ARTICLE

The Design, Development, and Testing of Learning Supports for the Physics Playground Game

Valerie J. Shute[1] · Ginny Smith[1] · Renata Kuba[1] · Chih-Pu Dai[1] · Seyedahmad Rahimi[1] · Zhichun Liu[1] · Russell Almond[1]

Abstract

In honor of Jim Greer, we share our recent work—a design and development study of various learning supports embedded within the game *Physics Playground*. This 2-dimensional computer game is designed to help students learn Newtonian physics and uses stealth assessment to measure, in real-time, their physics understanding. The game operates according to Newtonian mechanics where students draw simple machines (e.g., ramps, levers, pendulums, and springboards), or adjust sliders to manipulate physics parameters to solve problems. Over the past three years, a team of experts in measurement, assessment, learning sciences, game design, and physics education has iteratively examined how to enhance students' learning during gameplay. We incorporated supports for both physics understanding and game mechanics. Some of our designs showed promise while others did not and were discarded. Through three usability studies and a final learning study, we document the design, development, and testing of the final set of supports that involve modeling (Worked Examples, Physics Videos), modality (Animations, Formulas, Definitions, Hewitt Videos), and advice (Hints). We conclude with a discussion of the impact of incorporating learning supports, students' perceptions of the learning supports, and the lessons learned through our iterative process. Jim shared our passion for supporting student learning and we hope that by sharing our journey we spur others on to continue the work he loved.

Keywords Game-based learning · Stealth assessment · Learning supports · Physics education · Physics playground

✉ Valerie J. Shute
vshute@fsu.edu

[1] Florida State University, Tallahassee, FL, USA

Introduction

Going back to the mid-1980s, Jim Greer and I (first author) have engaged in parallel and often overlapping research streams in the field of artificial intelligence and education. To illustrate, here are some of the areas in which we both did extensive research and published papers (typically citing each other's work in the process):

(a)   Intelligent tutoring systems (e.g., design and evaluation)
(b)   Student modeling (e.g., Bayesian networks for student models, visualizing Bayesian networks, externalizing student models, and granularity issues in diagnosis)
(c)   Collaboration (e.g., in learning and problem solving)
(d)   Modeling scientific reasoning/inquiry (within our own respective intelligent tutoring systems)
(e)   Adaptivity and personalization in instructional/learning systems (including learning analytics, feedback, and other types of support)
(f)   Best practices and lessons gleaned from the aforementioned areas.

The main theme that we both deeply cared about boils down to this: *How can we accurately measure student learning (in real time, at various grain sizes, and as transparently as possible) of targeted knowledge and skills, and then use that information to provide personalized support for further development?* I will sorely miss Jim's creative and thoughtful contributions in the field moving forward.

In addition to simply being aware of each other's work, we also liked each other as people. When searching through some pictures, I found two pictures (both taken at AIED conferences; see Fig. 1) where we hung out, along with Julita Vassileva.

This paper is intended as a tribute to an amazing scholar and friend—Jim Greer. Our goal is to share some of our current work in the area of game-based assessment for learning—warts and all. Our hope is that if Jim were here, he would approve.

Purpose

In 2016, the National Science Foundation (NSF) awarded us a three-year grant for improving STEM education through game-based assessment. The project used a learning game, *Physics Playground* (Shute et al. 2019a, 2019b, 2019c, 2019d),

Fig. 1 Photograph of three friends (Greer, Vassileva, and Shute) across time

embodying Newtonian physics (e.g., Newton's laws of force and motion, torque, linear momentum, and energy). One main component of the NSF project was the design and development of in-game learning supports. Over the three-year project timeline, we used an iterative design process to create and test the supports. This paper shares our design and development journey. We conclude with some preliminary results from a recent study testing the impact of our learning supports on physics understanding.

Background

Physics Playground (PP) is a 2-dimensional computer game assessing students' Newtonian physics understanding using stealth assessment (Shute 2011). The game is appropriate for learners from 7th grade into adulthood. The overarching goal of the game and all levels is to help students understand Newtonian physics—relative to particular physics competencies, which are shown in Fig. 2. We define "competency" broadly—as understanding relevant knowledge and rules related to targeted physics content (e.g., understanding that *Energy can transfer*). Each competency has a set of associated terms (e.g., the term "elastic potential energy" is linked to the physics competency "energy can transfer").

The goal across all the hundreds of game levels (i.e., individual puzzles) we have created is always the same—move the ball so that it hits the balloon. Students solve a level by either drawing (sketching levels) or manipulating variables (manipulation levels). That is, sketching levels are solved by drawing ramps, levers, pendulums, and/or springboards on the screen using a mouse or stylus. Manipulation levels are solved by moving sliders of various physics parameters, such as changing the ball's mass, the level's air resistance, gravity, and/or bounciness of the ball. Students may also
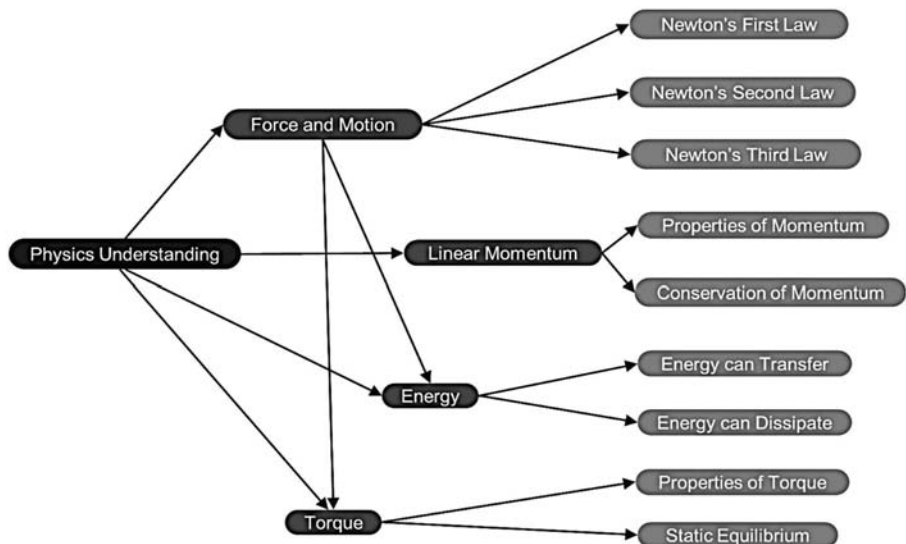

Fig. 2 Competency model for revised version of *Physics Playground*

change the amount of force exerted on the ball by using a puffer and/or adjusting a blower in certain manipulation levels.

All game levels were designed using Evidence Centered Design (ECD) (Mislevy et al. 2003) to ensure alignment between game tasks (i.e., game levels) and assessment of the targeted physics competencies. The levels were physically created using a level editor we built to accompany the game.

The original version of the game (developed prior to this NSF project), assessed only a few physics competencies (e.g., angular momentum, potential and kinetic energy). During the current NSF project, we expanded the game to assess more physics competencies—nine competencies (on the right) nested within four main physics topics (see Fig. 2).

The structure of the Bayesian network in Fig. 2 served as the initial student model for *Physics Playground* and stems from Jim Greer's seminal work using such networks for student modeling (e.g., Zapata-Rivera and Greer 2004). Following the scoring design process and scoring algorithm in Almond, Mislevy, Steinberg, Yan & Williamson (Almond et al. 2015), we built a collection of evidence models for each game level. The evidence models identified both the competencies addressed in the game levels and the key observables that could provide evidence about those competencies. For example, in a sketching level, if a student draws a lever to solve the level, this provides different evidence than if the student draws a pendulum. Linking students' gameplay to their understanding of the physics competencies, the stealth assessment methodology (Shute 2011) allows for adaptivity of level selection and/or appropriate learning support. Students' current understanding of the nine competencies is updated after completing each game level. An adaptive algorithm is used to select the next game level relative to the student's need. The adaptive algorithm also determines if the student needs to see a targeted learning support before receiving the next level. For more details on adaptivity in *Physics Playground*, see Shute et al. (2019a, 2019b, 2019c, 2019d) and Shute et al. (2020).

To represent the full set of links between all game levels and their associated competencies (as well as their terms and applicable learning supports), we created a large Q-matrix which allowed us to map each of the nine competencies (columns in the Q-matrix) with its corresponding game level (rows in the Q-matrix) in *PP*. Our two physics experts identified a primary (shown as "1" in Table 1) and occasionally a secondary (shown as "2" in Table 1) competency for each game level. All game levels were designed to ensure complete coverage of the competency model.

Once we finalized the expanded physics competency model, the learning support design and development process began. The learning supports for each level were keyed to the

Table 1. Q-matrix example in PP with three game levels as examples

Levels	N 1st L	N 2nd L	N 3rd L	ECT	ECD	POM	COM	POT	Equil
Down Hill	1	0	0	2	0	0	0	0	0
Scale	0	0	0	0	0	0	0	1	0
Twin Peaks	0	1	0	0	0	0	0	0	0

Note. ECT = energy can transfer; ECD = energy can dissipate; POM = properties of momentum; COM = conservation of momentum; POT = properties of torque; Equil = equilibrium

competencies associated with that level in the evidence model. According to Roll and Wylie (2016), students need support in interactive learning environments. Although we found in previous studies that *Physics Playground* can improve learning (e.g., Shute et al. 2013), we wanted to enhance learning even more, particularly for struggling students.

Incorporating learning supports into game play—especially without disrupting the flow state (Csikszentmihalyi 1990) —is tricky, but has been shown to increase students' learning and engagement in learning games. For example, in a recent meta-analysis on the use of learning supports in such games, Wouters and Van Oostendorp (2013) found a significant positive effect of learning supports. Students who played learning games with supports typically outperformed students who played games without supports. They also identified ten distinct types of learning supports. For the NSF project described herein, we focused on three types: modeling, modality, and advice. Modeling supports provide example solutions or explanations about solutions. Modality refers to offering multiple representations, such as auditory and visual presentation of information. Advice supports aim to focus the students' attention on an important aspect or aspects of the game level or activity.

Method

Data Sources

The current study is a design and development study of the learning supports for *Physics Playground* in the past three years. We collected data on the design and development processes of the PP learning supports through two main sources: (a) content analysis of detailed notes from our research team meetings and related documents, and (b) usability testing summaries and reports. The research team was comprised of experts in measurement, assessment, learning game design, and physics education. Full research team meetings occurred biweekly throughout the three-year project with meeting notes recorded during the meetings. Throughout the project, subsets of the research team met for various purposes. The learning support (LS) design team was one subset, and their specific meeting notes and other artifacts (e.g., paper prototypes) were also included in the analysis.

Three iterations of usability testing occurred, each with a different focus. The first usability test examined students' satisfaction with the initial version of learning supports through observations, think-aloud protocols, and student interviews. In the second and third usability studies, we administered a post-game survey to students to assess their satisfaction with the revised learning supports. The second usability study also compared students who used learning supports to students who did not use learning supports, providing preliminary evidence on the impact of the supports. Following the three usability studies, we conducted a quasi-experimental study on the final set of learning supports.

Results

Over the course of the three-year project, we iteratively designed and developed a set of learning supports for our game *Physics Playground*. Each phase of design and

development included a usability test to gather feedback on the students' perceptions and the effectiveness of the various learning supports. The team made design decisions based on the results of these studies, learning/instructional theories, and the recommendations of our physics experts. Below is a description of our iterative design and development process.

Initial Learning Support Design

The first set of learning supports we developed were short expert solution videos ("Worked Examples"), a dictionary containing definitions and examples of key physics terms ("Physics Facts"), short cartoon videos originally created by physics educator Paul Hewitt ("Hewitt Videos"), and short text messages that hinted at the physics needed to solve a level ("Advice").

Worked Examples Video demonstrations were one of the key types of learning supports we planned to develop for PP. We discussed the possibility of showing students how an expert would solve the game levels, (i.e., modeling). We called these supports Worked Examples (WEs). The purpose of WEs is to model and illustrate how a problem can be solved (Lang and O'Neil 2008; Shute et al. 2019a, 2019b, 2019c, 2019d; Wouters and Van Oostendorp 2013). Members of the research team recorded expert solution videos for each game level. When a game level had multiple solutions, we created WEs for each one, and students could choose which solution method to view. The physics experts provided a script, and narration was added to highlight the physics rationale per solution.

Physics Facts A second major focus for game support development included physics explanations. One of our first ideas was to create a dictionary explaining the key physics terms relevant to the game. We created the list of 28 terms based on the physics experts' recommendations of the physics covered through gameplay. We dubbed the support "Physics Facts" as each entry (i.e., term) included a definition, an example, and a "see also" section listing related terms. We envisioned Physics Facts to be a searchable or scrollable webpage with hyperlinks to the related terms. However, the first iteration was a scrollable document of static text.

Hewitt Videos Hewitt videos are cartoon animations explaining general physics competencies such as Newton's first, second, and third laws of force and motion. Paul Hewitt created the videos, and with his permission, we selected and edited several videos for each competency to incorporate into our first set of learning supports. We selected the videos according to engagement and relevance (i.e., direct association to our nine competencies).

Advice The advice support was the first set of level specific hints we developed. The idea of Advice was to help students perceive the physics behind each level so they would choose the right solution method. Our physics experts wrote 2–3 pieces of advice (i.e., short sentences explaining the physics competency underlying a specific game level; see Fig. 3) to address each of the nine competencies. While we included
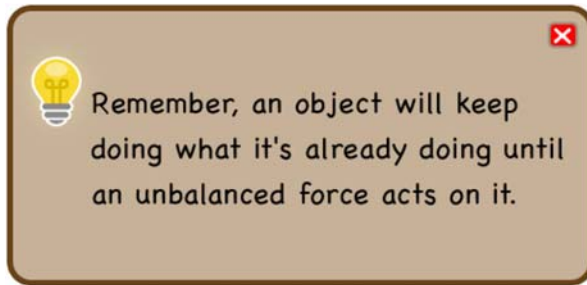
Fig. 3 Advice pop-up window about Newton's first law of motion

Advice in the first set of learning supports, it was not available until after students spent more than five minutes in a level.

User Experience Design One additional element of designing the learning supports for PP was determining how students would access the various learning supports in the game. User experience design is the process of designing products to provide meaningful and high-quality experiences (Norman 2013). The focus is to improve learner's understanding of what can be done. Several principles of multimedia design (Mayer 2009; Mayer 2017), game design (Clark et al. 2016), and usability design (Norman 2013) guided our decisions about our learning supports. In particular, real estate on the gameplay screen is a precious commodity. When it comes to visual stimuli, less is often more. We wanted to limit extraneous processing (e.g., irrelevant content) and encourage generative and focused processing while students were working to solve a game level. Thus, we did not want the screen littered with buttons and menus. However, for students to access the supports, they must be able to find them.

Our first user experience design housed the learning supports in a drawer that opened from the left-hand side of the screen (see Fig. 4). The "handle" for the drawer was always visible as a green triangle labeled with the word "support." When students clicked the green triangle, it opened to reveal icons for each of the embedded learning supports. To help students remember the learning supports when struggling, the green handle would glow after students had been working on a level for five minutes. This is an example of the signaling principle of multimedia design, which is used to direct students' attention to a specific area or item on the screen (Mayer 2009; Mayer 2017).
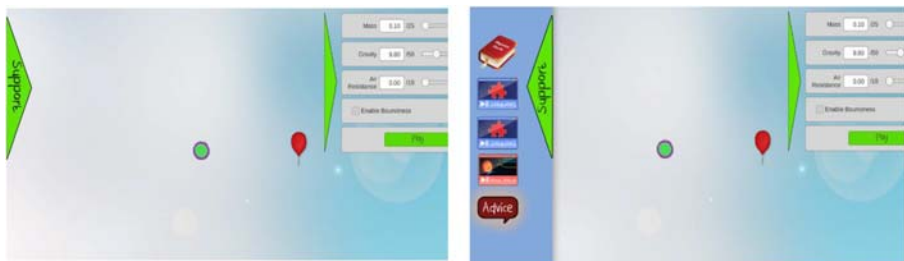

Fig. 4 Initial user experience for the learning supports with "Support" drawer closed (left image) and opened (right image)

In this case, the glowing handle let students know there was a new support available (i.e., Advice).

Usability 1 Findings

Our first usability study occurred in September 2017. Participants included 24 high school students (9th–11th grades) in a public K-12 school in Florida. The students played PP (consisting of both sketching and manipulation levels) for 150 min across three days.

For the first usability study, we included the four previously discussed learning supports. Most students worked in pairs, sharing one laptop, and switching who was in control for each new game level. We gathered data on students' thoughts and attitudes when playing the game through a think-aloud protocol designed for this study. We also recorded players' actions and their articulated thoughts in two ways: (a) via video, using Open Broadcaster software for all actions and utterances; and (b) via a paper-and-pencil observational checklist that we created. Because this was our first usability test, we focused our observations on playability and student reactions to the learning supports.

Worked Examples According to the qualitative feedback from the students' gameplay sessions, WEs were the overall favorite learning support. Viewing the WEs elicited some eureka moments (e.g., shouting "*Ohhhh!*" and raising both hands in the air). As challenges are often good for learning (Shute and Ke 2012), some students indicated they only wanted to watch WEs if they were completely stuck or struggling. Other students preferred to figure out the solution entirely on their own, instead of watching the WE (e.g., one student noted that, "*Watching the Worked Examples would ruin the fun of playing the game!*").

Physics Facts Not surprisingly, students rated the Physics Facts document as their least favorite learning support. Students stated that it required too much reading and should be more interactive. Only a few students spent time viewing the Physics Facts after opening it initially.

Hewitt Videos While the Hewitt videos were infrequently accessed, students rated them positively. Most students reported that the videos were helpful. Some stated they already knew the content. In one case, a student exclaimed after watching a Hewitt video, "*Ahhhaaaa! The science behind it!*" One researcher also observed a case when a student was struggling with a level, but then solved it after watching a Hewitt video. When asked, some students were able to explain the connection between the video and the associated game level. Other students, however, did not watch the entire video because they felt it did not help them solve game levels or it was too long.

Advice Students reported mixed opinions about the Advice support in the first usability study. Some students thought it was mildly helpful. One student liked Advice because it did not give away the answer. Most students, however, commented that the Advice statements were too general and did not help them solve the level, or that the statements

were hard to understand. Based on the feedback from the students, we decided to remove the Advice support.

User Experience Students typically did not initially access the learning supports during gameplay. However, they opened the learning supports drawer when reminded by the researchers. After the initial reminder, students sometimes accessed the supports in subsequent levels without prompting from a researcher. Students reported they did not see the glowing handle. This suggests it was an inept signal for the action we wanted them to perform. Students who solved or exited levels before the five-minute threshold did not even get the opportunity to see the glowing handle. Based on these results, we decided to re-design the way students access the learning support for our next usability test.

Second Iteration of Learning Supports

Based on the findings from the first usability study, we embarked on a second round of learning support design and development. We separated the supports into two main types: gameplay support and physics support. Some of the original supports remained. Hewitt Videos continued to be a part of the set of learning supports, as did Worked Examples. The narration was removed from the WEs to keep the support focused on game mechanics. The original Physics Facts support morphed into three separate supports: physics animations, interactive definitions, and a glossary.

Physics Animations Since the original Physics Facts support was not well received, we explored other ways to support physics understanding in the game. One key aspect of learning support design discussed in the original project proposal entailed providing multiple representations of the targeted physics competencies. Over the course of several weeks, the team worked on developing visual representations (i.e., static pictures) and symbolic representations (i.e., formulas—discussed in the next section) to illustrate the 28 physics terms from the original support. However, we discovered many of the terms did not lend themselves well to static images. Consequently, we used the examples from the original Physics Facts document to create visual representations of the physics terms in the game environment.[1]

In the example targeting "momentum," students can clearly see the relationship between mass and momentum—more mass equals greater momentum when velocity is constant. The two green balls move with the same velocity, but the one with 10 kg mass has greater momentum than the 1 kg mass and can easily push the blue box out of its way (see Fig. 5).

The first step in designing the new physics animations was to create prototypes for the whole team—especially our physics experts—to evaluate. We started with two terms, gravity and coefficient of restitution (i.e., bounciness), and created game footage illustrating them. The videos were presented to the whole team, who liked the idea, and tasked the LS design team to storyboard ideas for all 26 remaining terms. We decided

---

[1] For an example of a physics animation, see: https://www.youtube.com/watch?v=QAocTAAHjt4&feature=youtu.be.
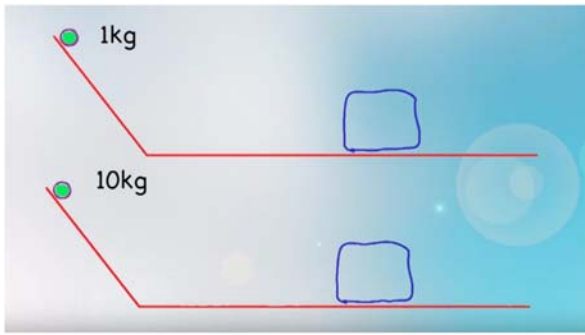
Fig. 5 Physics Animation example about momentum

on the following parameters for the remaining animations: (a) create one animation per term, (b) keep the videos silent (i.e., with no narration or background music to minimize noise in the classroom), (c) provide a replay button at the end, and (d) use minimal text on screen. We also discussed the evaluation criteria for each animation— e.g., the animation must highlight only the relevant physics competency accurately at a middle school level of understanding.

With the production protocol and evaluation criteria in place, we developed animations for all relevant terms. We used the following iterative process: (1) the LS design team created a storyboard with input from physics experts, (2) a prototype was created from the storyboard, (3) the prototype was presented to the physics experts for feedback, (4) the prototype was revised, (5) the animation was presented to the entire team for feedback, (6) steps 4 and 5 were repeated until the final animation was approved, and (7) the approved animations were incorporated into the game by the technical team. Overall, the team created 23 individual animations related to the pertinent physics terms in the game, as a few highly associated terms used the same animation (e.g., motion and kinematics). All animations were implemented in the game before the second usability study. The animations appeared in two places, as a standalone support when students clicked animation and coupled with the cloze tasks.

Interactive Definitions Based on students' feedback in the first usability study, we prototyped several mini games to try to gamify the Physics Facts support. After almost a dozen discarded prototypes, the team decided to go with a simpler solution—make the definition component of Physics Facts interactive based on the physics animation we created. Our first idea was to use drag-and-drop functionality for students to create the definitions. Instead of chunking the whole definition, we decided to use cloze tasks (i.e., fill in the blank) and kept the drag and drop mechanic. With the guidance and feedback of our physics experts, we turned each term's definition into a statement with a series of missing pieces.

Once we finalized the cloze tasks for all 28 physics terms, we gave them to a local high school physics class in Florida for feedback. The feedback was mainly positive. Students suggested we change the order and reduce the options for certain terms. Based on their feedback, each cloze task was revised to have five blanks with short missing pieces (i.e., ≤ 5 words). The final versions of the cloze tasks were implemented into the game with each physics term (e.g., elastic potential energy—EPE) linked to at least one

game level (i.e., in this case EPE was linked to levels with a springboard solution). To help students accurately complete the cloze tasks, we included a second learning support, short animations of the physics terms, with the definitions (see Fig. 6). We discuss the design and development of the physics animation support in the following section.

Formulas In many fields, like math and science, students need symbolic representations to interpret meaning (e.g., Bruner 1964; Plass et al. 2015; Uttal et al. 2009). As part of our overhaul to the original Physics Facts support, we added symbolic representations or formulas to the physics learning supports. An example of a formula representation is shown in Fig. 7.

During the design and development processes, we fine-tuned the representation format and settled on the following parameters. First, center the formula at the top in one line. Second, use a vertical fraction format to represent division. Third, define all pertinent variables below the formula in one column using smaller font. Finally, use a simulated blackboard background to present all formulas.

Glossary With our revision of the Physics Facts support, each game level was linked to one of the 28 physics terms and included the supports discussed above for that term. We did not want to limit student inquiry to just one term, so we created a Glossary support that contained shortened definitions for all 28 terms. When students clicked on the Glossary support, they could scroll through all the physics terms related to gameplay.

User Experience Based on the results of the first usability study, we revised the way students accessed the supports. We still wanted learning support access to be the student's choice, so we created a help button. The help button was placed in the lower right-hand corner of the screen. When clicked, it opened a pop-up window with buttons representing the two main branches of learning supports: gameplay support and physics support (see schematic representation of supports in Fig. 8).
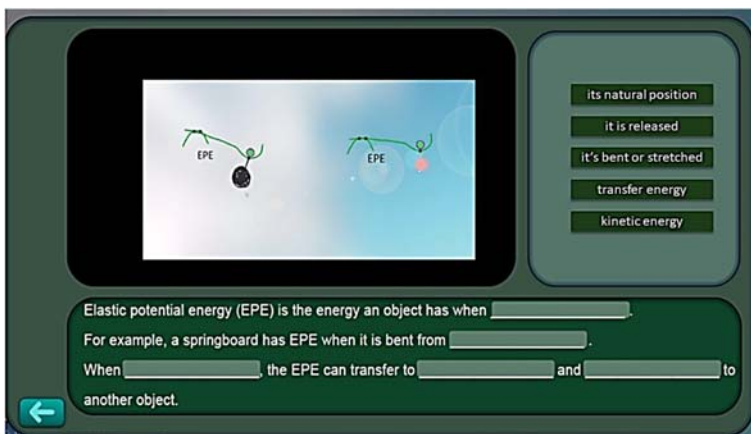

Fig. 6 Screen shot of physics learning support, "Interactive Definition" showing the layout of the term's animation and the cloze task

$$a = \frac{\Delta v}{\Delta t} = \frac{v_f - v_i}{t_f - t_i}$$

v_f: final velocity
v_i: initial velocity
t_f: final time
t_i: initial time
Δ: change

Fig. 7 An example of the Formula support

The gameplay support was housed between two buttons, "Review the Tutorials" and "Show me a Solution." The physics supports were accessed through one button, "Show me the Physics" that opened a second pop-up menu with icons linking to the five different types of physics supports (see Fig. 9).

Usability 2 Findings

We implemented our revised game for the second usability study in April 2018 with forty-four 8th grade public school students from a K-12 school in Florida. We randomly assigned students to either the learning support (LS) group or non-learning support (Non-LS) group. We observed the students' gameplay for four days (i.e., about 50 min per day). We measured students learning through a comparison of a pretest before gameplay and a posttest after gameplay. Students also completed a game satisfaction questionnaire and a learning-support satisfaction questionnaire (LS group only) after gameplay. Students who completed the study received a $10 gift card.

The students in both groups started with game tutorials and played both sketching and manipulation levels. Contrary to our expectations, an ANCOVA revealed that the
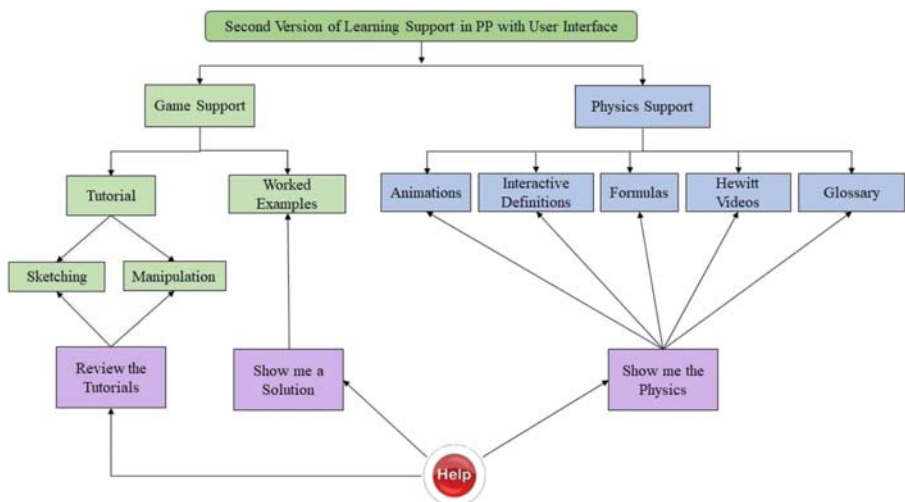

Fig. 8 Second version of learning supports in *Physics Playground* for Usability Study 2

Fig. 9 Screenshot of physics learning supports menu from second usability study

non-LS group had higher posttest scores than the LS group, holding pretest scores constant ($F(1,43) = 4.06$, $p = 0.05$, $d = 0.61$). We analyzed the test items and found that the pre- and posttests were difficult for the students, with low reliabilities ($\alpha_{pretest} = 0.43$ and $\alpha_{posttest} = 0.40$), which have implications for the validity of the tests. As a result of our item analysis, we revised the problematic items for subsequent testing, and present the new, improved reliabilities in the Final Study section of this paper.

The satisfaction questionnaire results showed that all students were excited about and engaged with playing the game. Students in the LS condition also reported that the learning supports were helpful. Worked Examples continued to be the favorite and most frequently used support. Students additionally reported that the tutorials were too long, and observations confirmed that students did not accurately recall the tutorial information later during gameplay. Given the limitations of the study, more testing of this second set of learning supports was warranted, along with revising the pretest and posttest items in line with an item analysis to improve reliabilities.

Third Iteration of Learning Supports

Based on the second usability study results, we turned our focus to the alignment between our learning supports, game level solutions, and the targeted physics content. We developed one final learning support connecting how students solve levels to the physics involved in the solution ("Physics Videos"). We revisited the original idea of hints and completely changed the game tutorials.

Physics Videos Feedback from the first and second usability studies revealed that we needed physics supports that were conceptually aligned to the solution of a level. This type of scaffolding consists of showing examples of failed and successful attempts that are related to the target competency to assist problem solving (Muldner and Conati 2010). Thus, we decided to create Physics Videos for each intersection of solution (e.g., ramp) and competency (e.g., Newton's 1st Law). The physics experts examined these intersections for feasibility (i.e., the connection could be illustrated in a short gameplay video). Overall, we found 16 intersections appropriate for the Physics Video support.

We discussed using tutorial levels as the gameplay backdrop for the physics videos. Based on the coherence principle of multimedia design, using the tutorial levels would

limit the extraneous elements in the videos and enhance learning (Mayer and Fiorella 2014). Consistency across videos would allow students to attend to what is different in the video (i.e., the physics explanations). However, the similarity across videos could cause confusion, (i.e., students could close the video because they thought they have already seen it). As a solution, we decided to create a title slide at the beginning of each video to state the solution and competency illustrated, highlighting the unique intersection of each video.

Each Physics Video focused on how the variables within a solution (e.g., height of the pendulum) affect the solution. We used the same format per video: (1) Introduce the competency to be presented in the video (e.g., "*Here you are going to see how to transfer energy to the ball using a pendulum*"); (2) Define terms (e.g., "*Gravitational potential energy is the energy of height…*"); (3) Show a failed attempt to solve the level (e.g., "*The pendulum does not have enough angular height…*"); and (4) Adjust the variable (i.e., the height of the pendulum) to show a successful attempt to solve the level.[2]

Emphasizing how variables are changing within the video was an important part of the visual design. We first prototyped meters that would fill and empty, accordingly. However, the distance between the meters and the ball's movement made it hard to concentrate on both at the same time, causing a split-attention effect where learners had to split their attention between the words and the animated graphics, making mental integration of the two sources of information difficult (Chandler and Sweller 1992; Johnson and Mayer 2012). To avoid this issue, we drew on the contiguity principle (Mayer and Fiorella 2014). That is, students learn more when words and graphics are located near each other, and closer elements are perceived as related to each other, while separated elements as less related (Johnson and Mayer 2012; Mayer and Fiorella 2014; Mayer 2003). Thus, we synchronized the visual representations of loss and gain with the ball's movement. For example, to show changes in energy, the energy acronym (e.g., GPE = gravitational potential energy) would move with the ball and the font size changed to represent the change in magnitude (see Fig. 10). Likewise, to increase the relationship between acronym and elements, we used green color for the acronyms related with the ball state, and we matched the colors between other graphic elements (e.g., weight) to their related acronym. This strategy was based on the visual principles of unity and similarity, which states that elements with shared characteristics are perceived as belonging together (Lauer and Pentak 2011). We also used variance in font-size to show the relationship between variables in formulas. For instance, if variable X decreases when Y increases, we would decrease the font size of X and increase the font size of Y, leaving the remaining variables untouched.

We added audio narration to explain specific physics terms (e.g., Gravitational Potential Energy and Kinetic Energy) associated with both failed and successful attempts, and we minimized on-screen text. This decision was based on research by Moreno and Mayer (2002) who reported that students perform better and have higher recall and satisfaction levels when learning environments use speech rather than on-screen text to deliver instructional content. In some physics videos, the on-screen change between the failed and successful attempt was not as noticeable as in others. To solve this problem, we added a focal point (highlighted area) to make certain

---

[2] To see an example of a physics video: https://www.youtube.com/watch?v=jp8b5fAF3H4&feature=youtu.be
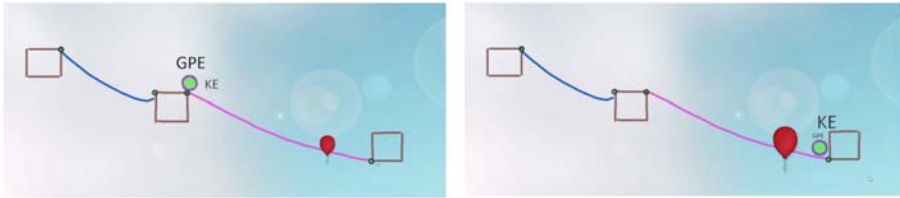
Fig. 10 Example of Physics Video showing the movement and change of font size used to highlight the change in energy

movements more obvious (see Fig. 11). The focal point is a visual design principle used to draw the viewers' attention to a specific element and encourage them to follow the highlighted part. In general, a focal point is achieved by increasing contrast, like when an object has a different color from the overall elements in a composition (Lauer & Pentak, 2012). In our videos, we decided to place a semi-transparent black layer on the screen except for where students should focus attention, creating contrast in color tones.

The LS team and physics experts iteratively revised all physics videos to ensure they followed our design parameters and accurately explained the physics content. The development of each video followed five stages:

(1) *Scripting* – The physics experts created a script for each physics video. The scripts provided the narration for the competency definition, the failed attempt, and the successful attempt, as well as direction for the game actions (i.e., game footage) needed to illustrate the narration.

(2) *Storyboarding* – The LS design team created storyboards for each video based on the physics experts' scripts. Storyboards consisted of gameplay footage with the proposed overlays. In a few instances, the LS design team first created slides representing the game action for each segment of the narration before generating any gameplay footage. Once developed, members of the LS design team presented the storyboards to the whole team for feedback and approval.
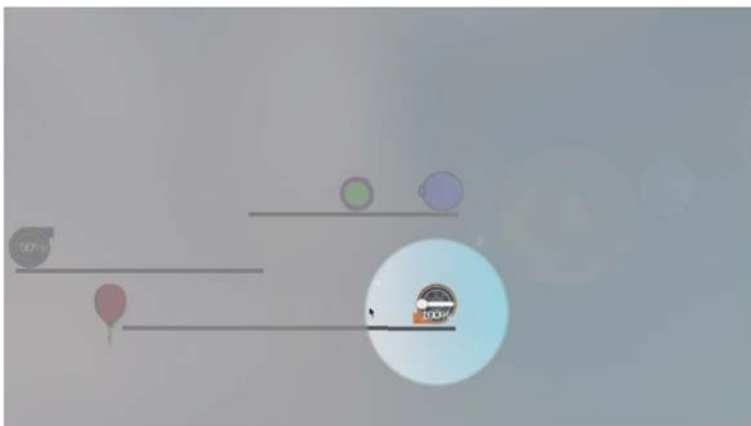

Fig. 11 Example of Focal Point in the Physics Videos showing the contrast used to highlight the changes made to a dynamic blower

(3) *Audio recording* – Once the storyboard was approved by the team, we recorded the narration. One member of the team acted as the narrator for all physics videos. We used the same computer and software to record all audio files.

(4) *Video editing* – Members of the LS design team edited the gameplay footage, created and animated the overlays, and edited the audio file to create each video.

(5) *Revising* – The whole research team iteratively revised each new video. As more videos were developed, we gained insight for improvement. At times, these insights were applicable to previously developed videos. Therefore, each physics video went through several rounds of revision during the design and development process.

Hints Providing hints regarding a level's solution was one of the original learning supports mentioned in the project proposal. During the first iteration of supports, we developed the Advice support, but it was not successful. After the second usability study, we revisited the idea of using Hints as partial solutions—directing students to the correct solution (e.g. "*Try using a lever*") without disclosing the full solution.

The early version of Hints linked the physics with the game action for the level. For example, the hint for the game level named Downhill stated, "The ball will obey Newton's 1st Law unless something changes its motion." However, the format for these hints were inconsistent and we wanted to focus on the required game action. The final version of Hints used the same format. That is, for sketching levels, the Hints support states: "Try drawing a [the name of an applicable simple machine]." For manipulation levels, the Hints support states: "Try adjusting the [the name of one of the sliders]." The user interaction for Hints also changed over time. Ultimately, we decided to use a pop-up window, like the one originally used for the Advice support (shown in Fig. 3).

Game Tutorials Originally, the game tutorials were video demonstrations that included opportunities for students to try the previously demonstrated action. Based on our field observations and usability studies, the game tutorials were complicated, long, easy to forget, and demotivating for the students. Therefore, we decided to change the format of the game tutorials to address these concerns. We made the tutorials into a series of game levels, each focused on one solution method. Students play the tutorial levels with on-screen, step-by-step instructions (see Fig. 12).

Game Tips Based on our observations, students often forgot what they learned in the tutorials (e.g., the steps needed to draw a springboard). Our first idea was to add a new button to the help menu, "Review the Tutorial." When students clicked the button, they would be taken back to the appropriate tutorial for the level (i.e., sketching or manipulation). A "Back to the Level" button was added so students could return to the game level they were playing. However, this design was inefficient. Students needed a quick reminder, not another round of tutorials. We addressed this issue by adding a new area of learning supports accessed by clicking "Show me Game Tips" (see button on the right side of Fig. 13).
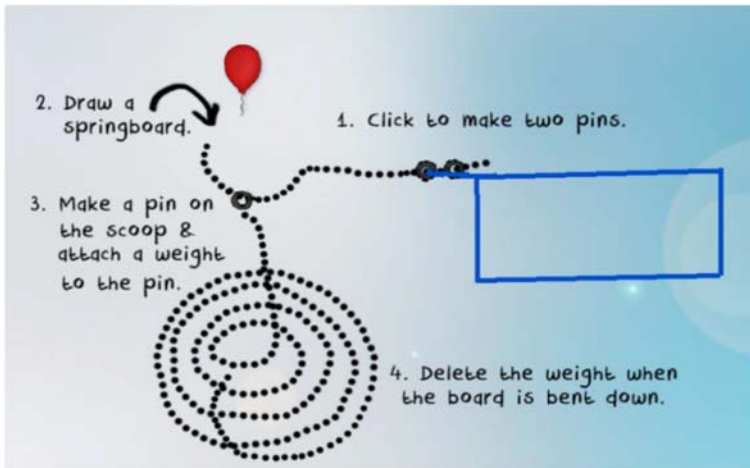
Fig. 12 Screenshot of tutorial level for drawing a springboard solution

By clicking "Show me Game Tips" in any sketching level students would see a menu with three tabs:

(1) *Controls* – which explain the various mechanics of the game (e.g., how to nudge the ball, draw/delete an object, make a pin, see Fig. 14)

(2) *Simple Machines* – which contains images of the four sketching tutorial levels that when clicked enlarge to full screen.

(3) *My Backpack* – briefly explains the functionality of My Backpack (i.e., the progress-tracking and incentive system of the game, which is outside the scope of this paper).

By clicking "Show me Game Tips" in any manipulation level, students would see a menu with two tabs: (1) *Tools* – which includes images and brief explanations about the functionality of the game tools pertaining to the manipulation levels (e.g., puffer, blowers, and sliders); and (2) *My Backpack* – which briefly explains the functionality of My Backpack (the same as in sketching levels).

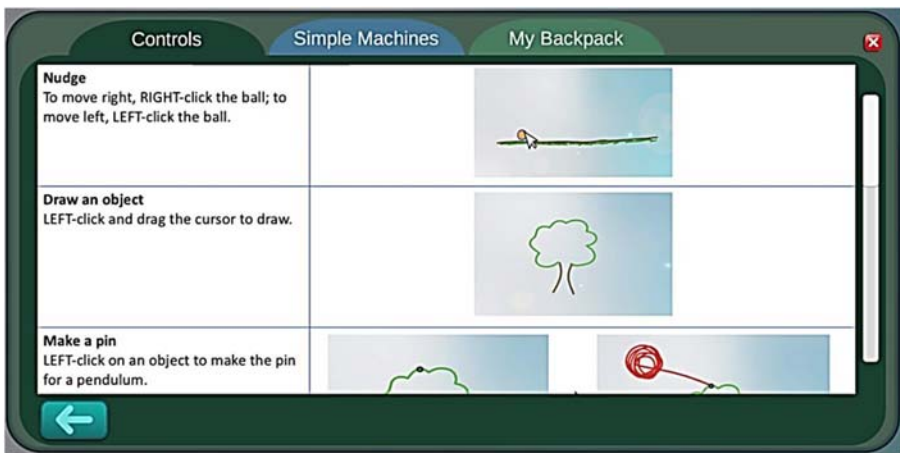Fig. 13 Screenshot of the final Help button main menu

Fig. 14 Screenshot from sketching level Game Tips support showing the Controls tab

Usability 3 Findings

We conducted our third usability test in September 2018 with fourteen middle school students in a charter school in Florida. Participants were selected through a convenience sampling method and included six 7th graders and eight 8th graders. Our focus for this third iteration of testing was our new learning support—physics videos. In the usability study, we used seven physics videos connected to two of the game's nine physics competencies: energy can transfer and properties of torque. We selected 30 sketching levels of varying difficulty for students to play that focused on our two focal competencies. We matched the most appropriate physics video to each level. The results from this third iteration showed significant learning gains via a paired-sample t-test analysis ($M_{pre} = 7.93$, $SD_{pre} = .99$, $M_{post} = 8.79$, $SD_{post} = 1.31$, $t(13) = 2.20$, $p < 0.05$, $d = .60$). Moreover, students found the learning supports satisfying and useful, and found the physics videos in particular to be helpful for learning physics. Students also continued to report high levels of enjoyment playing the game. We also included the new sketching tutorial levels in the third usability study. Students responded favorably and it took much less time for students to complete the game tutorial than in previous studies.

Final Version of the Learning Supports

After the third usability study, we completed our design and development of the learning supports for *Physics Playground*. We kept the help button as the mechanism for accessing the supports. When clicked, it opens a pop-up window with three options: "Show me the Physics," "Show me a Solution or a Hint," and "Show me Game Tips." Each of the supports included in the final set of learning supports is accessed through one of the three buttons (see Fig. 15). To evaluate the final set of learning supports, we conducted one final quasi-experimental study.
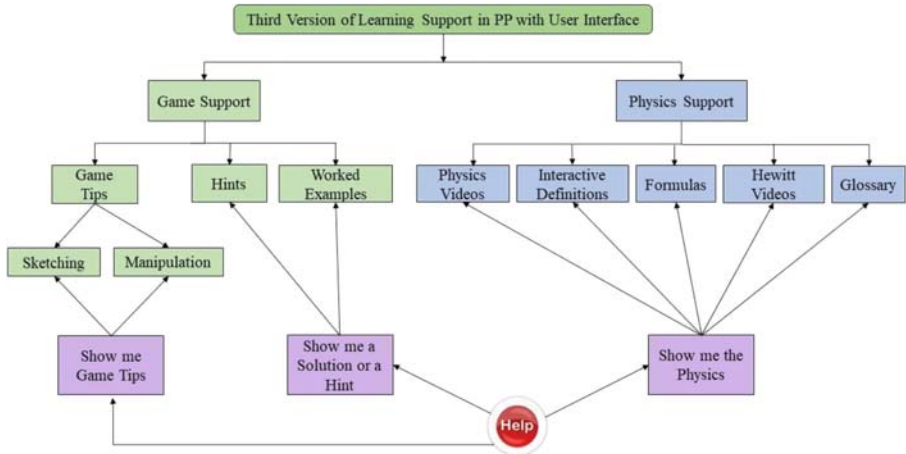
Fig. 15 Final version of learning supports implemented in *Physics Playground*

Final Study The three main goals of the final study were to: (1) validate our in-game stealth assessment measures of physics competencies, (2) test the influence of adaptive sequencing of game levels on learning, and (3) test the effectiveness and perception of the learning supports. Only the results of goal three are pertinent for our discussion of the design and development of the learning supports. For a discussion of goals one and two, see Shute et al. (2020). We conducted the final study in May 2019 with 263 high school students from a large K-12 school in Florida. The students engaged in more than four hours of gameplay across six days in 50-min sessions per day. We conducted a pretest at the start of the first session, and a posttest along with two surveys (i.e., a game- and support-satisfaction questionnaire) at the end of the last session. The tests (pretest and posttest) have been revised across two years of testing, and the current reliabilities (Cronbach's α values) were: pretest = .77; posttest = .82; $n = 263$. Students who completed the study received a $30 gift card at the end.

The game consisted of ten brief tutorial levels and 81 game levels (38 sketching and 43 manipulation levels) covering all nine competencies. We ordered the competencies from easy to difficult based on the conceptual difficulty of the physics content, in line with the recommendation of our two physics experts working on the project. The ordering was: Newton's 1st law, energy can transfer, energy can dissipate, properties of momentum, conservation of momentum, properties of torque, equilibrium, Newton's 2nd law, and Newton's third law. Within each competency, we then ordered the game levels from easy to difficult. All participants had access to the learning supports in all levels, and the help button was always activated.

Final Study Preliminary Results The first question we wanted to address was whether the students, overall, learned any physics from this support-enhanced version of the game. We computed a repeated measures ANOVA on pretest to posttest scores of the students who played the game and found that indeed, there were significant improvements ($F(1, 198) = 9.53$; $p < .01$). In contrast, our control condition ($n = 64$)—who did not play the game but did take the pretest and posttest—showed no gains ($F(1, 63) = 0.002$; $p = .97$).

Next, to examine students' perceptions of the learning supports, we reviewed their responses to our 5-point Likert scale survey questions. As shown in Table 2, overall, the students found the supports to be helpful, easy to use, and agreeable (i.e., not annoying). Specifically, they really liked the worked examples and the physics videos. Regarding students' preferences for playing the game with or without supports, the data are mixed. That is, the mean of 3.28 was about in the middle of the 1–5 scale with a somewhat large standard deviation suggesting that people differed in their attitudes about using supports in the game—some preferred to play the game on their own while others preferred to access supports when struggling.

Our third question examined the degree to which students' attitudes toward the supports predicted their gain (posttest minus pretest) score. Results from a linear regression analysis entering all six variables showed that among the support-attitude variables, one (*supports were helpful*) was a strong positive predictor of gain score ($t = 2.71$; $p < .01$; $\beta = .23$; see Table 3).

This suggests that students who perceived the learning supports as helpful gained more physics knowledge than those who did not. The result further highlights the potential effectiveness of the learning supports. Two other variables, however, were inversely related to gain: *supports were easy to use* ($t = -2.06$, $p = .04$, $\beta = -.17$) and *I prefer solving levels with supports* ($t = -2.33$; $p = .02$, $\beta = -.19$). Note that these results may be related to the player's self-efficacy in using computers, playing games, and solving physics problems; none of which were measured. This suggests that the relationship between student ability and self-efficacy, support use, and learning is complex and needs more study. We are currently processing the event logs to get a more detailed picture of how students used the learning supports during the game.

Conclusions and Future Directions

Our primary goal in writing this paper – in addition to honoring Jim Greer's memory – was to highlight the value of iterative design in game-based learning studies, and to share the specific types of learning supports that are suitable for use in educational games. The design experience should shed light on how to design next-generation learning games that blur the distinction between learning and assessment.

Table 2. *Questionnaire items relating to learning supports (N = 195).*

Item	Mean	SD
1. *Show Solutions* (e.g., worked examples) helped me solve the level	3.82	1.25
2. *Show Physics* (e.g., Physics videos) helped me solve the level	3.60	1.09
3. In general, the supports were agreeable/unobtrusive	3.46	1.08
4. In general, the supports were easy to use	3.78	0.83
5. In general, the supports were helpful	3.56	1.09
6. I prefer solving levels with supports	3.28	1.28

Table 3. *Regression Analysis Summary for Learning Gain.*

Variables	Unstd. Coef. B (SE)	Std. Coef. β	t
(Constant)	−.62 (1.10)		−0.57
Show solutions helped me solve the level	.29 (.22)	.12	1.29
Show physics helped me solve the level	.09 (.25)	.03	0.35
Supports were not annoying	.37 (.21)	.14	1.71
Supports were easy to use	−.59 (.29)	−.17	−2.06*
Supports were helpful	.61 (.23)	.23	2.71**
I prefer solving levels with supports	−.42 (.18)	−.19	−2.33*

Note. Unstd. Coef. = Unstandardized Coefficient; Stand. Coef. = Standardized Coefficient. SE = Standard Error. The dependent variable is gain; * indicates $p < .05$; ** indicates $p < .01$

In general, we have confirmed that students' physics understanding can be improved through gameplay. The results of previous studies using *Physics Playground* (e.g., Kim and Shute 2015; Shute et al. 2013; Shute et al. 2015) were replicated herein. That is, playing the game has an overall positive effect on students' physics understanding. After three rounds of iteration in the past three years, we have developed and implemented in-game learning supports that helped to lay a conceptual foundation of physics for students. The supports with multiple representations have further enhanced students' physics understanding, beyond just playing the game. However, the supports were still not as effective as we expected. Potential explanations include the dosage of learning support access and the motivation of using learning supports. That is, while the learning supports were available at all times, it was up to the student to access them. Some chose to do so; others did not.

Some students gained little in terms of physics knowledge (i.e., lower than the mean gain) but rated the supports as easy to use. It is possible that these students did not engage with the supports extensively but rated them as easy-to-use because the mechanism was intuitive. Others demonstrated good physics gains but reported a preference *not* to use supports. These students could have been engaged in gameplay, performing well in the game, but perhaps viewed the supports as a kind of "cheat." Again, these preliminary results and potential explanations suggest the need for in-depth analyses of students' engagement with learning supports. The relationship between learning, gameplay, support access, and attitude may not be simple and linear. Instead, moderators such as the context of tasks, player typology, and prerequisite skills likely complicate the relationship. Better understanding of students' use of learning supports will allow for better adaptivity to meet students' needs. In the future, other variables beyond students' understanding of physics competencies could be included in the calculations for adaptive delivery of learning supports.

We took care to integrate the learning supports into the game to avoid disruption of flow. From the outset, we decided there would be no talking heads delivering content. Although some supports (e.g., Hewitt videos, glossary, and formula) involved a certain level of formalization, they did not scare or bore the students. Again, students could access the supports voluntarily to facilitate their own gameplay. Therefore, learners' autonomy was preserved, which in turn can increase the motivation and enjoyment of

learning (Vansteenkiste et al. 2004). Further exploration into the frequency and duration of support usage can provide evidence about what supports provide the largest impact on learning and student performance (e.g., Shute et al. 2020). In addition, our ongoing efforts to personalize and automate the delivery of learning supports to provide just-in-time learning has the potential to create even larger learning gains.

This paper focused on its learning support feature, but other research has been conducted using this game, such as stealth assessment of creativity (Shute and Wang 2016), persistence (Ventura and Shute 2013), as well as affective states (Shute et al. 2015). Educational data mining has been used within *Physics Playground* to determine when a player is likely to quit a level (Karumbaiah et al. 2018; Spann, Shute, Rahimi, & D'Mello, 2019), and the game has been used to examine collaborative problem-solving skills (Chen, Shute, Stewart, Yonehiro, Duran, & D'Mello, 2019). Future research may consider inclusive design in educational games, such as creating in-game learning supports that are enticing to students with diverse characteristics, needs, and motivations, while not disrupting the flow. In summary, accurately modeling and supporting all kinds of learners using engaging, interactive environments is Jim Greer's lovely legacy to us all.

References

Almond, R. G., Mislevy, R. J, Steinberg, L.S., Yan, D. & Williamson, D.M. (2015). Bayesian networks in educational assessment. Springer. ISBN: 978-1-4939-2124-9.

Bruner, J. S. (1964). The course of cognitive growth. *American Psychologist, 19*, 1–15.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*, 233–246. https://doi.org/10.1111/j.2044-8279.1992.tb01017.x.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79–122.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optical experience.* New York: Harper Perrennial.

Johnson, C. I., & Mayer, R. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology, 18*(2), 178–191. https://doi.org/10.1037/a0026923.

Karumbaiah, S., Baker, R. S., & Shute, V. (2018). Predicting quitting in students playing a learning game. In *11th International Conference on Educational Data Mining*, (pp. 1–10). Buffalo, NY.

Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education, 87*, 340–356.

Lang, J., & O'Neil, H. (2008). *The effect of presenting just-in-time worked examples for problem solving in a computer game.* Paper presented at the American Educational Research Association, New York, US.

Lauer, D. A., & Pentak, S. (2011). *Design basics* (eighth ed.). Boston, MA: Wadsworth.

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction, 13*(2), 125–139. https://doi.org/10.1016/S0959-47520200016-6.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge, NY: Cambridge University Press.

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403–423. https://doi.org/10.1111/jcal.12197.

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). New York: Cambridge University Press.

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_02.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94*, 598–610.

Muldner, K., & Conati, C. (2010). Scaffolding meta-cognitive skills for effective analogical problem solving via tailored example selection. *International Journal of Artificial Intelligence in Education, 20*(2), 99–136. https://doi.org/10.3233/JAI-2010-0004.

Norman, D. (2013). *The design of everyday things*. New York, NY: Basic Books.

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist, 50*(4), 258–283.

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3.

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. 503–524). Charlotte, NC: Information Age Publishers.

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..

Shute, V. J., Rahimi, S., & D'Mello, S. (2019). The productive role of cognitive reappraisal to regulate frustration during game-based learning. *Computers in Human Behavior, 100*, 358–369. https://doi.org/10.1016/j.chb.2019.03.00.

Shute, V. J., Almond, R. G., & Rahimi, S. (2019a). *Physics Playground* (version 1.3) [computer software]. Tallahassee, FL: Retrieved from https://pluto.coe.fsu.edu/ppteam/pp-links/

Shute, V. J., D'Mello, S. K., Baker, R., Cho, K., Bosch, N., Ocumpaugh, J., Ventura, M., & Almeda, V. (2015). Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. *Computers & Education, 86*, 224–235.

Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 43–58). New York, NY: Springer.

Shute, V. J., Rahimi, S., & Lu, X. (2019b). Supporting learning in educational games: Promises and challenges. In P. Díaz, A. Ioannou, K. K. Bhagat, & J. M. Spector (Eds.), *Learning in a digital world - smart computing and intelligence* (pp. 59–81). Singapore: Springer.

Shute, V. J., Rahimi, S., & Smith, G. (2019c). Game-based learning analytics in physics playground. In M. Chang & A. Tlili (Eds.), *Data analytics approaches in educational games and gamification systems* (pp. 69–93). New York: Springer.

Shute, V. J., Ventura, M., & Kim, Y. J. (2013). Assessment and learning of qualitative physics in Newton's playground. *The Journal of Educational Research, 106*, 423–430.

Shute, V. J., Rahimi S., Smith, G., Ke, F., Almond, R., Dai, C-P, Kamikabeya, R., Liu, Z., Yang, X., & Sun, C. (2020). Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity, and learning supports in Physics Playground. Manuscript submitted for publication.

Shute, C. A., Rahimi, V. J., & D'Mello, S. K. (2019d). The productive role of cognitive reappraisal in regulating frustration during game-based learning. *Computers in Human Behavior, 100*, 358–369 https://doi-org.proxy.lib.fsu.edu/10.1016/j.chb.2019.03.002.

Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives, 3*, 156–159. https://doi.org/10.1111/j.1750-8606.2009.00097.x.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology, 87*(2), 246–260.

Ventura, M., & Shute, V. J. (2013). The validity of a game-based assessment of persistence. *Computers in Human Behavior, 29*, 2568–2572.

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

Zapata-Rivera, D., & Greer, J. E. (2004). Interacting with inspectable Bayesian student models. *International Journal of Artificial Intelligence in Education, 14*(2), 127–163.