszentmihalyi, "The Domain of Creativity," In *Theories of Creativity*, edited by Mark A. Runco and Robert S Albert, 190-212 (Newbury Park, CA: Sage, 1990).

24. Vygotsky, 1978.

25. Shute, 2011; Shute, Ventura, Bauer, & Zapata-Rivera, 2009.

26. Catherine H. Crouch, and Eric Mazur, "Peer Instruction: Ten years of Experience and Results," *American Journal of Physics*, no. 69 (2001): 970-977; Richard R. Hake, "Interactive-engagement Versus Traditional Methods: A Six-thousand-student Survey of Mecahnics Test Data for Introductory Physics Courses," *American Journal of Physics* 1, no. 66 (1998): 64-74; Ibrahim Halloun, and David Hestenes, "The Initial Knowledge State of College Physics Students," *American Journal of Physics*, no. 53 (1985): 1043-1055.

27. Halloun and Hestenes, 1985.

28. Paul Hewitt, *Conceptual Physics.* 11th ed. (San Francisco: Addision Wesley, 2009).

29. David Hestenes, Malcolm Wells and Gregg Swackhamer, "Force Concept Inventory," *The Physics Teacher*, no. 30 (1992): 141-158.

30. David Hestenes and Malcolm Wells, "A Mechanics Baseline Test," *The Physics Teacher*, no. 30 (1992): 159-166.

31. Andrea A. DiSessa, "Unlearning Aristotelian Physics: A Study of Knowledge-based Learning," *Cognitive Science* 1, no. 6 (1982): 37-75.

32. Barbara Y. White, “Designing Computer Games to Help Physics Students Understand Newton's Laws of Motion,” *Cognition and Instruction* 1, no.1 (1994): 69-108.

33. Richard P. Feynman. *The Character of Physical Law* (Cornell University Press, 1964); Richard P. Feynman, Robert B. Leighton and Matthew Sands, *The Feynman Lectures in Physics (*Addison Wesley, 1964).

34. Valerie J. Shute, Matthew Ventura and Diego Zapata-Rivera*, Stealth assessment in digital games* (The MacArthur Foundation, in press).

35. Arthur E. Poropat, "A Meta-analysis of the Five-factor Model of Personality and Academic Performance," *Psychological Bulletin*, no. 135 (2009): 322-338.

36. Brent W. Roberts, Nathan R. Kuncel, Rebecca L. Shiner, Avshalom Caspi and Lewis R. Goldberg, "The Power of Personality: the Comparative Validity of Personality Traits, Socioeconomics Status, and Cognitive Ability for Predicting Important Life Events," *Perspectives on Psychological Science* 4, no. 2 (2007): 313-345.

37. Brent W. Roberts, Oleksandr S. Chernyshenko, Stephen Stark and Lewis R. Goldberg, "The Structure of Conscientiousness: An Empirical Investigation Based on Seven Major Personality Questionnaires," *Personnel Psychology*, no. 58 (2005): 103-139.

38. Jo A. Abe, "The Predictive Validity of the Five-factor Model of Personality with Preschool Age Children: A nine year Follow-up Study," *Journal of Research in Personality* 4, no. 39 (2005): 423-442; Erik E. Noftle and Richard W. Robins,

"Personality Predictors of Academic Outcomes: Big Five Correlates of GPA and SAT Scores," *Journal of Personality and Social Psychology* 1, no. 93 (2007): 116-130; Poropat, 2009; Melissa O'Connor and Sampo V. Paunonen, "Big Five Personality Predictors of Post-secondary Academic Performance," *Personality and Individual Differences*, no. 43 (2007): 971-990; Sabrina Trapmann, Benedikt Hell, Jan-Oliver W. Hirn and Heinz Schuler, "Meta-analysis of the Relationship Between the Big Five and Academic Success at University," *Journal of Psychology*, no. 215 (2007): 132-151; Filip De Fruyt and Ivan Mervielde, "Personality and Interests as Predictors of Educational Streaming and Achievement," *European Journal of Personality*, no. 10 (1996): 405-425; Rebecca L. Shiner, Ann S. Masten and Jennifer M. Roberts, "Childhood Personality Foreshadows Adult Personality and Life Outcomes Two Decades Later," *Journal of Personality* 6, no. 71 (2003): 1145-1170.

39. Noftle & Robins, 2007; Proporat, 2009.

40. Jill Casner-Lotto and Linda Barrington, *Are They Really Ready to Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century Work Force* (New York: The Conference Board, 2006).

FIGURES AND TABLE

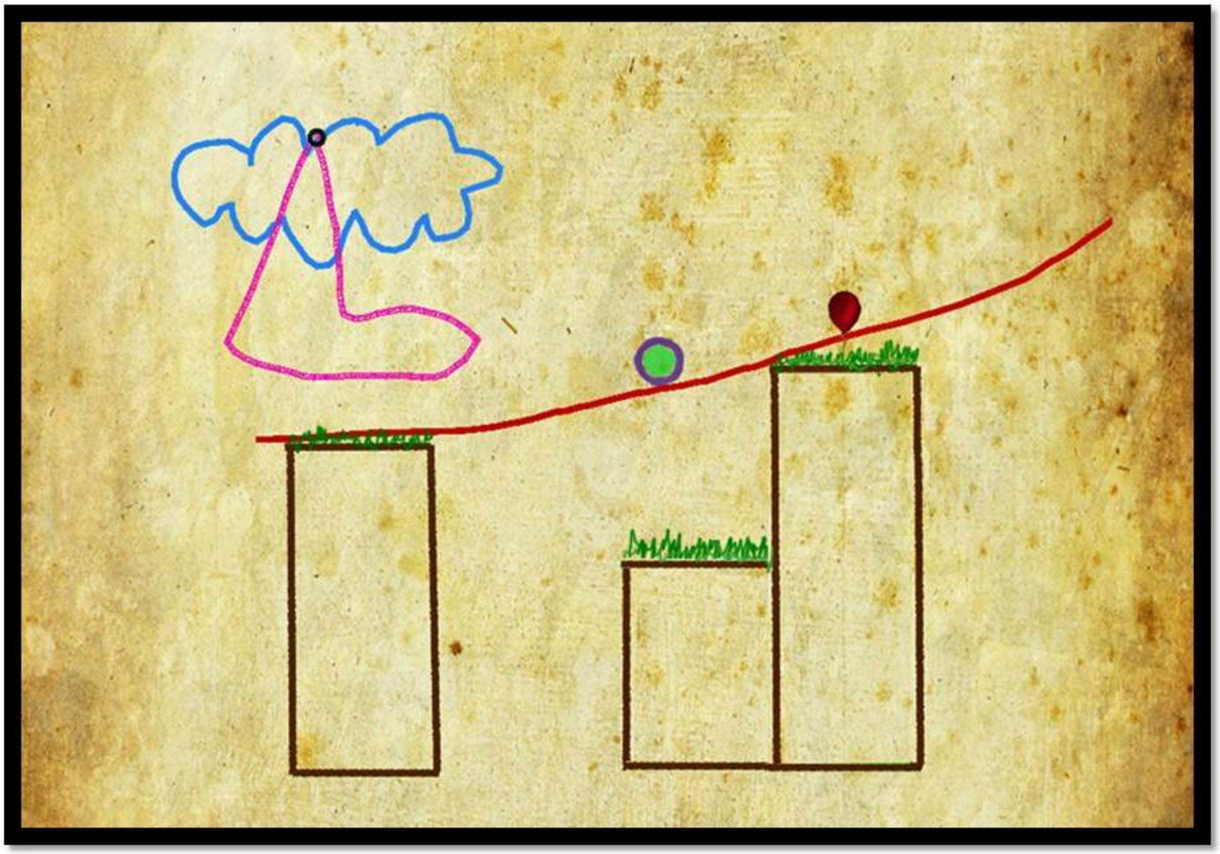

Figure 1: Golf problem in Newton's Playground

Agents	Micro-Indicators
Ramp	1. Number of bends (or tubes, i.e., tortuosity) 2. Angle of each bend 3. Length of ramp
Lever	1. Length of the lever 2. Position of fulcrum 3. Height through which object falls before hitting lever 4. Mass of object 5. Location of the dropped object on lever (distance from fulcrum)
Pendulum	1. Angle of pendulum relative to horizontal fulcrum (180 degrees is max) 2. Length between the axis point and the fulcrum (Moment of inertia) 3. Mass (important when the pendulum hits something) 4. Position of pin
Springboard	1. Length of springboard 2. Mass of the object to weight it down 3. Position of the ball at release 4. Delete object or let fall off springboard 5. Angle of springboard at release (90 degrees max)

Table 1. Micro-indicators for CPD agents of force and motion

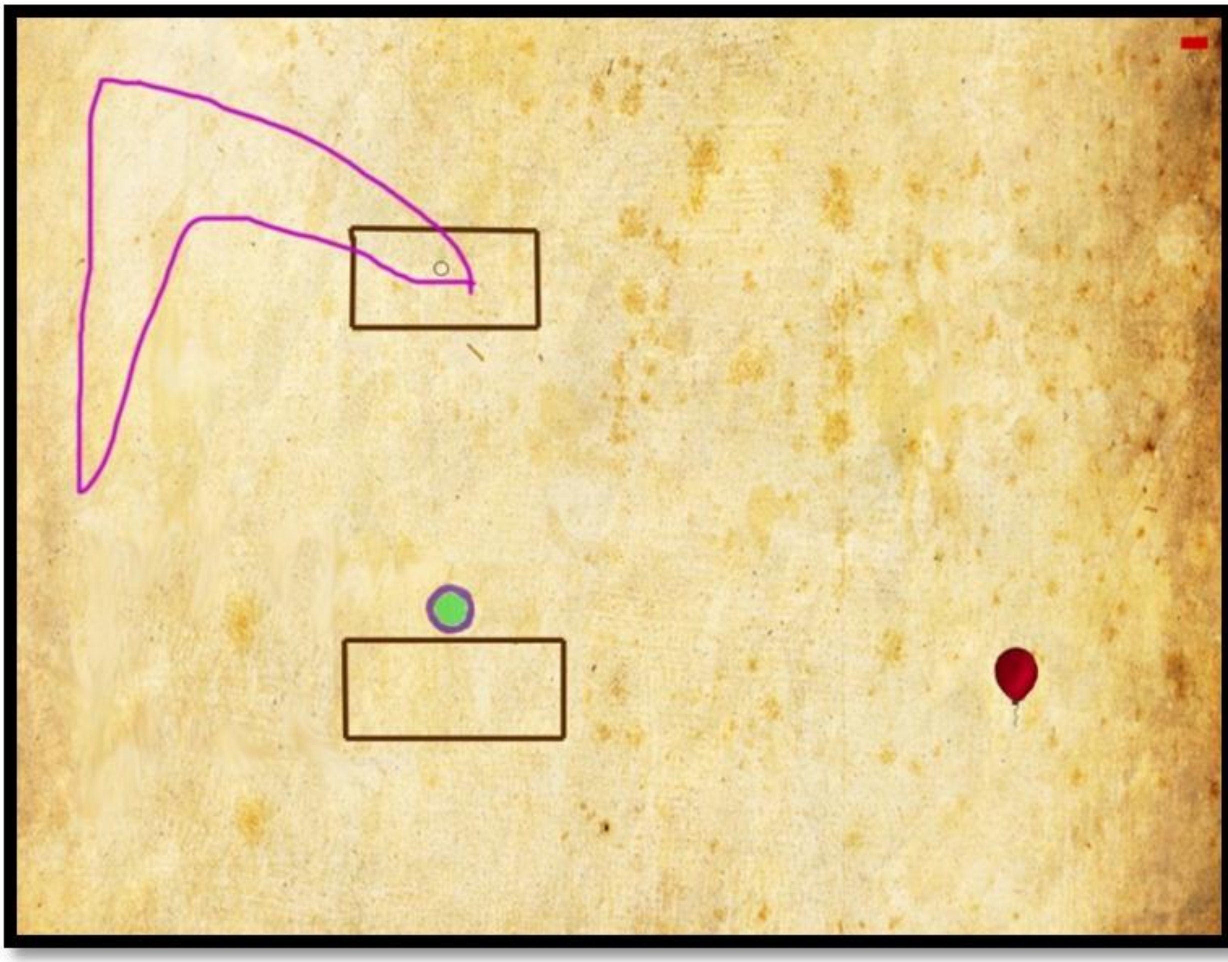

Figure 2. Ballistic pendulum problem in Newton's Playground

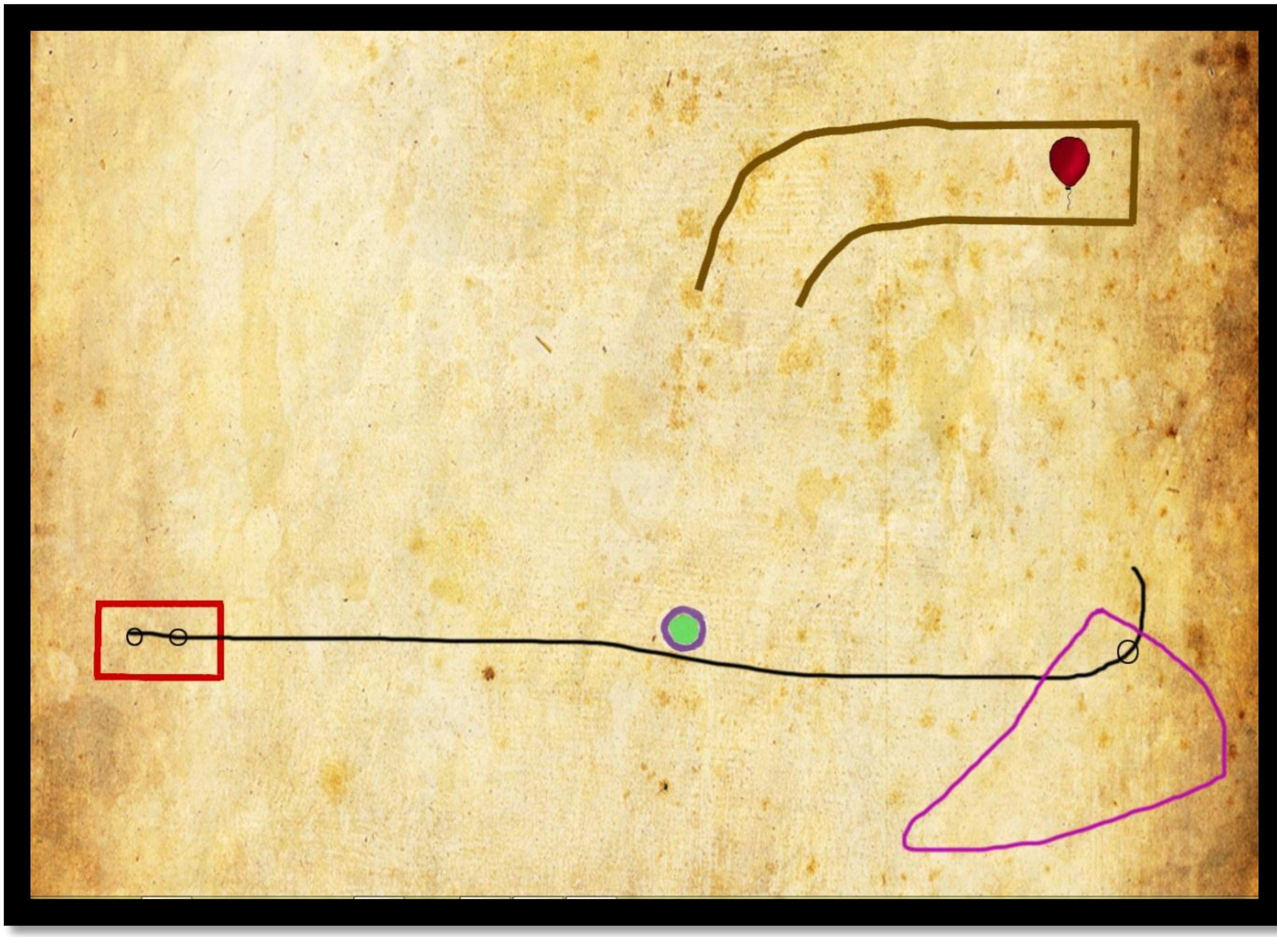

Figure 3. Diving board problem in Newton's Playground

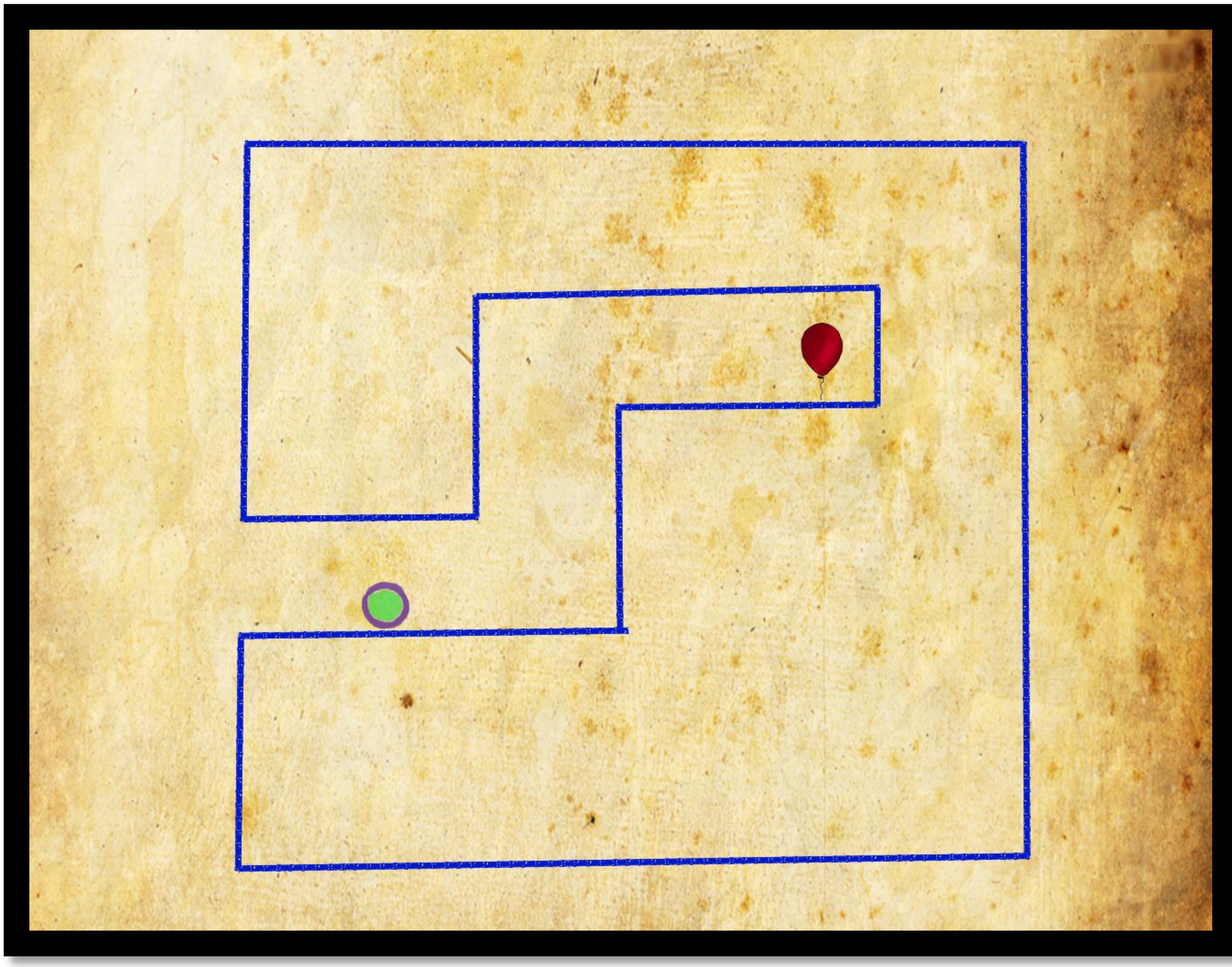

Figure 4: Cave story in Newton's Playground