International Journal *of* Designs for Learning
2016 | Volume 7, Issue 1 | Pages 107-146

SURGE'S EVOLUTION DEEPER INTO FORMAL REPRESENTATIONS: THE SIREN'S CALL OF POPULAR GAME-PLAY MECHANICS

Douglas B. Clark[1], Satyugjit Virk[1], Pratim Sengupta[2], Corey Brady[3], Mario Martinez-Garza[1], Kara Krinks[1], Stephen S. Killingsworth[1], John Kinnebrew[1], Gautam Biswas[1], Jacqueline Barnes[4], James Minstrell[5], Brian C. Nelson[6], Kent Slack[6], & Cynthia M. D'Angelo[7]
[1]Vanderbilt University; [2]University of Calgary; [3]Northwestern University; [4]Northeastern University; [5]FACET Innovations; [6]Arizona State University; [7]SRI International

We have iteratively designed and researched five digital games focusing on Newtonian dynamics for middle school classrooms during the past seven years. The designs have evolved dramatically in terms of the roles and relationships of the formal representations, phenomenological representations, and control schemes. Phenomenological representations can be thought of as the "world" representations that depict the actual actions and motion of a game as they occur (i.e., the central representations in most recreational games). Formal representations highlight the disciplinary relationships of interest from a pedagogical perspective (such as vector arrows, graphs, and dot traces). Our initial design perspective focused on highlighting the formal physics relationships within popular game-play mechanics. This perspective prioritized a commitment to the phenomenological representations and controls of recreational games, specifically marble-genre games. We designed formal representations around and over the phenomenological representations of that genre. Over the next seven years, we navigated the tensions between the original recreational genre and creating a new genre situated within the formal representations themselves. More specifically, our designs evolved to situate the game-play squarely in the formal representations in terms of the controls as well as in terms of the communication of goals and challenges. We backgrounded phenomenological representations and streamlined visual complexity to focus on key relationships. Our discussion compares our design evolution to the *SimCalc* design evolution recounted in IJDL's recent special issue on historic design cases.

Team Overview. The author list of this paper includes the research team members involved in the design of each phase of SURGE game development including SURGE Classic, SURGE Next, Fuzzy Chronicles, SURGE NextG, and SURGE Symbolic. Our interdisciplinary team includes learning sciences researchers, science education researchers, computer scientists, cognitive scientists, educational technology researchers, and games researchers. The membership of the team has evolved and changed over each phase, but all team members have been central designers across one or more of the phases of research, design, and development. More information about SURGE games, research, personnel, and publications is available at www.surgeuniverse.com

INTRODUCTION

SURGE Classic was funded by an exploratory NSF grant in 2008. As stated in the proposal, the original design plans focused on developing a game that would highlight the salience of the core physics relationships underlying popular game mechanics in a manner that would increase students' understanding of these core concepts without damaging the intrinsic motivating factors of the popular game mechanics. More specifically, the design approach focused on popular commercial game mechanics from marble genre games like *Switchball* (see Figure 1 top right). Reviews and reception were very positive for *Switchball* when it was released in 2007 on PC and Xbox360, and marble genre games have been popular in various forms for three decades.

The original *SURGE* grant was followed by two subsequent grants, one of which is still ongoing. Our design foci have evolved through iterative cycles of design, development, and research. The top left image of Figure 1 presents a version of *SURGE Classic* that faithfully implemented the design approaches articulated in the original grant proposal. The lower image of Figure 1 presents the evolutionary descendent of that original design vision (*SURGE Symbolic*). The evolution can be thought of in terms of an ongoing design tension between popular game mechanics and disciplinary learning goals that emphasize disciplinary representations. The evolution can also be thought of as a shift from focusing primarily on the phenomenological representations of the game "world" (with formal representations layered on top) to focusing centrally on the formal disciplinary representations as the heart of the game (with the phenomenological representations playing a supporting role).

Phenomenological representations are the representations that depict the actual motion and actions of the game as they occur (the central representations of *SURGE Classic*, *Switchball*, and most recreational games). Formal representations highlight the disciplinary relationships of interest from a pedagogical perspective (such as vector arrows, graphs, and dot traces) as well as any intermediate representations created for pedagogical purposes in support of the disciplinary relationships and representations. As can be seen

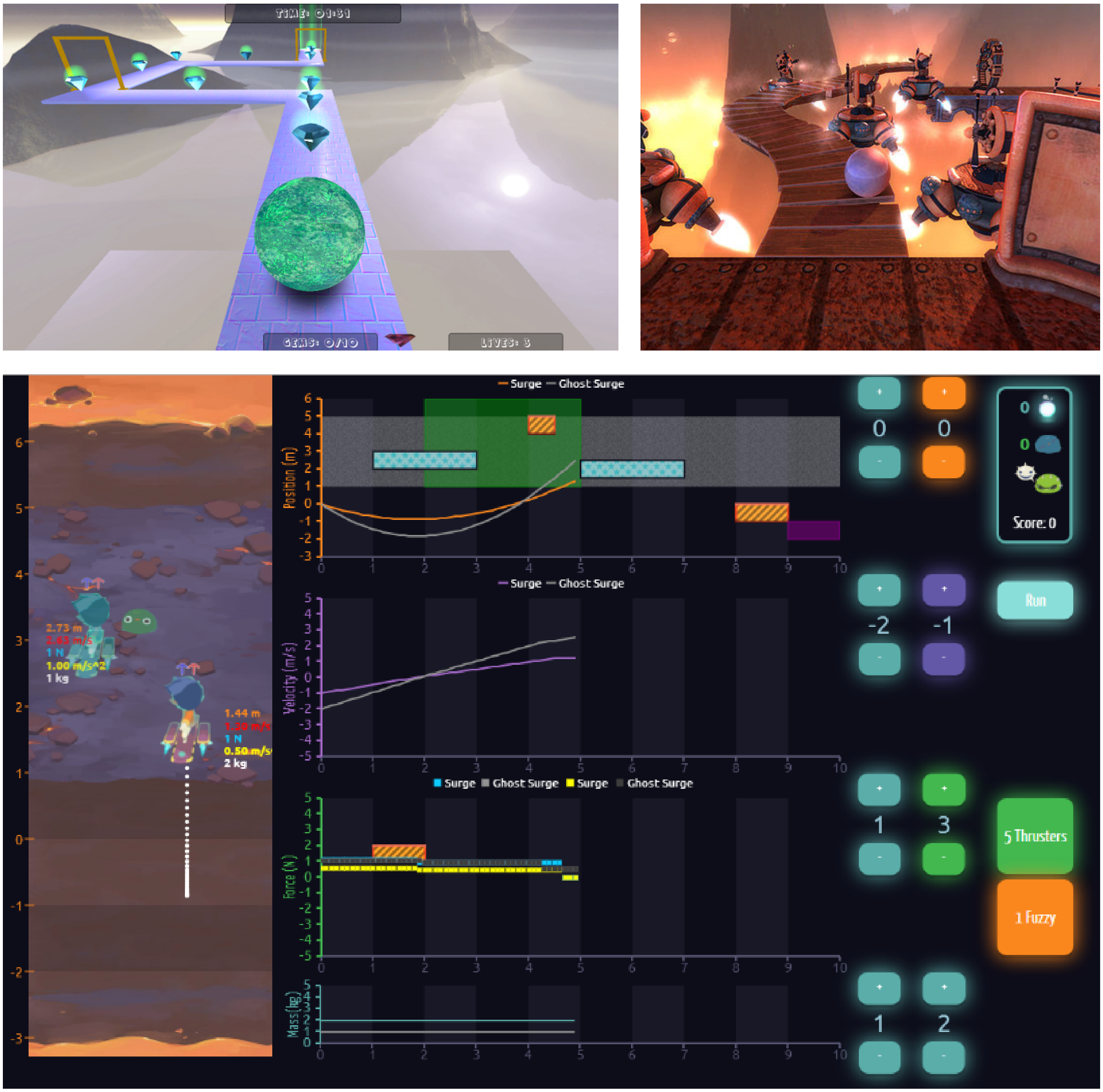


FIGURE 1. Our original vision for *SURGE Classic* (top left) was based on popular recreational marble games like *Switchball* (top right). This original vision focused game play on the "world" representation with formal representations in a supporting role. The vision for *SURGE Symbolic* Phase 3 (bottom) has evolved to focus game play in the formal representations with the "world" representations playing a supporting role.

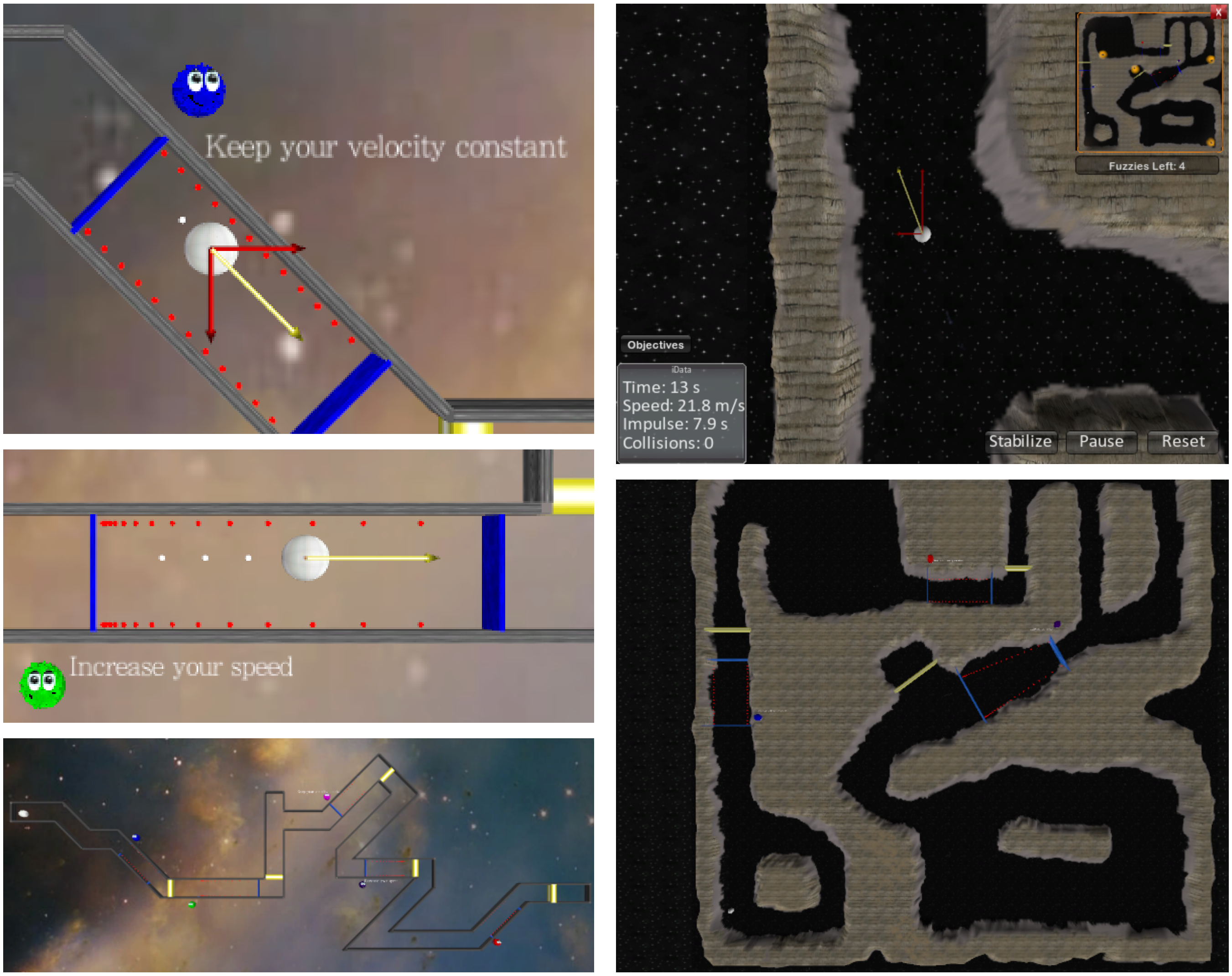


FIGURE 2. Students must guide Surge in her spherical spaceship through maze-like prisons to rescue Fuzzies in *SURGE Classic*. Screenshots from an impulse level (left) and a constant force level (right). The screenshots at the bottom show full levels.

in the top images of Figure 1 and in all of the images of Figure 2, the designs from *SURGE Classic* centrally emphasize phenomenological representations, and these phenomenological representations emphasize the trappings and structures of the popular game genres upon which they are based. The history of our design work can thus be thought of as the progressive ascendance and centrality of the formal representations in the designs of the games.

In Clark, Sengupta, Brady, Martinez-Garza, and Killingsworth (2015), we wrote a synthesis of our work with a focus on *Fuzzy Chronicles*, *SURGE Next*, and *SURGE NextG* (the middle three games of our series), but also including *SURGE Classic* and an early version of *SURGE Symbolic* Phase 2. We wrote that synthesis as part of a theoretical argument for adopting a Science as Practice perspective (Duschl, Schweingruber, & Shouse, 2007; Lehrer & Schauble, 2006a, 2006b; Pickering, 1995) as opposed to a Knowledge In Pieces perspective (Clark, 2006; diSessa, 1993; Hammer, 1996; Sherin, 2001) as the underlying framework for our game design.

Essentially, the theoretical framing underlying *SURGE Classic* and early versions of *SURGE Next* and *Fuzzy Chronicles* involved what we named conceptual integration (Clark & Martinez-Garza, 2012). Clark, Sengupta, et al. (2015) argued for a shift beyond conceptual integration to what we termed disciplinary integration. We refined our theoretical arguments about disciplinary integration in Clark, Sengupta, and Virk (2016) and Sengupta and Clark (in press). Essentially, disciplinary-integration focuses game design on students' manipulation and interpretation of formal disciplinary representations as the central game representations for communicating challenges and opportunities to the player and for the player to model, control, and execute her plans and actions. While Clark, Sengupta, et al. (2015) explored this evolution from the perspective of the theoretical implications of the designs, the current article focuses more

ASSUMPTIONS	REPRESENTATIONS AND CONTROLS
Original Grant Proposal (2007): Overlay formal representations on popular game mechanics to highlight the formal physics relationships inherent in the popular game mechanics.	**Representations:** 3D "marble world" phenomenological representations based on popular games from the genre with the overlay of vector formal representations. **Controls:** Clicking arrow keys on the keyboard to apply constant force, which is a popular control mechanic in marble games.
***SURGE Classic* (2008):** Adjust popular game mechanics to highlight and clarify formal representations.	**Representations:** Develop "marble world" in a 3D engine but lock motion to 2D to clarify and highlight the vector formal representation overlays as well as to simplify the physics. **Controls:** Clicking arrow keys to apply constant forces or impulses to allow learning progressions that highlight formal relationships.

TABLE 1. Design assumptions driving the original *SURGE Classic* grant proposal and how those design assumptions shifted during iterative cycles of development and research.

pragmatically on the evolution of the phenomenological representations, formal representations, control schemes, and relationships between them in those designs. Essentially, whereas Clark, Sengupta, et al. (2015) focused on a shift in theoretical implications, the current article focuses on the evolution of the pragmatic design details and thinking that drove those designs and arguments. We conduct this design analysis based on a review of our grant proposals, annual reports, and internal and external design documents across the three grants and seven years of our design work.

FIRST STEPS: *SURGE CLASSIC* DESIGN AND RATIONALE

SURGE Classic takes places in an outer-space environment (i.e., there is no gravitational force, friction, fluid resistance, or heat loss). Students use the arrow keys on their keyboards to navigate their rocket-powered spaceship (occupied by avatar "Surge") around barriers and through corridors in search of non-player characters in need of rescue (see Figure 2). We created *SURGE Classic* in the Unity game engine (www.unity3d.com) so that we could generate a polished environment faithful to other current recreational marble games of the genre (as evidenced in the top left image of Figure 1) while still providing flexibility for deployment in classrooms across platforms without requiring installation.

Overlaid on the screen are different read-outs of information for players, including their ship's current speed, the number of impulses used, the number of collisions with the walls, and elapsed time on a given game level. Players are told to minimize their collisions, level completion time, and number of impulses in order to get a high score. There are also on-screen buttons used to reset or pause the level and to stabilize the player's ship if it starts moving out of control. A vector representation of players' velocity is also on the screen, showing their current speed and direction. Some levels include a Motion Map Region (highlighted in the upper and middle left images of Figure 2), where students must maintain a constant velocity, increase their speed, or decrease their speed (a "fuzzy" tells them which one to do) in order to continue in the level. The first set of levels uses an impulse control system (left images of Figure 2), where every time the student pushes an arrow key a fixed impulse is applied to Surge's ship (represented as a white ball in the game). The second set of levels uses a constant force control system, where students can hold down an arrow key to apply a constant force in that direction (right images of Figure 2). *SURGE Classic* can be played with no installation beyond downloading to the desktop. You can download for Mac or PC from www.surgeuniverse.com by clicking on "Play SURGE!" and then clicking on the link for *SURGE Classic*. Four sets of levels are available: 2D Impulse (see Figure 2 left), 2D Constant Force (see Figure 2 right), 2D Projectile Motion, and 3D with Friction (see Figure 1 top left).

Table 1 summarizes our assumptions and approaches to the representations and controls in the proposal and then in the actual development during the grant. The table highlights how our initial heavy emphasis honored the commitments of the popular game-play mechanics and privileged their phenomenological representations and controls. As the language makes clear, the proposed design focused on maintaining high fidelity with the popular game mechanics and representations. The formal representations were intended to highlight the formal relationships inherent in the popular game mechanics without adulterating the popular game mechanics. During development and piloting, however, we immediately encountered tensions that pushed us toward representations and controls that better supported disciplinary representations and learning needs (see Figure 2). While the versions of *SURGE Classic* in Figures

1 and 2 are built in the same engine and software codebase, the visual differences and tradeoffs are quite apparent. These differences represent the shifts in design commitments from the proposal to the actual research and development phase. Most obvious is the fact that we purposely "locked" the third dimension of the 3D world in terms of movement to (a) simplify the physics that students were investigating and (b) provide a viewing perspective that facilitated making sense of the formal representations. Essentially, (a) three dimensions of motion made the physics too complicated and (b) the viewing angle perspective employed for 3D made it very challenging for players to make sense of the formal representations, particularly when we mapped the vectors directly onto the player's ship.

Thus, almost immediately our learning goals and designs came into tension with the popular game mechanics that inspired the project. We discuss the specific design thinking and experiences that led to these shifts in the sections below. We first focus the discussion in terms of anticipated design challenges and imperatives. We then focus on the unanticipated assumptions and imperatives that arose during development and research.

ANTICIPATED DESIGN ASSUMPTIONS AND IMPERATIVES: OVERLAY POPULAR GAME-PLAY MECHANICS WITH FORMAL PHYSICS REPRESENTATIONS AND CONCEPTS

As outlined in Table 1, the original proposal focused on overlaying popular game-play mechanics with key formal physics representations. This design focus remained central throughout the actual development and research during the first grant. Core ideas from recreational game design conventions included (a) supporting engagement and approachable entry (Koster, 2004; Squire, 2011), (b) situating the player with a principled stance and perspective (McGonigal, 2011), (c) providing context and identification for the player with a role and narrative (Aarseth, 2007; Gee, 2007; Pelletier, 2008; Squire, 2011), (d) monitoring and providing actionable feedback for the player (Annetta, Minogue, Holmes, & Cheng, 2009; Garris, Ahlers & Driskell, 2002; Kuo, 2007; Munz, Schumm, Wiesebrock & Allgower, 2007), and (e) using pacing and gatekeeping to guide the player through cycles of performance (Squire, 2006). In terms of popular game-play, we also worked to embed the game in a storyline and art styles with broad appeal. Based on research about girls and gaming (e.g., Kafai, 2008), we chose to focus on a rescue theme. We also chose to steer between Japanese cultures of cute anime (e.g., *Hello Kitty*) and violent anime to instead emulate the proven middle ground leveraged by early Nintendo and Sega games such as *Mario Brothers* and *Sonic* in terms of art style (given the enormous appeal of those styles across genders and ages). As part of this effort, for example, we focused on a strong female heroine who was not ethnically specific (see Figure 3).

In terms of formal representations, we focused on vector representations and dot traces. We wanted to overlay the formal representations such that using the representations would be useful and advantageous to players. Our vector display, for example, included both composite and component vectors so that players could more easily determine how many impulses or how much acceleration would be required in a given direction to achieve their goals (see Figure 2). Similarly, the dot trace representations were designed to help players visualize constant and changing velocities as part of helping them master these ideas and techniques for puzzles within the game involving "velocity detection zones" that served as keys the players used to unlock passageways. In our designs we drew on earlier work on simulations, particularly ideas

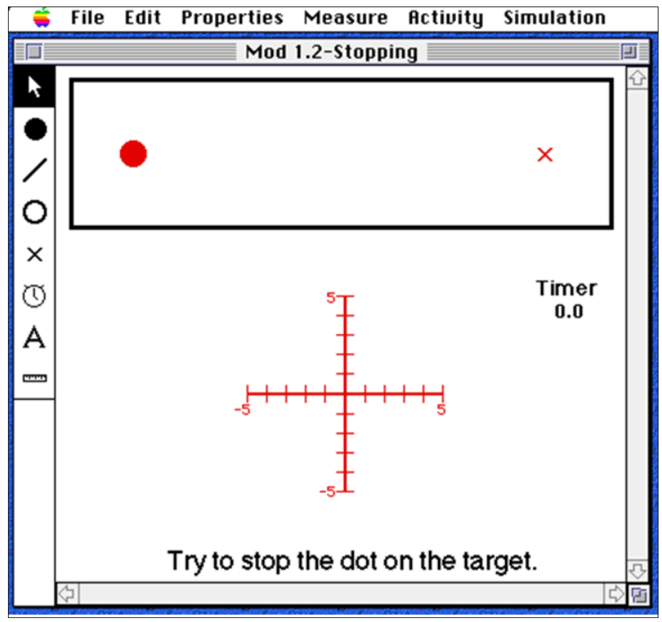


FIGURE 4. A Thinker Tools challenge (from http://thinkertools.org)

from *ThinkerTools* (e.g., White, 1993; Figure 4).

In terms of designing the formal representations into the game, we quickly realized that even relatively simple recreational games, such as our focal marble genre, tended to involve contexts that are visually more rich than those found in other multimedia formats for learning (such as simulations). We quickly observed in our pilot runs that this visual richness often confused players in terms of which aspects or details of the screen were salient and which were just environmental detail. We observed, for example, that players sometimes didn't even realize that they had key formal representations unless the representations were centered on their ship (the sphere they are moving through the game). If the representation was placed in the lower corner of the

FIGURE 5. Physics ideas and terminology were integrated into pre-level and post-level story and feedback screens.

screen, for example, players sometimes didn't notice it at all. In one study where players in each classroom were randomly assigned to versions of the game that included the vector representation either in the corner of the screen or centered on their sphere, a player who had the representation in the corner pointed to the screen of his neighbor (whose vector representation was centered on the sphere) and asked why he didn't get a "speed representation" – When we pointed out his representation, he said, "Wow! I can't believe I didn't see it!" Thus, we found that careful application of multimedia principles to signal and cue attention, such as the contiguity principle, may be even more important in game design than in the design of other multimedia formats for learning.

To support the formal representations, we worked to design each level to focus on specific challenges directly linked to physics concepts. To complete these challenges, students needed to learn and apply many principles related to mechanics (e.g., impulse, inertia, vector addition, elastic collision, gravity, velocity, acceleration, free-fall, mass, force, projectile motion). We focused on what we termed "conceptually-integrated games" for learning (Clark & Martinez-Garza, 2012), rather than "conceptually-embedded games." The science to be learned in conceptually-integrated games is integrated directly into the mechanics of navigating through the game world. In levels with the "velocity detection zones," for example, the puzzle involved figuring out what constant velocity entails in order to maneuver safely through the level and unlock passageways. In order to navigate through the level, a player needed to understand characteristics of constant and changing velocities. Other levels focused their challenges on combining vector components. All of these challenges, however, were enacted through the player's navigation through the game world.

Building on these ideas, we tried to design each level to highlight one or two topics and to connect concepts together across levels so that students can notice the relations that exist among the topics. For example, in the multiple dimensional motion levels, students learn and apply the concept of applying impulses at right angles to produce motion in two dimensions. This builds on students' knowledge of additive and canceling impulses and motion in one dimension, and extends that knowledge to include the resultant motion of impulses at right angles. Our goal was to help students to gain a firm grasp of a concept before new concepts were introduced.

We also worked to integrate physics ideas and terminology into pre-level and post-level story and feedback screens (see Figure 5) and within the phenomenological representations of the game itself. Several levels, for example, include "detection" corridors where the player needs to maintain a constant velocity, increase her velocity, or decrease her velocity in order to open a velocity gate (e.g., the left images in Figure 1). This specific terminology is explicitly delivered as instructions by a friendly non-player character (a "fuzzy") and is central to succeeding in the level.

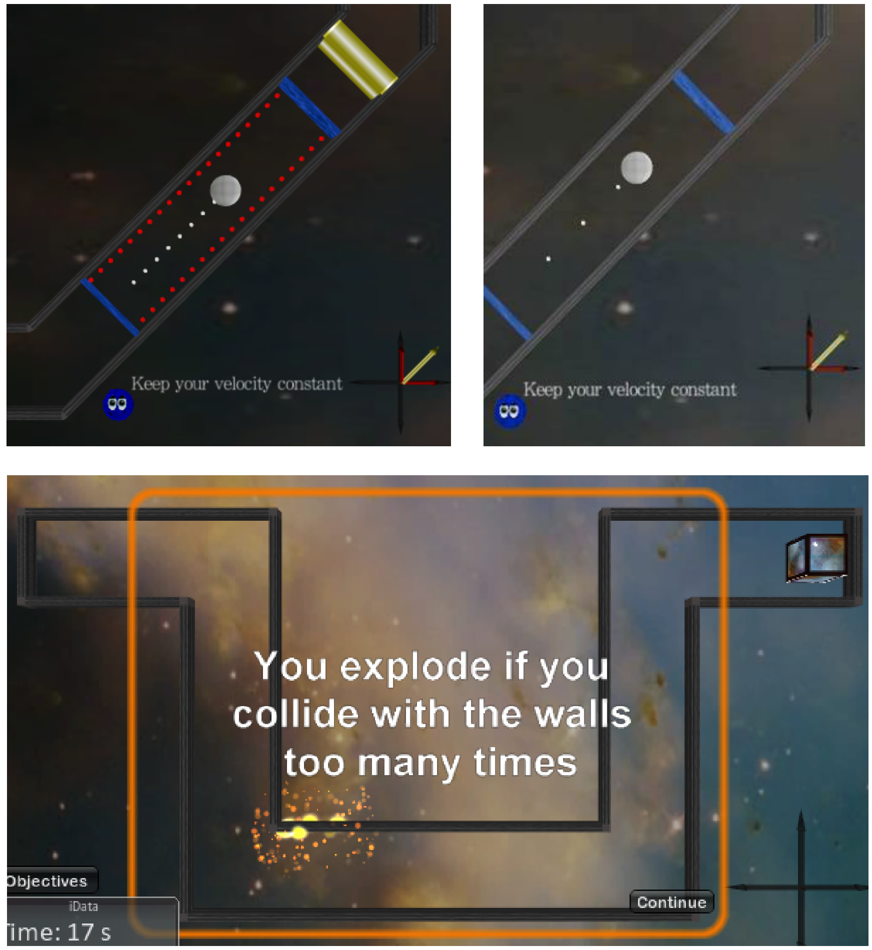


FIGURE 6. Adding gates to velocity zones and explosions for excessive collisions.

Unanticipated Design Challenges and Imperatives

The previous sections discuss the design challenges we were expecting to explore in the proposal in terms of overlaying popular game-play mechanics with formal representations to highlight the underlying physics. In addition, however, we also encountered unanticipated design challenges and imperatives resulting from the shift from voluntary out-of-school participation to non-voluntary in-school participation. More specifically, players in recreational out-of-school contexts self-select into recreational games that are appropriate to their skills and interests. Games designed for the classroom, however, ethically and pragmatically need to support all students as players. We had not considered this tension in the proposal, and these tensions led to unanticipated design realizations, challenges, and imperatives. Each of these pragmatic design discoveries in turn influenced the design of the popular and formal mechanics and representations during development in *SURGE Classic*.

Protecting novice players from frustration cannot allow progress without mastery

We did not want to create a game that would be productive for players with extensive gaming experience but that was frustrating or less productive for less experienced players. At the beginning of *SURGE Classic* development, we focused on minimizing frustration and scaffolding success for less experienced players. In the early versions of *SURGE Classic*, if a player could eventually reach the end of a level, the player earned at least a bronze medal for that completion. Silver medals were intended to be fairly challenging to attain, and the gold medals were intended to be very challenging. This resulted in (a) players not necessarily learning what they were supposed to learn in the levels and (b) progressing into levels that presumed that players had already developed some of that understanding. Players could thus reach levels for which they were not prepared to learn or succeed. We struggled with this tension between protecting students from failure while also requiring certain levels of mastery for advancement. As an example, velocity detection zones in early versions of *SURGE Classic* affected score but did not have physics gates attached to them (see Figure 6 top left). Later versions added a gate that players needed to solve and unlock with the velocity detection zones before progressing further in the level (see Figure 6 top right). Similarly, levels did not initially include fixed failure triggers forcing the level to reset (e.g., our initial versions of *SURGE Classic* did not include the possibility of Surge's ship exploding after a set number of collisions that was added later as shown in bottom image in Figure 6). We thus realized early in *SURGE Classic* that protecting players from undue frustration cannot allow progress to subsequent levels without ensuring a certain baseline of mastery. Interestingly, and perhaps not surprisingly, players greatly preferred and enjoyed the versions that added these new constraints and challenges.

Keep people from falling off with "just in time" support

We also worked to increase "just-in-time" support in *SURGE Classic*. Our initial plans involved including scaffolding for physics ideas and terminology in the scenes before and after levels. We soon learned however, that even though we were working to minimize the length of such supports, players tended to skip right over them. Only when players encountered difficulty and were told that the scenes between levels contained helpful hints did some players start to read them, but generally the cut-scenes were not effective because they did not provide the help when it was most needed. We then began to increase emphasis on just-in-time supports. Figure 6 shows how messages are provided to players when they collide with the walls too many times. Another revision involved moving gates discussed in previous section to be placed directly after each detection zone (as shown in Figure 6) rather than having a single gate at the end of the level in front of the final fuzzy to be rescued. By placing the gates immediately after each detection zone, players received concrete and immediate feedback on their performance. Furthermore, players received messages for failed attempts in detection zones using the same pop-up format shown in Figure 6 for the explosions. These text-based just-in-time hints provide a sequence of hints and tips relevant to the specific problem encountered. We also added other visual signals, such as the red dots in the right screenshot of Figure 6 that model the spacing of the dot trace to be created by the player while also signaling the fact that the detection zone is different from other areas of a level. We worked to refine these components in *SURGE Next* and *Fuzzy Chronicles*.

Supporting a broad challenge curve

The previous two design considerations are part of a larger imperative of supporting a broad challenge curve. This is useful even in games intended for voluntary out-of-school contexts where students self-select, but we realized through our work on *SURGE Classic* that supporting a broad challenge curve becomes imperative for games intended for classroom use. When a student is appropriately challenged, that student will have an increased likelihood of learning and being engaged (see Figure 7). When the challenge curve is beyond a student's current capability, the student is likely to be frustrated rather than engaged, and that student is unlikely to learn. When the challenge curve is behind the current capabilities of a student, the student is likely to be bored rather than engaged, and the student is also unlikely to learn efficiently. Rather than focusing on where to center the challenge curve for a game, therefore, it is much more important to expand the curve, or range, of challenge (see Figure 7). Toward these ends, during *SURGE Classic*, we thought about expanding the challenge curve in multiple ways based on examples from various recreational game genres and exemplars. Some of these goals we managed to explore in *SURGE Classic*, but others became part of the design thinking for *SURGE Next* and *Fuzzy Chronicles*: (a) minimize costs of failure and experimentation, (b) provide multiple paths or solutions of varying difficulty and reward, (c) increase challenge correlated to performance, and (d) encourage improved performance through non-game-mechanic influencing incentives.

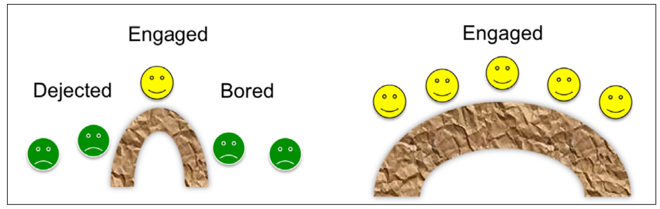


FIGURE 7. Expanding the challenge curve rather than simply re-centering the challenge curve as conceptualized in a 2011 keynote (Clark, 2011).

Minimizing costs of failure and experimentation. Minimizing costs of failure and experimentation is a core idea to popular game design, but counter to traditional structures in schools. Placing a high cost on not achieving or succeeding on the first attempt not only hurts engagement, but also dramatically truncates the challenge curve. This approach does not, however, over-rule the previous design imperative of "protecting novice players from frustration cannot allow progress without mastery." There must be a core level of performance requisite for progression to subsequent levels of a game, but players should be encouraged to experiment and try new approaches and consider different interpretations as they play to support learning of new approaches and skills. If the cost of failure in the game is high, players will be unlikely to experiment and learn. We did not get this balance right in the early versions of *SURGE Classic*, but we made progress in our subsequent *SURGE Classic* designs. By the conclusion of our *SURGE Classic* work, one area with which we still struggled was the length of our levels. As a result, exploding near the end of a level meant that players needed to start over at the very beginning of the level. This imposed a higher cost for experimenting than we wanted. This thinking contributed substantially to the focus in *SURGE Next* on shorter puzzle-like levels that were more amenable to experimentation and failure. These levels allowed the player to immediately reattempt more rapidly with lower cost in terms of needing to redo extensive parts of a level that they have already successfully negotiated.

Providing multiple paths or solutions of varying difficulty and reward. SURGE Classic made minimal progress in terms of providing multiple paths or solutions of varying difficulty and reward because working in the game engine created development bottlenecks. Essentially, the programming team needed to create and revise all level designs proposed or desired by the research team members. As a result we were limited in our ability to expand and revise *SURGE Classic*. As a result, all players followed the same path (with performance being differentiated based on efficiency in terms of time, fuel, and collisions). This resulted in a narrower challenge curve than we wanted. In the designs for *SURGE Next*, we worked to design level structures that are more open-ended and to provide a range of possible solutions that vary in challenge and also vary correspondingly in reward. All solutions were designed to require the base level of mastery we deem critical for progression, but harder solutions (with greater rewards) explored more nuanced or complex applications of the core conceptual ideas at the heart of the level. Our ultimate goal was to allow players to choose a level of challenge appropriate to their current skills to effectively broaden the challenge curve of each level.

Increasing challenge in correlation with performance. We were not able in *SURGE Classic* to increase challenge in correlation with performance. We returned to this idea in *Fuzzy Chronicles* in terms of developing a computer-adaptive-testing type of engine that selected variants of a level based on a player's prior performance on the level. In this way, as we describe later, we tried to flexibly adjust the challenge curve to each specific player at each specific time, effectively creating a much broader challenge curve for all players.

Encouraging improved performance through non-game-mechanic influencing incentives. We were also not able in *SURGE Classic* to encourage improved performance through non-game-mechanic influencing incentives beyond the medals that were awarded. A common mechanic in games is to encourage players to replay levels, analyze their prior strategies, and reflect on their playing by offering rewards for improved performance. The difficulty from the perspective of supporting a broad challenge curve is that these rewards can effectively narrow the challenge curve on subsequent levels if the rewards enhance the capabilities of the player within the game. Essentially, if you reward a player for performing well on one level with capabilities that make performing well on subsequent levels easier, you have pushed that player ahead, and possibly off, of the challenge curve on those subsequent levels, effectively narrowing it. Instead, rewards can focus on non-game-mechanic incentives, such as the appearance of the player's avatar, medals, and other visual representations of their progress within the game world. Players in commercial games go to great lengths to collect these types of rewards both for their own enjoyment and as representations to other players of their progress. Another approach is to unlock tools or aspects of the storyline that will allow the player to explore new areas or aspects of the game, but these rewards are generally better connected to the minimal levels of performance required for progression so that all players will have access to them. As part of *Fuzzy Chronicles* and *SURGE Next*, as we will discuss later, we created a more flexible authoring system for the "star maps" that players used to choose levels. The star map allowed us (a) to provide more visually rewarding and satisfying progress indicators for players as they progressed in terms of expanding the representation of their base, followers, and awards and (b) to create non-linear chains of levels.

Reflections on *SURGE Classic*

Ultimately, students playing versions of *SURGE Classic* demonstrated high engagement and significant learning gains on items based on the highly-regarded Force Concept Inventory (FCI), which is a widely known benchmark assessment for conceptual understanding of Newtonian dynamics at the undergraduate level (Hestenes & Halloun, 1995; Hestenes, Wells, & Swackhamer, 1992). In terms of these popular game-play design goals and efforts, we were relatively successful. Girls and boys in the U.S. and Taiwan enjoyed the game and there were similarities across genders and countries in terms of affective and learning outcomes (Clark et al., 2011). The downside, however, was that these gains and increasing mastery focused on intuitive understanding (which is what the FCI largely measures) rather than explicit understanding. Essentially, players could more accurately predict the results of various actions, impulses, and interactions (which improves performance in the game and on FCI questions), but players were not being supported in explicitly articulating their mental models and the connections from choices made in game-play to formal disciplinary representations and concepts (Clark et al., 2011).

We interpreted these results as demonstrating that the players were developing intuitive rather than formal understandings while playing a game built mainly on commercial design approaches. This seemed sensible because the goal of commercial games involves helping players develop robust intuitive understanding that helps them enjoy increasing levels of mastery as they play the game, which naturally increases their engagement and desire to play more. If players are left confused and unable to learn to play the game, or if the learning process is overwhelming or poorly structured, players will disengage, making it very unlikely that they will recommend the game to others or purchase future versions of the game. Repeated designs of this type would naturally drive a game company into bankruptcy. Thus, strong evolutionary pressures in the gaming industry favor design conventions that support intuitive understanding. There is no immediate market need, however, for commercial games to support explicit articulation or connection to formal ideas. The intuitive understandings developed at the heart of commercial games generally are not intended to correspond with important understandings outside of those games. We took these thoughts and assumptions forward into the planning process for our next grants along with all of the pragmatic imperatives and lessons we had learned in terms of supporting the full range of students as part of a formal classroom setting.

SURGE NEXT AND *FUZZY CHRONICLES*: ARTICULATION, PREDICTION, EXPLANATION, AND DATA MINING

We entered into *SURGE Next* and *Fuzzy Chronicles* with these refined design considerations. Most foundationally, our experiences with *SURGE Classic* led us to focus on approaches for supporting students in consciously considering and articulating the intuitive understandings and resources that they were developing through game-play. Research in psychology, science education, and the learning sciences suggests a number of ways to support explicit articulation and reflection, but it quickly became clear to us that the design approaches developed through that research focus on contexts and mediums with different characteristics, affordances, and constraints than those of digital games. As a result, in order to be synergistic rather than disruptive, we found that these design approaches require adaptation and reinterpretation for the digital game medium. Subsequent *SURGE* grants focused on leveraging explicit articulation in synergy with recreational game design conventions. More specifically, these areas of research involved enhancing prediction within navigation interfaces and self-explanation within game dialog (Adams & Clark, 2014; Clark, Martinez-Garza, Biswas, Luecht, & Sengupta, 2012; Killingsworth, Clark & Adams, 2015). In addition, the grants also focused on approaches to data mining student game-play (Clark & Martinez-Garza, 2012; Clark et al., 2012; Kinnebrew et al., in press; Martinez-Garza, Clark, & Nelson, 2013). These aspects of the research and development are discussed in extensive detail in the referenced publications and others. These were (and are) major foci of the design development work, but for the purposes of this paper, we will maintain our focus on the implications and evolution of the representations and control schemes that evolved in support of these goals (see Table 2).

SURGE Next and *Fuzzy Chronicles*, players still navigate their avatar through the play area to collect fuzzies and deliver them to safe locations while avoiding obstacles and enemies (as in *SURGE Classic*). Rather than employing the real-time interfaces of the original *SURGE Classic* grant (where pressing an "arrow key" resulted in immediate application of an impulse or constant thrust in the direction of the arrow key), *SURGE Next* and *Fuzzy Chronicles* incentivized prediction by requiring the player to place all of the commands in advance. After players have placed and adjusted commands to create a plan to their satisfaction, players click a "launch" button to launch the plan. Players can then revise and re-launch as they experiment with new approaches to a given level. The goal was to require the player to make predictions about the results of each command in terms of the motion of the player's avatar rather than simply interacting with the game reactively.

SURGE Next focuses on spatial placement, where commands are placed directly on the play map. (see Figures 8a and 8b). The spatially placed commands are executed if and when the player's ship crosses the commands on the map. *Fuzzy Chronicles* began with a temporal approach to command placement (see Figure 9a). In this approach, commands are placed in a timeline bar. Each command is executed at the specified time independent of the location of the player's ship. Ultimately this temporal approach worked well for some players but was too complex for other players (~20% of students in a typical middle school classroom). We therefore redesigned *Fuzzy Chronicles* to employ spatial commands in the third year of that grant (see Figure 9b).

In addition to shifting control approaches, *SURGE Next* and *Fuzzy Chronicles* also reduced the total number of commands a player initiates in a given level (thereby increasing the salience and impact of each individual command) to (a) encourage players to think more carefully about the

ASSUMPTIONS	REPRESENTATIONS AND CONTROLS
***SURGE Next* (2011)** and ***Fuzzy Chronicles* (2012):** Support articulation and reflection through prediction and explanation.	**Representations:** 2D focus for phenomenological representations and formal vector representation overlays. Create smaller tighter levels as puzzles involving fewer total actions to incentivize decisions about each action. Explanation functionality integrated to support articulation. **Controls:** Place commands in advance on a timeline or directly on the phenomenological representation. Then launch plan. Pre-placing command intended to incentivize prediction.
***SURGE NextG* (2012):** Make formal representations more salient to game-play.	**Representations:** Maintain 2D focus for phenomenological representations and formal representation overlays, but also include Cartesian graphs of position and velocity integrated into activities for key levels. **Controls:** Same as SURGE Next.

TABLE 2. SURGE Next, Fuzzy Chronicles, and SURGE NextG design assumptions and design approaches to representations and controls.

outcomes and implications of each action and (b) minimize the costs of experimentation and failure as discussed earlier for *SURGE Classic* lessons. Game levels thus became more puzzle-like. We selected an independent game called *Gravitee 2* as our model for aesthetics and puzzle-like play (although the physics elements of *Gravitee 2* puzzles differed from the puzzles we envisioned) and we hired the developer of *Gravitee 2* (http://www.funkypear.com) to work with us on the project (see Figure 10).

We encountered similar tough decisions and tensions between our research goals for data mining and the design of the phenomenological and formal representations. Successful data mining also depends heavily on being able to infer students' goals from their actions. Inferring players' goals and intuitions from their actions becomes exponentially more difficult as level design becomes more open-ended. For example, if you know that a player is trying to make a 90 degree turn because the player is travelling down a corridor that bends 90 degrees, then it is easy to make inferences about the player's goal (turn 90 degrees at that location) and it is therefore easy to infer the player's intuitions based on his or her actions (as well as to provide explanation dialog appropriate to the player's intuitions). In an open-ended game level, however, it is much more difficult to be certain of a player's goals, and therefore it is much more difficult to infer the player's intuitions from a data log of actions. The images in Figure 11a are slides from the planning document shared with the broader research team and collaborators. They show the very open-ended levels we intended. The images in Figure 11b show the markup for data collection that we built into each level. These levels are much more constrained than the levels in the planning document as a result. These tensions thus undermined our intentions of providing multiple paths or solutions of varying difficulty and reward that we discussed as a lesson learned from *SURGE Classic*.

Designing for data mining also involves even further constraints in terms of level design across the levels of a game. We initially approached level sequences for *SURGE Next* and *Fuzzy Chronicles* in terms of thinking about optimal learning progressions. These learning progressions focused on designing and optimizing appropriate sequences of levels for students to refine and explore their intuitions. Each level introduced one or two new focal ideas. Data mining and comparing students' game-play actions across levels, however, are complicated by this approach because levels do not necessarily provide equivalent comparison points (because each level introduces new focal ideas or twists that subtly revise the challenges).

Essentially, comparing across the levels that result from a focus on optimizing the learning progression unintentionally resulted in "apples and oranges" comparison challenges from a data mining perspective when trying to trace the evolution of a student's thinking and game-play. We then developed an optimized data mining progression, where specific navigational maneuvers were systematically and consistently included across levels to provide comparable comparison points, but this resulted in bloating the levels and losing the focus on the learning progression.

We then created a catalog of maneuvers with associated rubrics for coding specific successful and unsuccessful combinations of actions that a student might apply in terms of inferring the student's underlying intuitions based on Minstrell's facet-based approach to analyzing student thinking about physics (Minstrell, 2000; Minstrell, Anderson, & Li, in press). This approach builds on a knowledge-in-pieces theoretical account of student learning (diSessa, 1993; diSessa & Minstrell, 1998). A facet-based account of student learning assumes that students bring a large number of resources and ideas with them based on everyday experiences as well as formal learning experiences. The specific ideas that students apply in a given context are cued by the particulars of the context. Learning from this perspective involves students refining these specific ideas (or facets) in terms of how they are combined, cued, and applied. For *SURGE Next* and *Fuzzy Chronicles* we carefully sequenced these maneuvers in a manner that kept levels parsimonious and "tight" but provided points of comparison across levels. While this sequence of levels was exciting from a data mining perspective, it unfortunately also lost the learning-progression optimization we had worked so hard to develop.

The major influences and evolution of the representations and control schemes were driven by these emphases on incentivizing prediction and reflection through pre-placement of commands and reducing total number of commands to encourage players to think more carefully about the outcomes and implications of each action. These shifts, however, represented only a portion of the goals driving development. As mentioned in the introduction, the grants focused heavily on developing dialog functionality to drive adaptive self-explanation (see Figures 12a, 12b, and 13) and data mining to assess players' game-play (see Figure 11b). These emphases are described and discussed in detail for *SURGE Next* and *Fuzzy Chronicles* in other articles and chapters (as referenced above). For the purposes of this article, we discuss only the implications of these foci for the design of the representations and controls. More specifically, designing for the integration of explanation and data mining functionality involved challenging design trade-offs in terms of the representations, controls, and overall level design. Essentially, designing for explanation functionality and data mining functionality pushed design away from open-ended game-play toward more constrained game-play. Constraining game levels supported our ability to interpret the meaning of specific player actions and to provide players with explanation dialog experiences appropriate to their actions and decisions in the game. Thus open-ended game-play, and even learning goals, were compromised at times by the research imperatives of those grants.

This was most apparent in the evolution of *Fuzzy Chronicles*. Early on, a tough decision was made during negotiations with the game company who worked with us to develop *Fuzzy Chronicles*. Essentially, the developers thought that if *Fuzzy Chronicles* employed instantaneous impulses only, and never employed continuous forces or friction, we could keep the player's ship on a grid at each second. By adopting this grid focus, the developers proposed that the software could be programmed to diagnose the source of any ship crash, thus allowing the software to prompt an explanation dialog appropriate to the cause of the crash. We worked with the developer on a preliminary rubric/decision tree for the diagnosis, and we agreed that it seemed promising and exciting. This was a tough decision because helping students understand Newtonian dynamics depends substantially on being able to engage students with continuous forces and friction. We therefore had to weigh that tradeoff against the advantages and opportunities for research provided by the proposed explanation functionality. We decided to accept the trade-off.

Unfortunately, creating the explanation functionality proved more challenging and elusive than expected in terms of matching causes and explanation dialog to crashes in real time. Furthermore, it turned out that even if the real-time diagnostic software were flawless, the system would still be problematic because players often design only the initial portion of a plan, run it, and then iteratively add more components to the plan. Thus, the portion of a plan that a player was focusing upon might work flawlessly, but the player would still receive explanation dialog for the crash that occurred after the focal segment. Essentially, players were receiving explanation dialog to help them "understand" crashes that occurred only because the players weren't working on that part of the puzzle yet – often for reasons they already understood. Thus (a) the explanation diagnostics were less accurate than anticipated and (b) the explanation dialog was often undesired by the players.

Based on these findings, we ultimately shifted to an entirely different software and interface approach for explanation that built the explanations into specific short missions where explanations were triggered upon players' successful completion of specific short navigation challenges (see Figure 13). We were stuck, however, with the grid decision and the loss of continuous forces and friction for the rest of the project. That said, while the collaboration stumbled at places, *Fuzzy Chronicles* was ultimately a polished game that contributed to our evolving understanding of developing games for science learning. Furthermore, *Fuzzy Chronicles* is currently in the process of being refined and expanded by the game company for commercialization. Therefore, while some hard lessons were learned, overall many valuable lesson were learned (as we discuss later in our discussion of our successful collaboration with the company on *SURGE Symbolic*), and an excellent game is ultimately in the process of being released commercially.

Thus, in our experience, there is a substantial tension between designing a game for optimal game-play experiences and designing it to support data mining, assessment, and self-explanation functionality. While we are continuing to work on *SURGE Next* and *Fuzzy Chronicles* as we write this manuscript, these design tensions remain challenging and central.

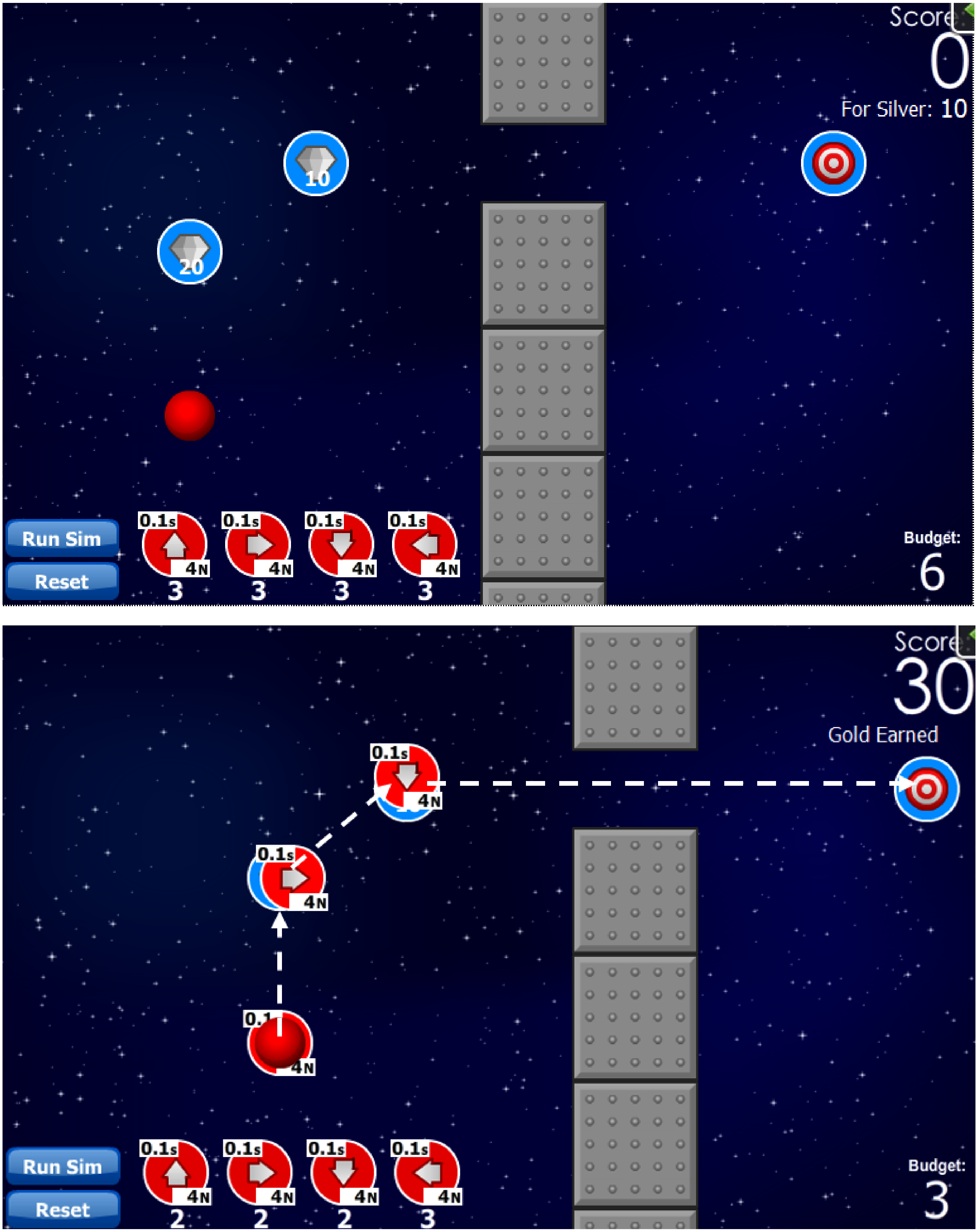


FIGURE 8A. Simple *SURGE Next* level (top) and a sophisticated solution (bottom).

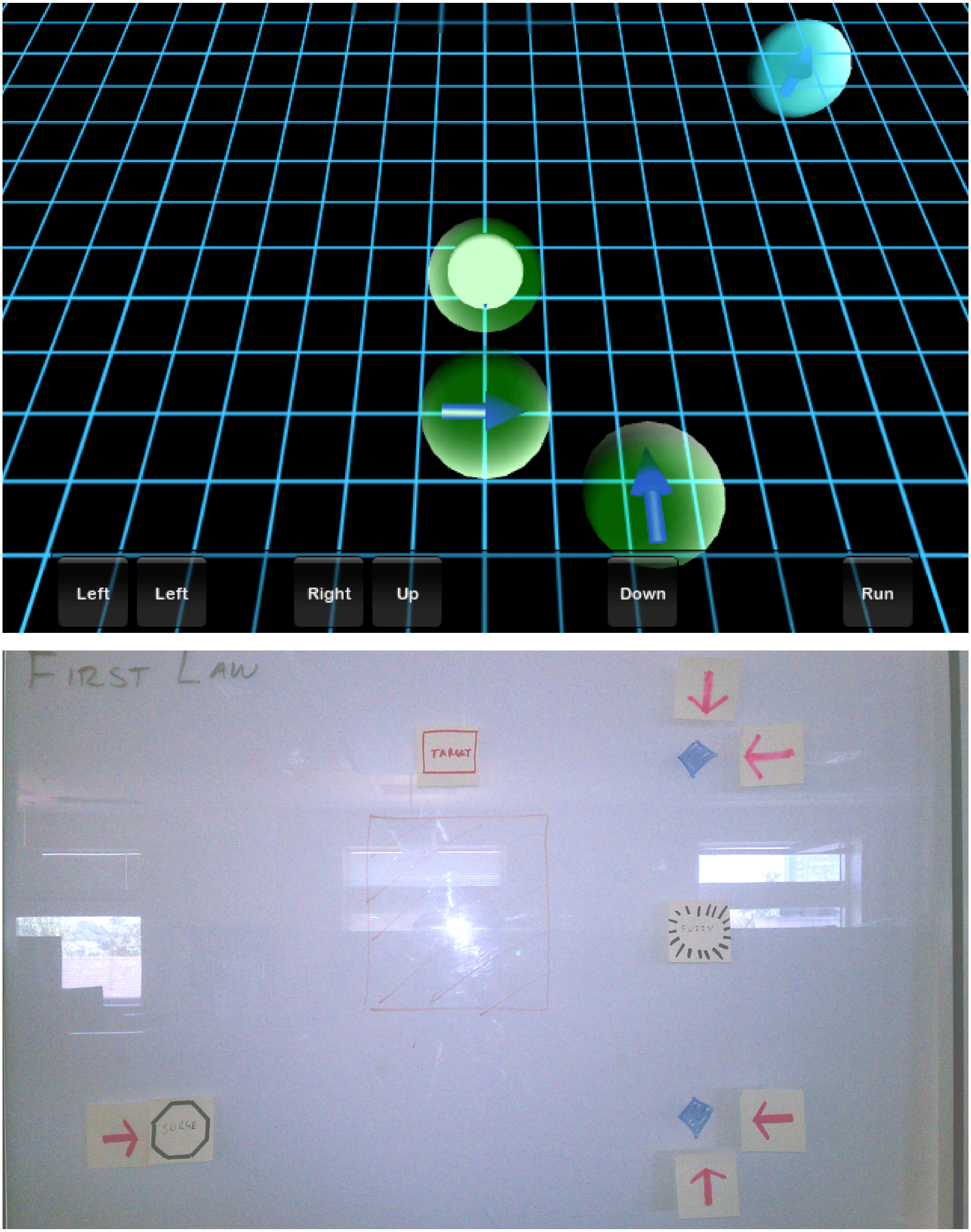


FIGURE 8B. An early prototype for SURGE Next created in Unity at the end of the SURGE Classic grant (top) and early level prototyping for SURGE Next (bottom).

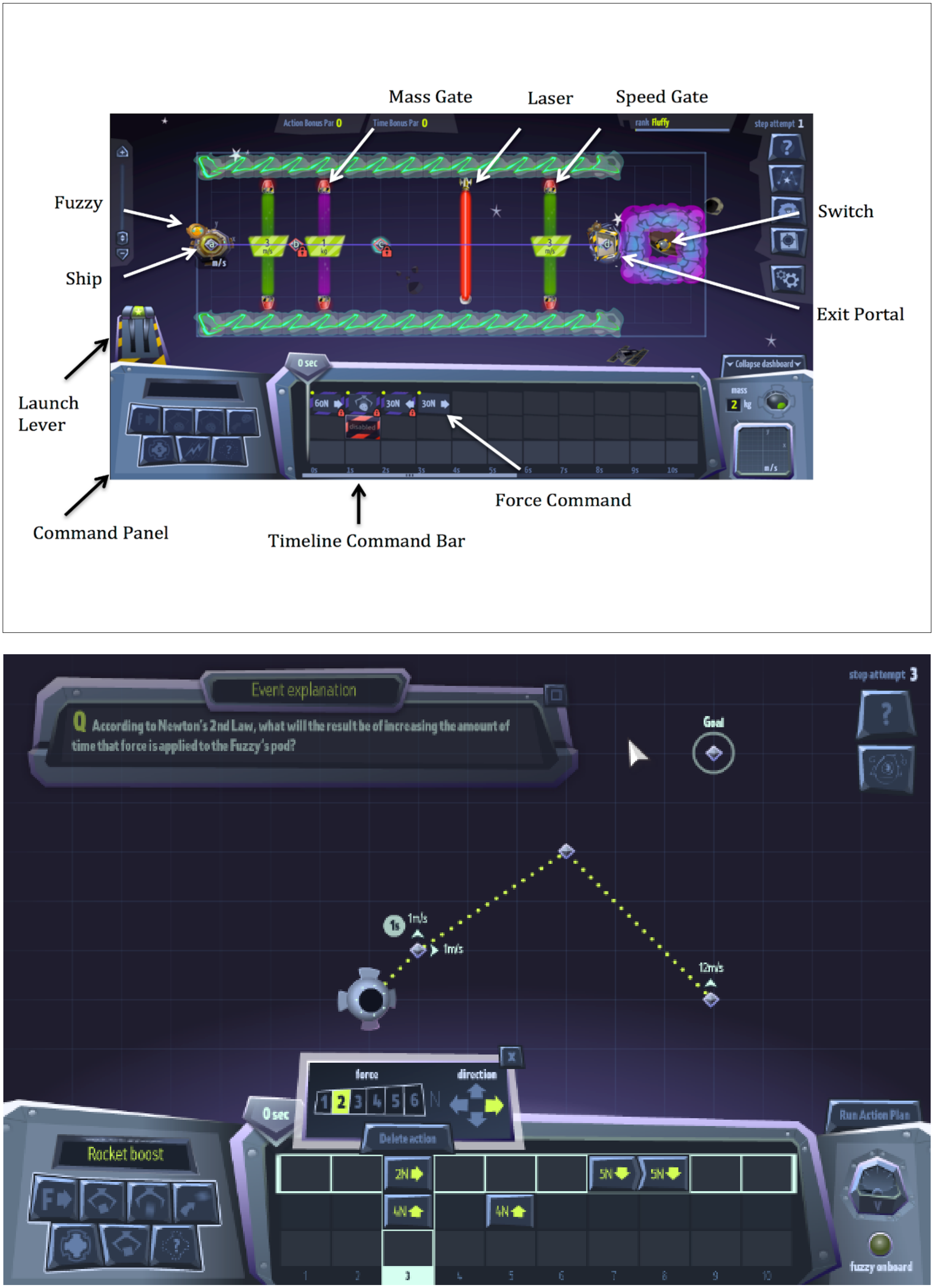


FIGURE 9A. Key parts of a Fuzzy Chronicles level with labels (left), prototype of timeline placement functionality (right).

FIGURE 9B. A Fuzzy Chronicles challenge level (top), and the revised spatial-placement control scheme (bottom).

FIGURE 10. *Gravitee 2* (top images) formed a primary inspiration for the look and feel of SURGE Next. The bottom image is a slide from a planning document shared with the broader research team and collaborators. *Gravitee 2* can be played at: http://www.kongregate.com/games/funkypear/gravitee-2

FIGURE 11A. Examples of metadata regions and their relationship to alternative solutions in *SURGE Next*.

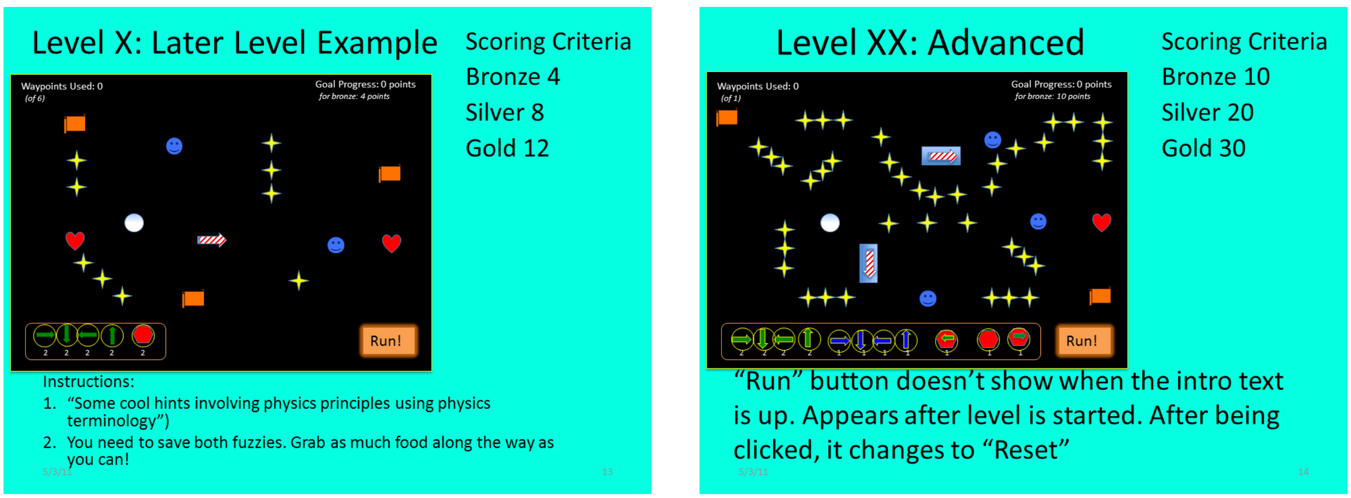


FIGURE 11B. Pages from the early planning doc that emphasize the idea of open-ended levels where players would be provided with finite numbers of available actions and be encouraged to explore different configurations of actions to maximize success.

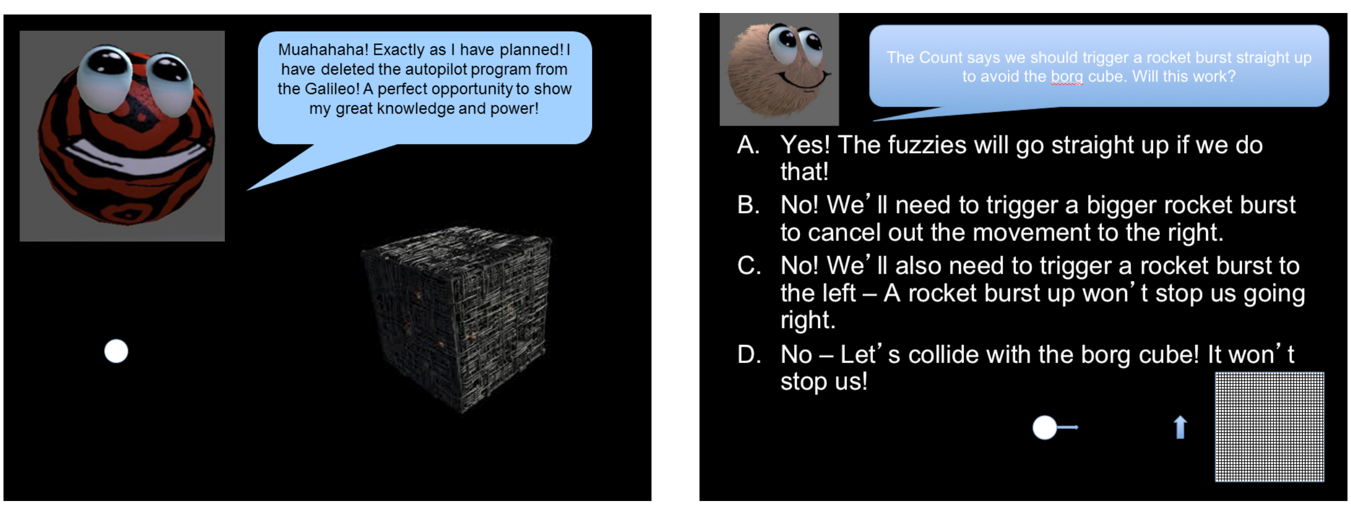


FIGURE 12A. Early prototyping for the explanation functionality. Our goal was to engage the player in self-explanation and reflection as part of the game narrative dialog in the context of helping the computer characters solve navigation challenges similar to those the player has just completed.

FIGURE 12B. Early prototyping for the explanation functionality (bottom) and the grid based explanation interface (top).

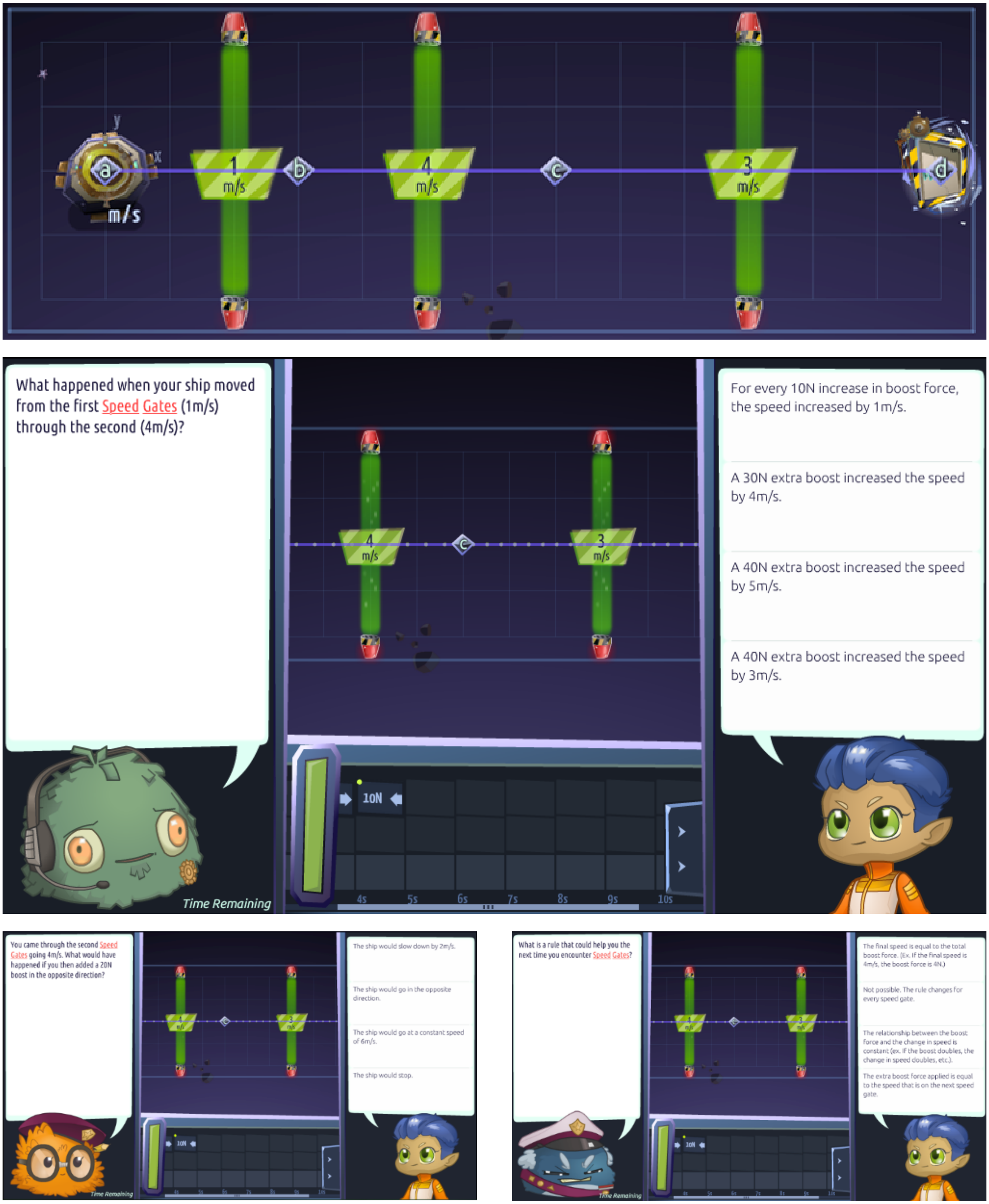


FIGURE 13. After succeeding in the navigation challenge in a warp mission, students encounter the explanation phase of the warp mission. In the full self-explanation condition, the first of the three questions asks the student to articulate the solution to the navigation challenge in a concrete manner (top), the second question asks the student to characterize the solution with a more abstract/generalizable relationship (bottom left), and the third question asks the student to articulate an even further abstracted a rule of thumb (bottom right).

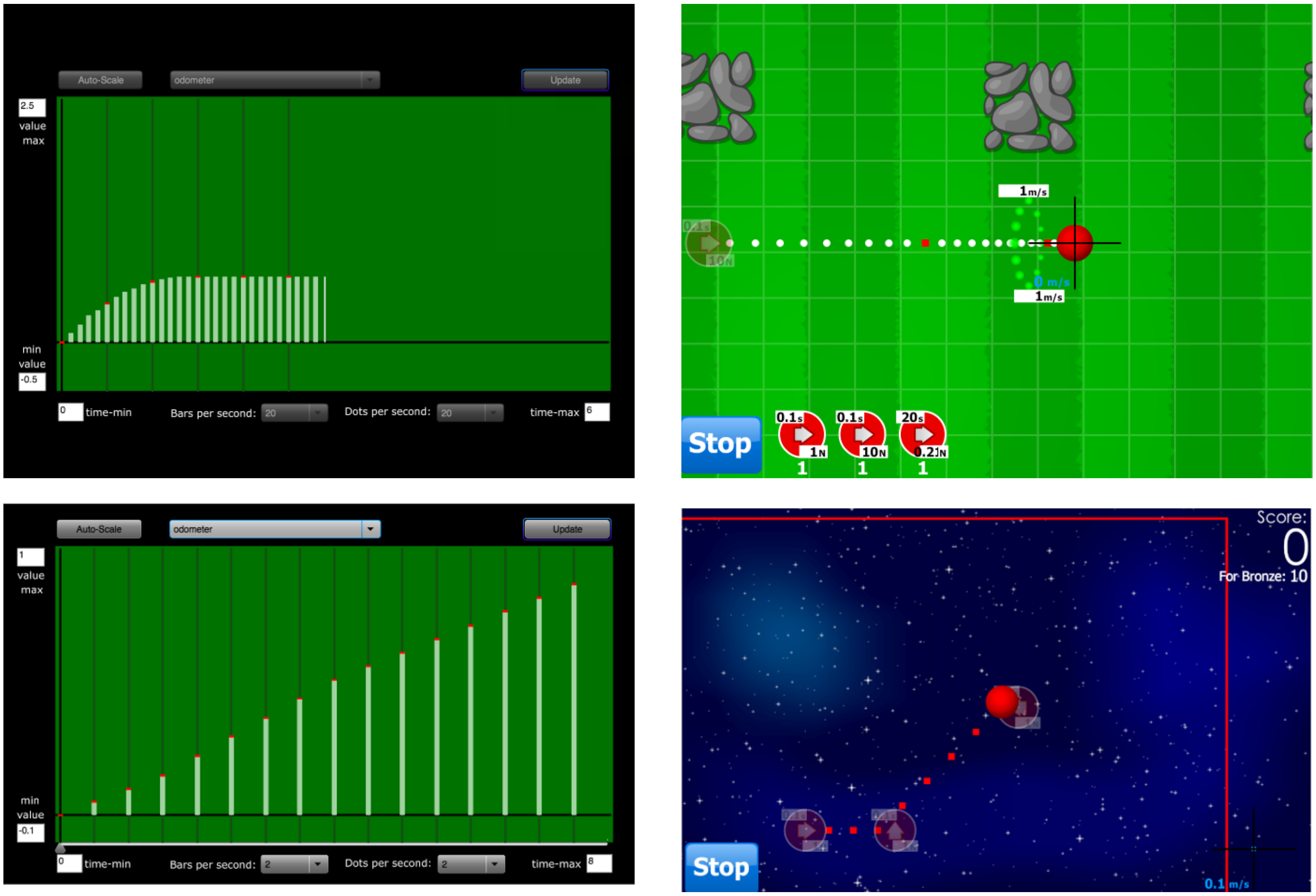


FIGURE 14. *SURGE NextG* paired Cartesian graphs with specific *SURGE Next* levels and activities to engage students explicitly in modeling to solve game challenges.

Reflections on *SURGE Next* and *Fuzzy Chronicles*: The Birth of *SURGE NextG*

The tension continued in *SURGE Next* and *Fuzzy Chronicles* between our original vision of a recreational marble-genre game as the focal representation as opposed to more primary foci on formal representations and our commitments to explanation functionality and data mining. The same was true for the evolution of the control schemes. In terms of learning, *SURGE Next* and *Fuzzy Chronicles* were an iterative improvement over *SURGE Classic,* but still did not support the level and depth of learning or integration of intuitive and formal understandings that had driven our proposals. We thus began to explore approaches that highlighted formal representations more explicitly through *SURGE NextG*, which explicitly incorporated Cartesian graphs into specific levels within the activity structure. This work began during the second academic year of the *SURGE Next* grant (2012-2013) and is written about in detail (Clark, Sengupta, et al. (2015); Krinks, Sengupta, & Clark, 2013; Krinks, Sengupta, & Clark, 2015; Sengupta & Clark, in press). We therefore will not report all of the details again here, but we will focus on the assumptions behind the design of the representations for the purposes of this manuscript (see Table 2 second row).

SURGE NextG represented important progress in encouraging students to reflect and articulate and to integrate their intuitive and formal understandings. More specifically, *SURGE NextG* focused on building explicit connections between the phenomenological representations and formal representations, and these explicit connections proved powerful in supporting student learning, particularly when embedded within a curriculum augmented by other digital and physical modeling opportunities in connection with the game (Krinks et al., 2013; Krinks et al., 2015; Sengupta & Clark, in press). Essentially, the students were challenged to create certain graph shapes with their ship in a level, and then in a later level they received a graph in a communication from the fuzzies as instructions for how to solve an upcoming level challenge (see Figure 14). While the integration of the Cartesian graphs was somewhat rough, we were excited by the opportunities provided in terms of connections from the game into the broader activities and materiality of the classroom and the teacher's sense of modeling and pedagogical practices (Sengupta & Clark, in press).

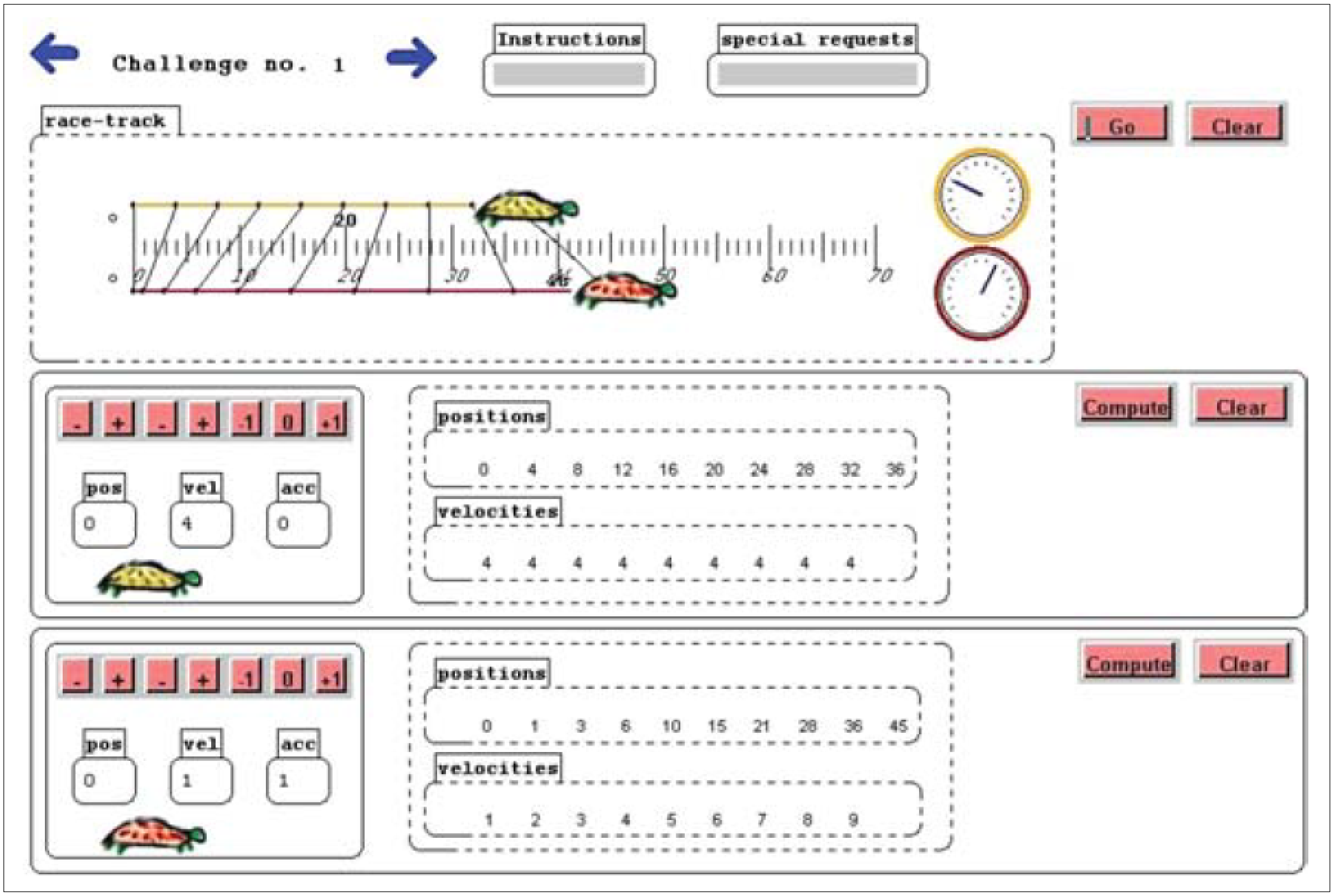


FIGURE 15. The original inspiration for *SURGE Symbolic* was *NumberSpeed* (Parnafes & diSessa, 2004). This figure is adapted from that article with permission from the authors and editor. Players set starting position, velocity, and acceleration for the two turtles to create specified patterns of relative motion.

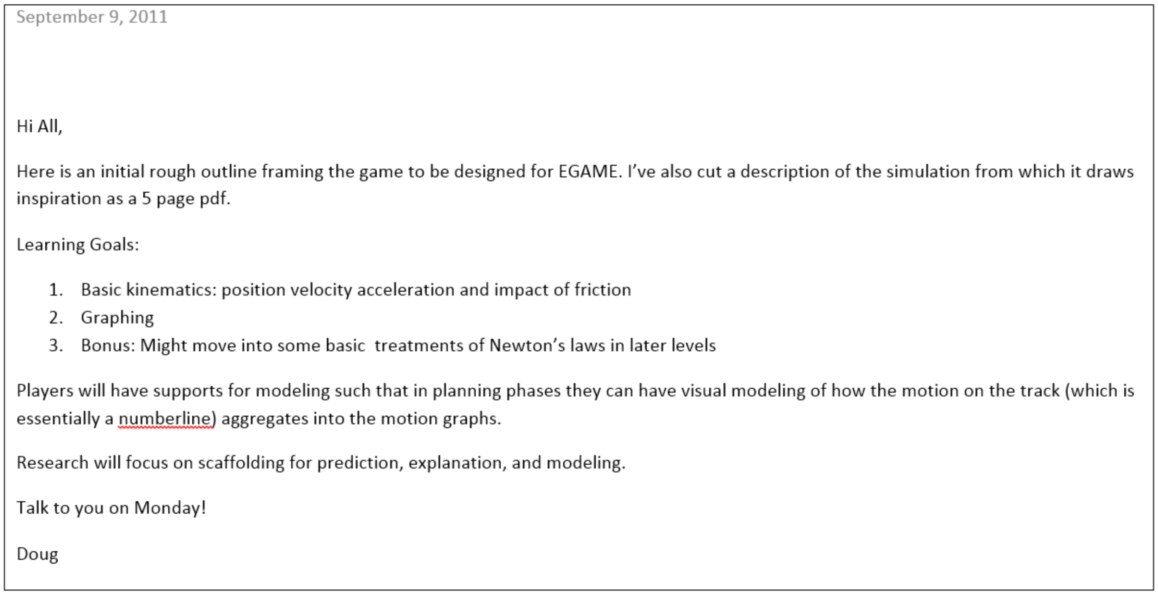
SURGE SYMBOLIC

SURGE Symbolic was developed as part of the same grant that funded *SURGE Next*. At the beginning of the grant in 2011, the initial plan for *SURGE Symbolic* Phase 1 was inspired by Parnafes and diSessa's *NumberSpeed* (see Figure 15 from Parnafes & diSessa, 2004). The original design of NumberSpeed was by Andy diSessa, with subsequent implementations and improvements by Steve Adams, Rodrigo Madanes, Andy diSessa, and Orit Parnafes. In *NumberSpeed*, players specify the starting acceleration, velocity, and position of two turtles. Players try to find sets of values for these parameters that will cause the turtles to move in specific sequences of motion relative to one another (e.g., a challenge might involve creating a set of values such that the top turtle starts in front, then the bottom turtle moves ahead, and then finally the top turtle moves ahead).

In our vision of *SURGE Symbolic* Phase 1, players would program two fuzzies and two enemy robots in various combinations such that the fuzzies would always be in safe locations relative to the robots while at the same time rescuing friends from specific locations. Figure 16 presents pages from an early planning document for Phase 1 from September 2011. The top page in Figure 16 explains the general ideas and goals in terms of focusing on prediction, explanation, and modeling. The bottom page in Figure 16 sketches out the interfaces and representations and also describes the functions and relationships of the controls and representations. As described in the bottom page, fuzzies were to be programmed by specifying acceleration, velocity, and starting position. Robots would be programmed (or "hacked") by adjusting their motion graphs.

The first row of Table 3 summarizes our driving assumptions for the design of *SURGE Symbolic* during Phase 1. The game company was resistant to trying to develop the game with us, however, because the vision seemed too far afield from their vision of digital games. Indeed, the September 2011 planning document in Figure 16 is a bit daunting from that perspective.

The game company responded with sketches of a game where a person ran through caves. Their vision for *SURGE Symbolic* emphasized the phenomenological representations and backgrounded the formal representations and the disciplinary learning goals and interactions. After further



September 9, 2011

Hi All,

Here is an initial rough outline framing the game to be designed for EGAME. I've also cut a description of the simulation from which it draws inspiration as a 5 page pdf.

Learning Goals:

1. Basic kinematics: position velocity acceleration and impact of friction
2. Graphing
3. Bonus: Might move into some basic treatments of Newton's laws in later levels

Players will have supports for modeling such that in planning phases they can have visual modeling of how the motion on the track (which is essentially a numberline) aggregates into the motion graphs.

Research will focus on scaffolding for prediction, explanation, and modeling.

Talk to you on Monday!

Doug

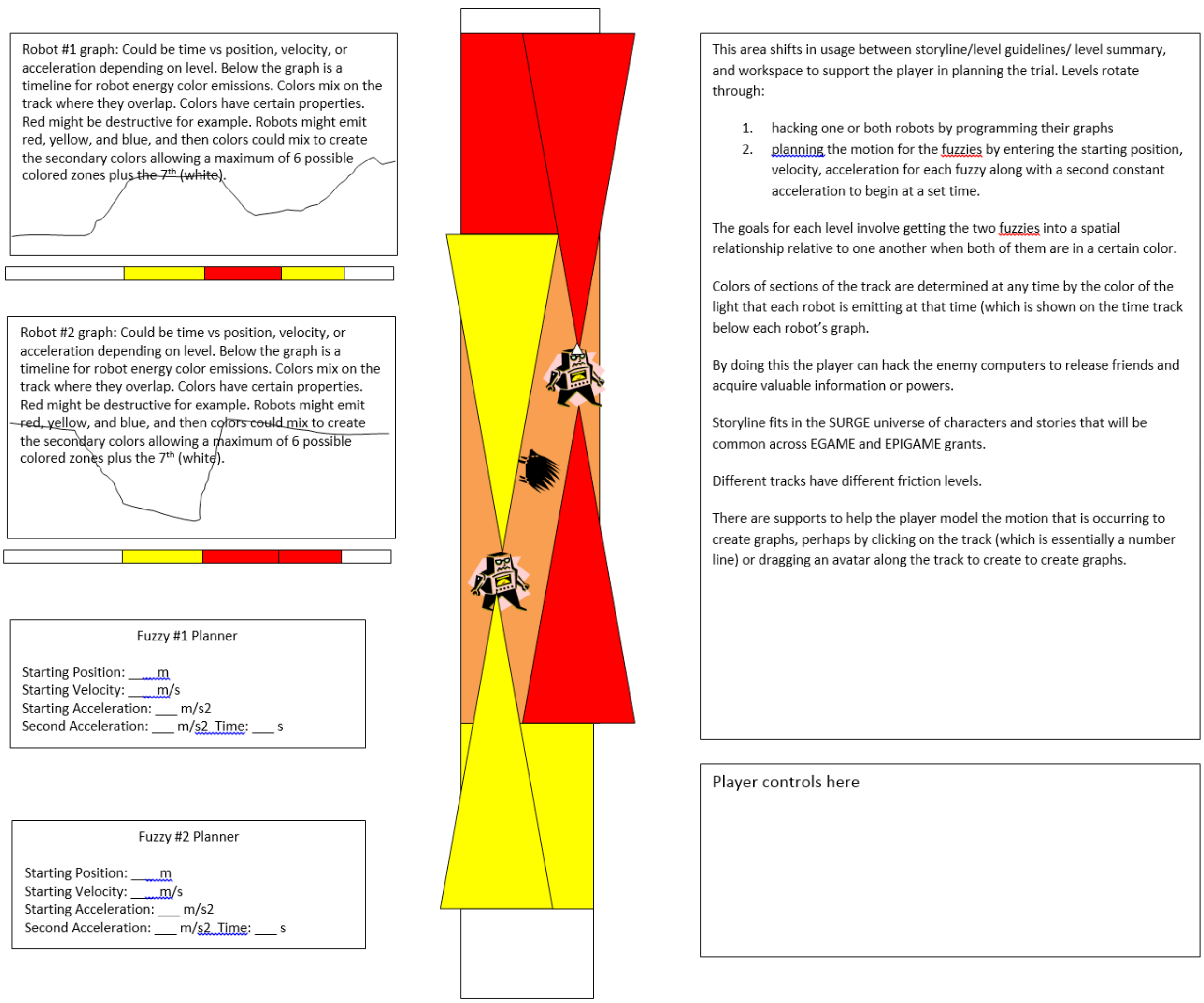


FIGURE 16. Original 2011 plans for *SURGE Symbolic 1*. The accompanying description highlights the focus on engaging the players in prediction, explanation, and modeling (top). The lower diagram explains the programming of the robots and the Fuzzies. The emphasis is clearly moving away from the phenomenological representations and into the formal representations (bottom).

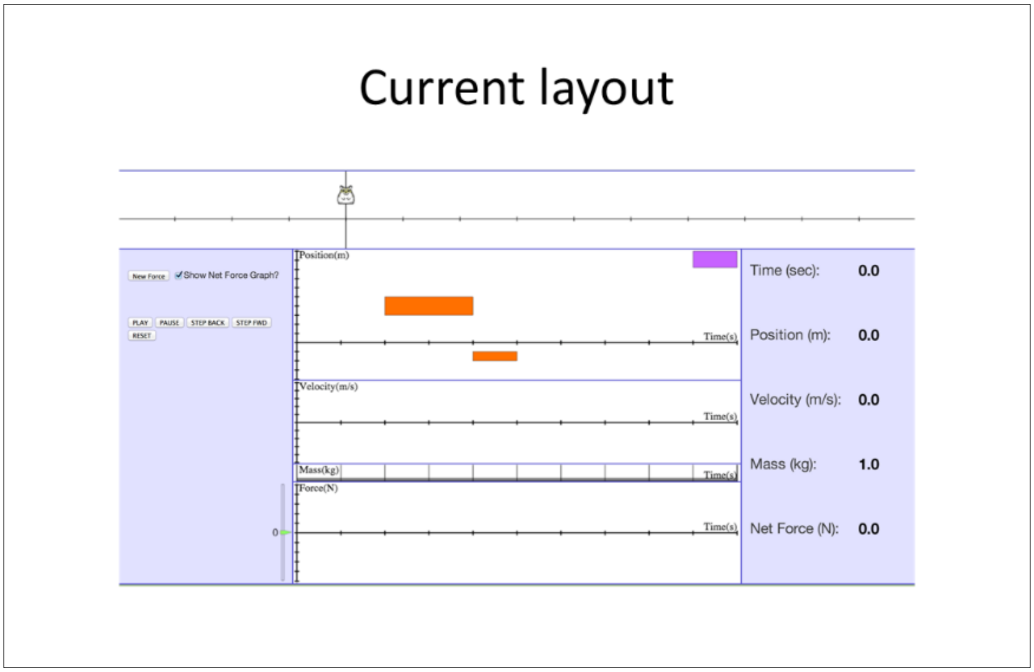
discussion and meetings with the game company, we decided to start with *Fuzzy Chronicles* (on which the game company was also contracted) and then return to *SURGE Symbolic* after the game company and our team had had more time to brainstorm and discuss.

In the meantime, we focused on *SURGE Symbolic* internally (along with our research on *SURGE Next*, *Fuzzy Chronicles*, and *SURGE NextG*). A major critique of our Phase 1 plans involved the complexity and profusion of representations and controls. Essentially, we had a very hard time bringing the game company on board because the controls and representations were (a) visually complex in combination with the visual complexity of the proposed phenomenological representation and (b) very different from the controls and representations of any popular game genre to which the company could relate. We realized that if we had trouble bringing the game company up to speed, bringing students and teachers up to speed would be equally challenging. During our Phase 2 planning, we decided to (a) streamline visual complexity to focus on key disciplinary relationships, (b) situate game-play squarely in the formal representations in terms of controls as well as in terms of the communication of goals and challenges, and (c) background the phenomenological representation. Essentially, the phenomenological representation would shift into an overlay/support role to the central game-play in the formal representations. Our assumptions and design approaches are presented in the second row of Table 3.

Figure 17 presents screenshots of a Phase 2 internal prototype developed in DART. During the prototyping and planning for Phase 2, we revised the software extensively as we worked to balance and clarify the connections between the formal and phenomenological representations and controls. Figure 17 depicts, for example, our decision to reorganize and reorient the phenomenological representation. Figure 17 also depicts a "throttle" control with which we were experimenting. The "throttle" slider to the left of the force graph allows players to adjust force in real time. We felt that a throttle would mirror similar controls in recreational games and thus provide players with an intuitive initial controller for exploring the representations and relationships. Essentially, we wanted an intuitive control option to help players begin to think about the force graph in relationship to Surge's motion and the other graphs. As we played with the prototype,

ASSUMPTIONS	REPRESENTATIONS AND CONTROLS
***SURGE Symbolic 1* (2011):** Layer informal representations over formal representations while organizing game play explicitly around navigating, translating, and coordinating across representations.	**Representations:** Shift to 1D focus on motion in phenomenological representations and formal representations. Challenges communicated to players in the graphs and in the phenomenological representations. **Controls:** Shift controls into Cartesian Graphs and explicit specification of position, velocity, and acceleration.
***SURGE Symbolic 2* (2013):** Situate game-play squarely in the formal representations in terms of controls as well as in terms of communication of goals and challenges. Phenomenological representation backgrounded. Streamline visual complexity to focus on key relationships.	**Representations:** Focus all levels on Cartesian graphs of position and velocity and acceleration that communicate all challenges and goals. **Controls:** Place forces in force/acceleration graph to create the graph to plan controls. Then launch plan.
***SURGE Symbolic 3* (2014):** Also need to provide control opportunities in Cartesian representations more proximal to the phenomenological representation.	**Controls:** Place forces in force/acceleration graph to create the force graph and/or adjust initial values in the position, velocity, or force graphs to plan controls. Then launch plan.
***SURGE Symbolic 4* (2015):** Support students in translating across intermediate representations, phenomenological representation, and formal representations. Focus game-play even more directly into the heart of the formal representations rather than focusing game-play on initial parameters.	**Representations:** Focus early game-play in the position graph, velocity graph, and dot trace representation instead of the force graph. **Controls:** Place interchangeable graph blocks, text blocks, or dot-trace blocks to create the position, velocity, or acceleration graphs to plan controls. Support players in interpreting and reflecting upon the meanings of the graphs beyond simply encouraging the students to manipulate the shapes of the graphs by using controls to engage the player in translating and interpreting from graphs to words and words to graphs for each segment of a graph.

TABLE 3. Design assumptions driving the evolution of *SURGE Symbolic*.



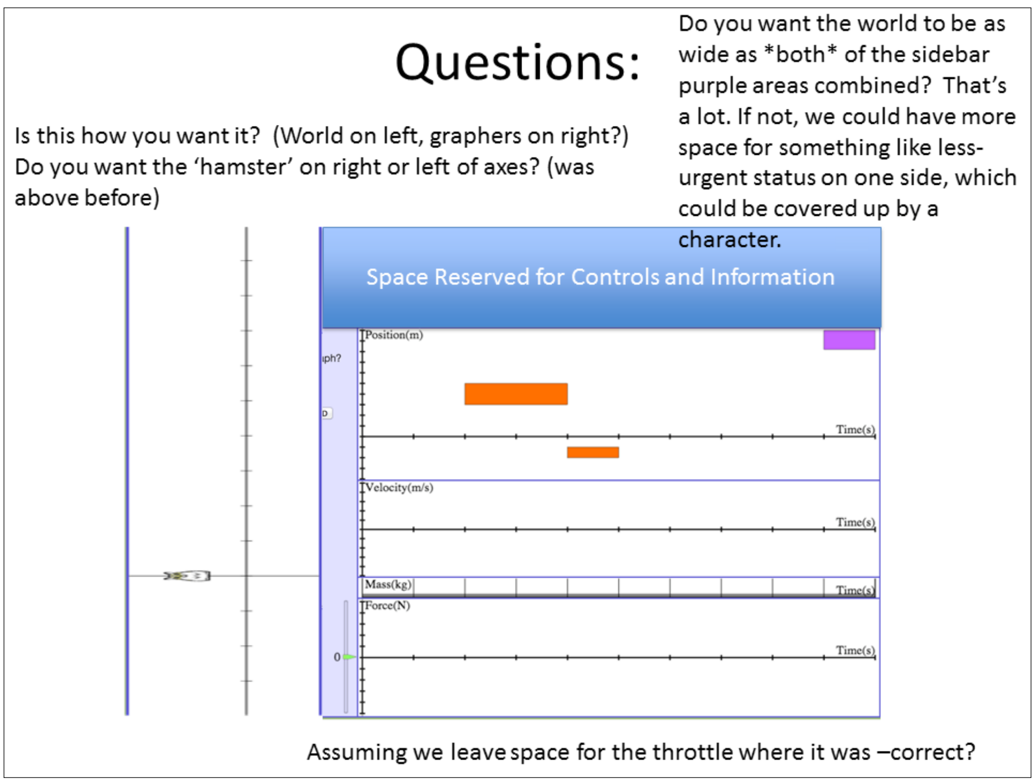


FIGURE 17. Planning for *SURGE Symbolic 2* with screenshots of a working prototype from September 2013 programmed in DART. At this juncture we decided to move the phenomenological representation from its prominent horizontal alignment at the top of the screen (upper image) to the less prominent vertical alignment on the left of the screen (bottom image). The vertical alignment facilitated connections with the graphs, but the horizontal alignment had felt more game-like. In terms of controls, players could place forces in the force graph at the bottom of the screen or manually manipulate the throttle slider in real time that is just to the left of the force graph.

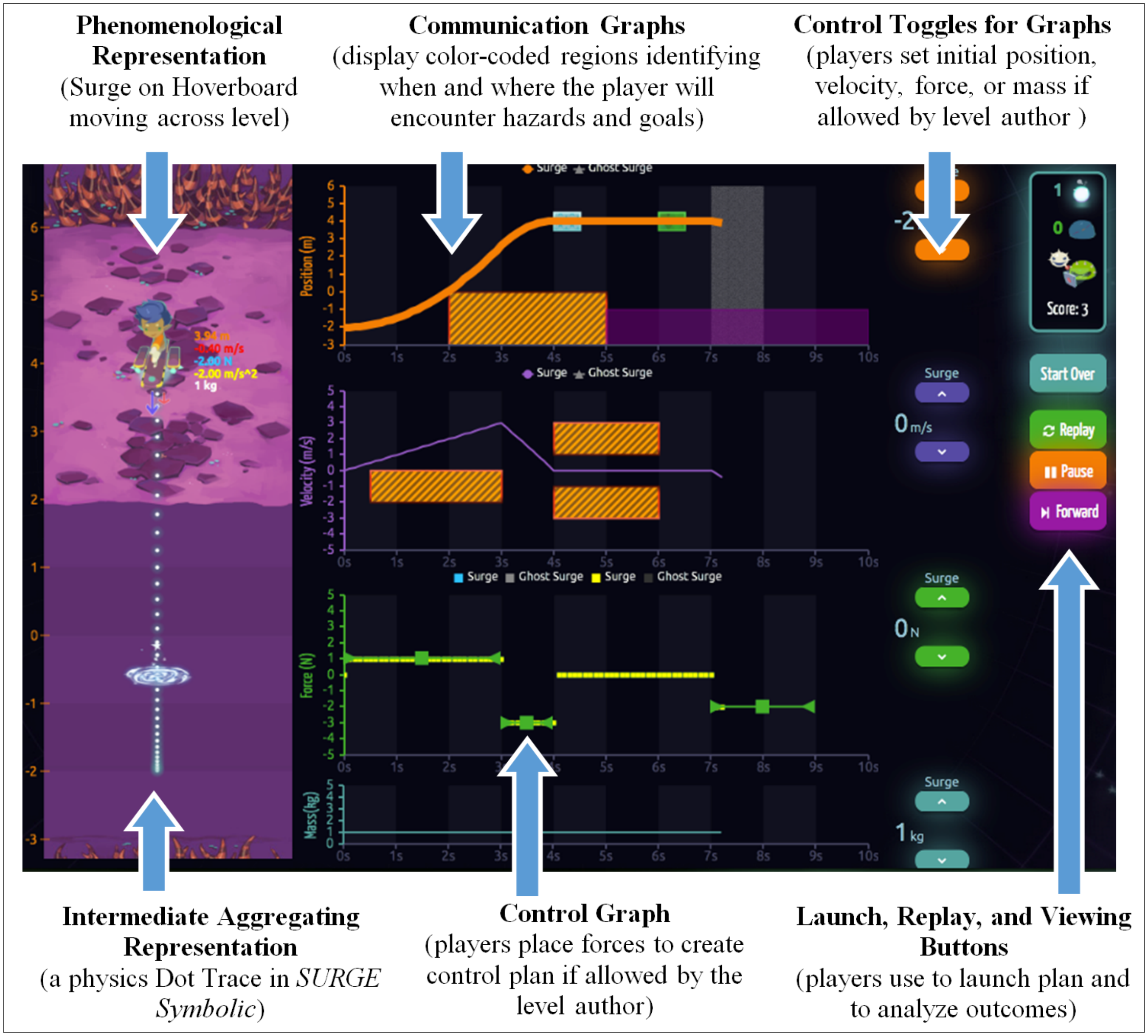


FIGURE 18. *SURGE Symbolic Phase 3.* Note that *SURGE Symbolic Phase 2* was essentially the same without the Control Toggles marked in the upper right.

though, it became clear that the throttle (a) encouraged the player to focus on the phenomenological representation rather than the formal representations and relationships, (b) frustrated the player because it promised a type of popular game-play that was not the focus or nature of the mechanics of our game, and (c) provided a crude control for the game mechanisms and relationships we wished to highlight. We therefore eventually dropped the throttle control, but we remained acutely aware of the importance of helping players connect their actions in the force graphs across all of the intermediate graphs to the phenomenological representation. Force and mass determine acceleration. Acceleration determines velocity. Velocity determines position. Controlling the forces, rather than directly controlling position or velocity, was too challenging for new players because force is so many steps removed from position or velocity conceptually.

After many cycles of iteration and substantial revision of the internal DART prototype, we decided we should create a clean code-base using JavaScript to move forward. We also began working with the game company on the project again. We and the game company had learned a great deal about working with one another through our experiences with *Fuzzy Chronicles*, and we organized the workflows, planning, and design processes accordingly. We decided to keep the development of the game engine internal to our group. We felt that (a) we understood the disciplinary and learning processes underlying those relationships in greater detail

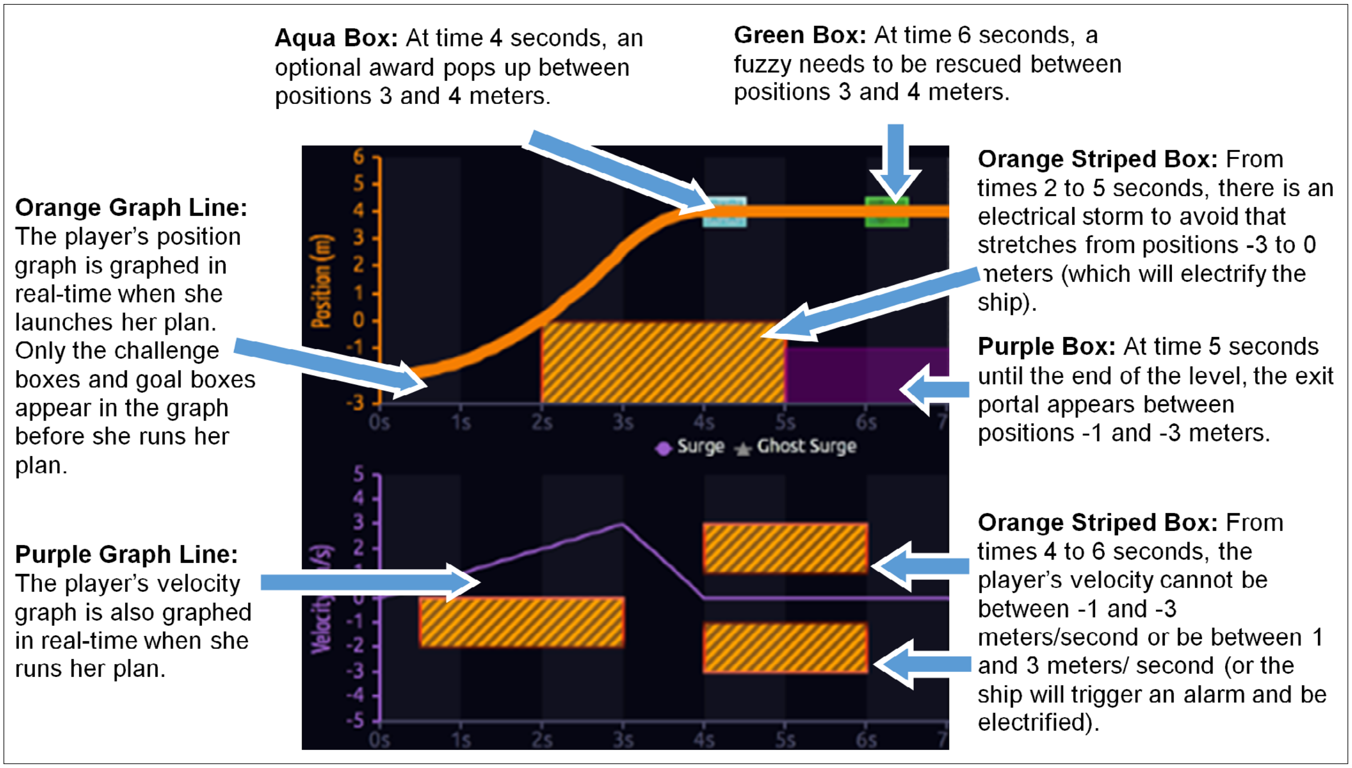


FIGURE 19. The goals and challenges of a game level in *SURGE Symbolic* are communicated in the Cartesian graphs.

than the game company could as a function of all the time we had spent thinking and working with the issues and (b) we would need to continue to refine the structures, designs, and relationships of the formal and phenomenological representations beyond our contract with the game company. We thus decided to focus the game company on the dialog engine, art assets, and interface assets for the game. We hoped that this organization of development efforts would concentrate resources in a manner to leverage the best of both worlds. The game company's artists are among the best in the industry, and we were also confident that we could collaborate with the game company to design an excellent dialog engine. The result was *SURGE Symbolic* Phases 2 and 3 as shown in Figure 18. Ultimately, we were very happy with the outcomes of the work plan and we would approach future projects in a similar manner.

SURGE Symbolic Phase 3 became our prototypical "disciplinarily-integrated game" in Clark, Sengupta, et al. (2015). Whereas *SURGE Classic*, *SURGE Next*, and *Fuzzy Chronicles* focused on layering formal representations over informal representations, *SURGE Symbolic* Phases 2 and 3 inverted this relationship, layering informal representations over formal representations while organizing game play explicitly around navigating, translating, and coordinating across formal representations. More specifically, earlier versions of *SURGE* encouraged players to reflect on formal representations to refine their game strategies, but the formal representations were not the medium through which players planned, implemented, and manipulated their game strategies. Earlier versions of *SURGE* provided vector representations, for example, to help students understand what was happening and how they might adjust their control strategy, but these formal representations communicated only information that a player might or might not use (or even notice). The challenges and opportunities in game levels were communicated through the phenomenological representation, not in the formal representations. Similarly, the player's controls for executing a strategy were also independent of the formal representations. Thus, attending to the formal representations might help a player succeed in a level (or improve on the level), but formal representations were not the medium through which challenges and opportunities were communicated to the player, nor were formal representations the medium of control.

As Clark, Sengupta, and Virk (2016) explain, *SURGE Symbolic* (and all disciplinarily-integrated games) use formal representations as the medium through which challenges and opportunities are communicated to the player (communication representations). *SURGE Symbolic* (and all disciplinarily-integrated games) also use formal representations as the medium through which the player implements strategies and exerts control over the game (control representations). Some disciplinarily-integrated games might use the same representation for both control and communication, while other disciplinarily-integrated games might use one or more formal representations for communication and one or more

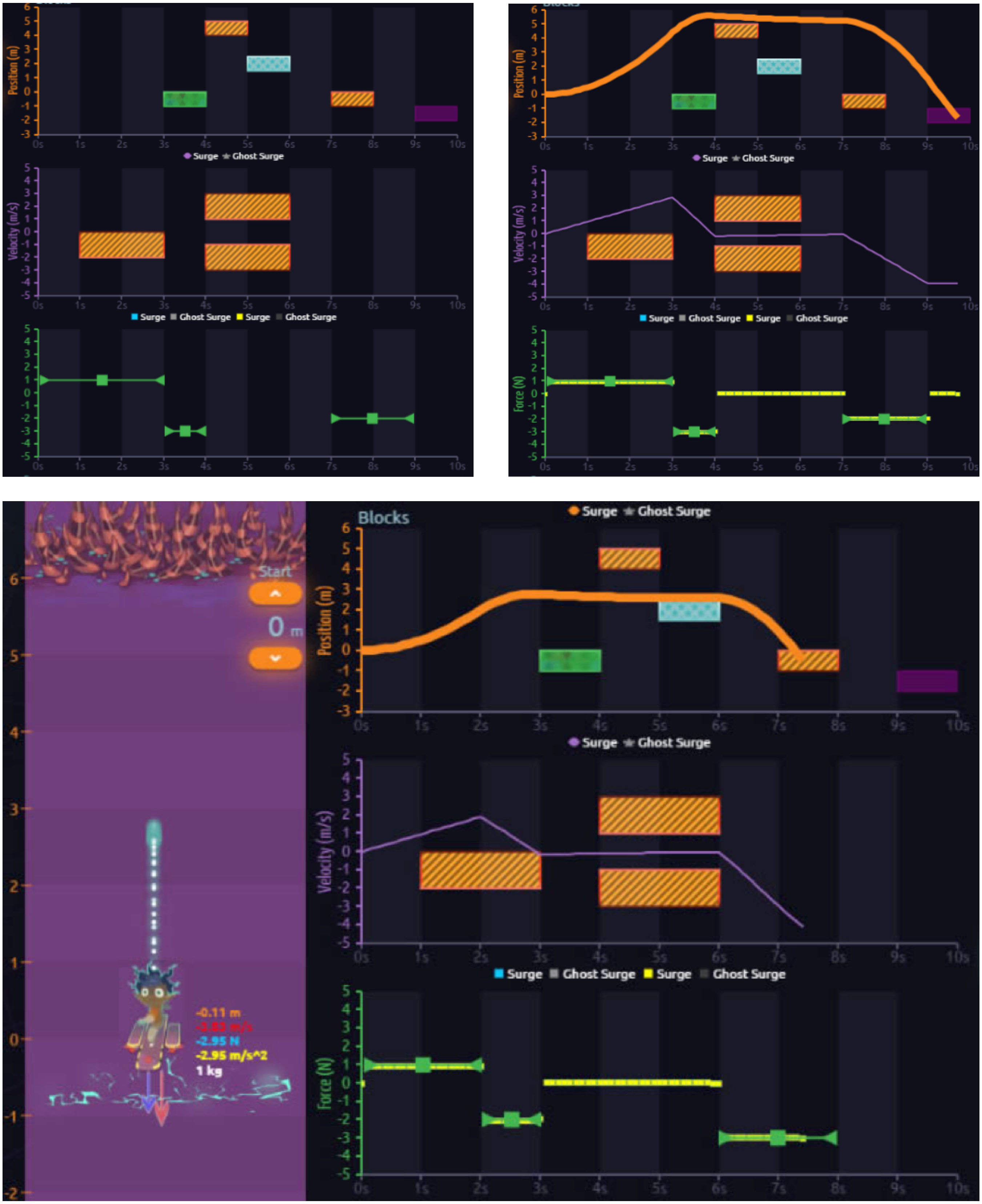


FIGURE 20. Plan before launching (top left) and the graphs of the plan after successful launch that reaches the purple exit portal that appears at the position between -1 and -2 meters at time 9 seconds, as marked in the position graph by the purple box – Note that the player navigated between the velocity traps in the velocity graph by maintaining acceptable velocities (top right). An unsuccessful plan ends up electrifying Surge because she enter the electrified area between 0 and -1 meters that appears there between times 7 and 8 seconds (bottom).

other representations for control. All disciplinarily-integrated games include a phenomenological representation (which in traditional digital games would be the primary focus). Furthermore, all disciplinarily-integrated games include an intermediate representation to support players in translating from the phenomenological representation to the formal representations and to constrain their interpretation of the formal representations. The goal in all disciplinarily-integrated games involves interpreting, creating, modifying, and translating across these formal and phenomenological representations.

The template for *SURGE Symbolic* in Figure 18, for example, presents the phenomenological representation on the left side. The phenomenological representation portrays the heroine, Surge, on her hoverboard moving forward and backward along a phenomenological representation of the game world. The formal Cartesian graphs on the right side are the communication and control representations. The position and velocity graphs in Figure 18, for example, can present information about the specific regions of the game-world that will be affected by dangerous electrical storms at given times, as well as information about locations and times where rewards will appear or allies will rendezvous with Surge. As a result of this design approach, the Cartesian space emerges as a set of scientific instruments for the player by communicating data about the game world that are not available through other means.

While the challenges and opportunities are communicated through the position and velocity graphs in Figure 18, any subset (or all) of the Cartesian graphs could serve this role. Simultaneously, the Cartesian graphs also play the role of an instrument panel or mission planner, offering control over the movement of the Surge spacecraft. In Figure 18, for example, the player can exert control by placing forces of various magnitudes and durations at different time points in the force graph. This force graph control scheme was the focus of Phase 2 development. Phase 3 development provided an alternative (or supplementary) control scheme for early game levels using the toggles to the right side of the graphs. The player could exert control through any of the graphs using the toggles to set the initial values for position, velocity, and acceleration on the respective graphs. The author of a game level designates which graphs are visible to the player, which graphs are used for which purposes (communication or control), which (if any) of the toggles are active, and what challenges and goals constitute the level. Figure 19 describes the challenges and goals communicated in the position and velocity graphs of Figure 18.

The top images of Figure 20 show what a level might look like before launching the plan (left) and after launching the plan (right). The player's plan is specified in the force graphs. The challenges and goals are communicated to the player in the position and velocity graphs. As shown in the top right

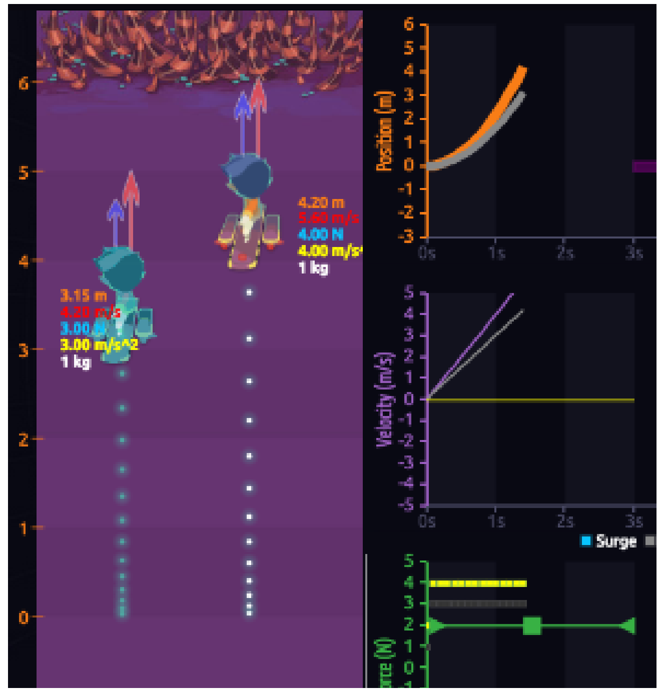


FIGURE 21. Ghost Surge and Surge can both be included to provide comparison points or challenges.

image, the plan is successful because Surge reaches the exit portal that appears between positions -1 and -2 meters at the time 9 seconds. The bottom image in Figure 20 depicts an unsuccessful plan that resulted in Surge getting electrified because she entered the electrified area between 0 and -1 meters that appears between times 7 and 8 seconds. Note that all of the action also unfolds in the phenomenological representation, but all communication and control are centered in the formal representations – the phenomenological representation plays a reinforcing and support role rather than a central role. Figure 21 depicts a level that includes Ghost Surge, a holographic companion that authors of levels can include to provide challenges for the player that engages the player in comparing graphs. Authors can specify the degree to which players can control Ghost Surge as part of these challenges.

During Phase 2 and Phase 3 development, we also continued with our work to integrate dialog into the game. While our attempts to integrate explanation dialog in *SURGE Next* and *Fuzzy Chronicles* caused compromises or constraints for the representations and controls, *SURGE Symbolic* thus far appears to have benefitted from the design plan of developing the dialog system separately with the game company. We designed the dialog system with the game company as an overlay that provides an API for embedding games like *SURGE Symbolic*. The dialog system is very flexible and can include text and images as shown in Figure 22 (top and middle). Furthermore, we can create as simple or long or complexly branched trees of dialog nodes as we would like (see Figure 22 bottom), with each node including whatever

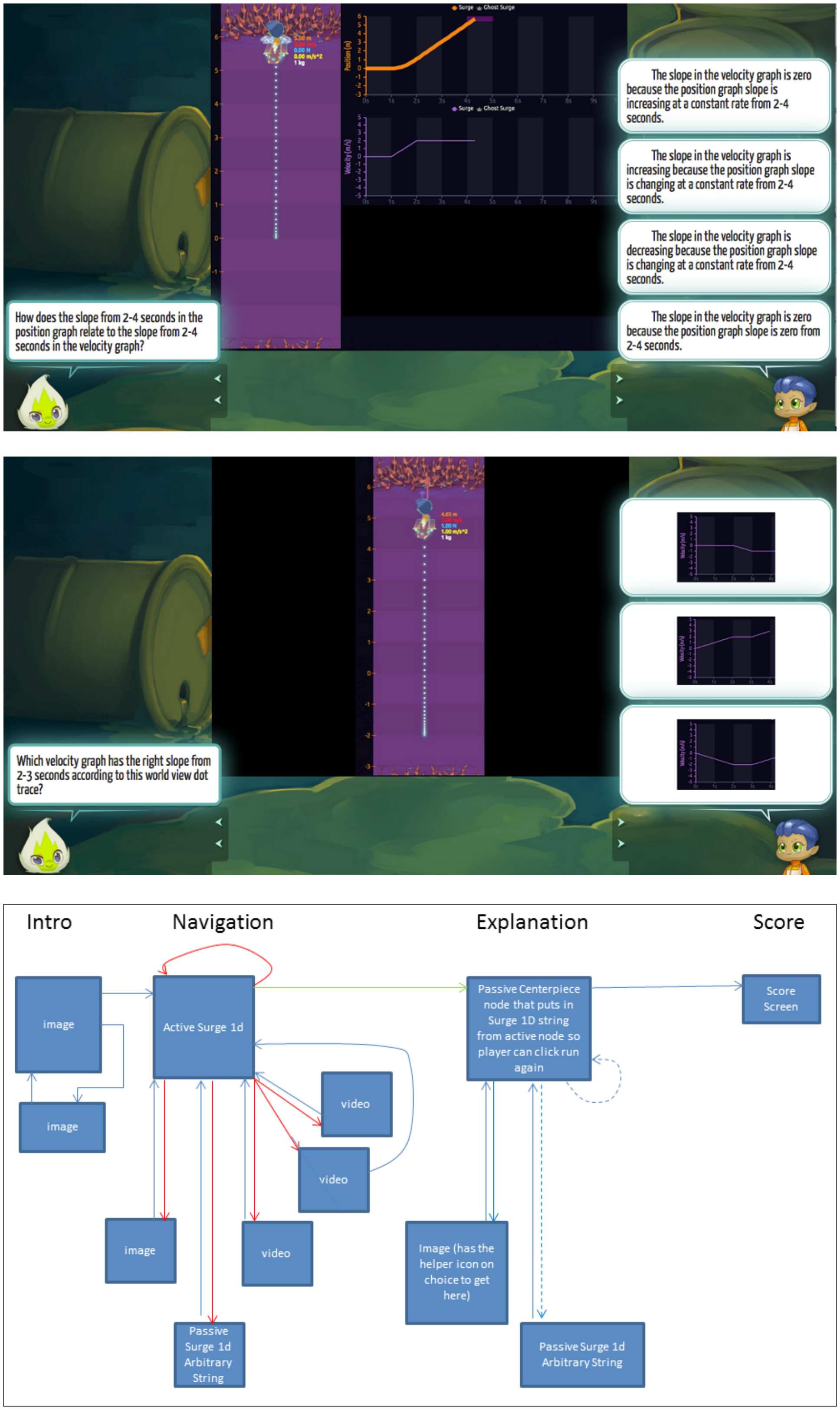


FIGURE 22. *SURGE Symbolic* game levels are embedded in a flexible dialog tree. The dialog options in a node can include any combination of images and text (top two images). The specific dialog options that appear in a node can be filtered by the player's actions in a game level in that node in terms of success, failure, greater history of attempts, or based on any performance tag that the game level provides back to the dialog tree based on the player's performance. Nodes can connect to any number of nodes, and nodes can contain game levels, videos, images, or text (bottom image). The goal is to allow the progression of levels to adapt to the needs, proficiency, and interests of the player.

type of centerpiece that we would like (images, video, *SURGE Symbolic*, etc.). Most interestingly, we worked with the game company to design a system whereby the dialog options of a node are filtered (i.e., change) depending on the tags that are passed back to it by an embedded game. Thus, a player's success or failure in a node's game level, as well as the cause of failure or the past history of successes, can be used to filter which dialog options appear in the node. This allows us to continue to iterate and refine our game engine however we want, as well as introduce new tags to describe what we determine to be salient about student play, and then add those tags into the nodes we author. We attribute the success of the dialog tree aspect of the project to the facts that (a) it was the focus of the game company's development for the project, (b) the game company could relate well to the nature of a dialog system as a concept because it did not involve disciplinarily or pedagogically specific aspects, and (c) the game company's project manager and programmer worked closely with us to spec out all the functionality in advance in great detail. The game company then had a solid foundation from which to work, and ultimately the game company created and delivered a system that fabulously exceeded our expectations. This was a polar opposite experience (and thrilling outcome) compared to working on the *Fuzzy Chronicles* project with the same company and really demonstrated the power and importance of the shared working understandings and clear expectations developed through that first project.

In terms of our internal work on the game itself, Phases 2, 3, and 4 represent iterations of the control schemes through which the player acts on the game. While Phase 2 focused entirely on the force graph for controls, we soon realized that this graph was too many steps conceptually removed from the phenomenological representation to support new players. We realized that we needed to provide intuitive controls that were conceptually closer to the phenomenological representation. Essentially, as discussed earlier, the force graph needs to be interpreted in conjunction with the mass graph to make sense of changes to the acceleration graph. The acceleration graph drives the velocity graph. The velocity graph drives the position graph. Changes and movement in the position and velocity graphs are conceptually much closer to changes in the phenomenological representation. Ultimately, we would like for players to translate across each and all of representations. As part of Phase 2,

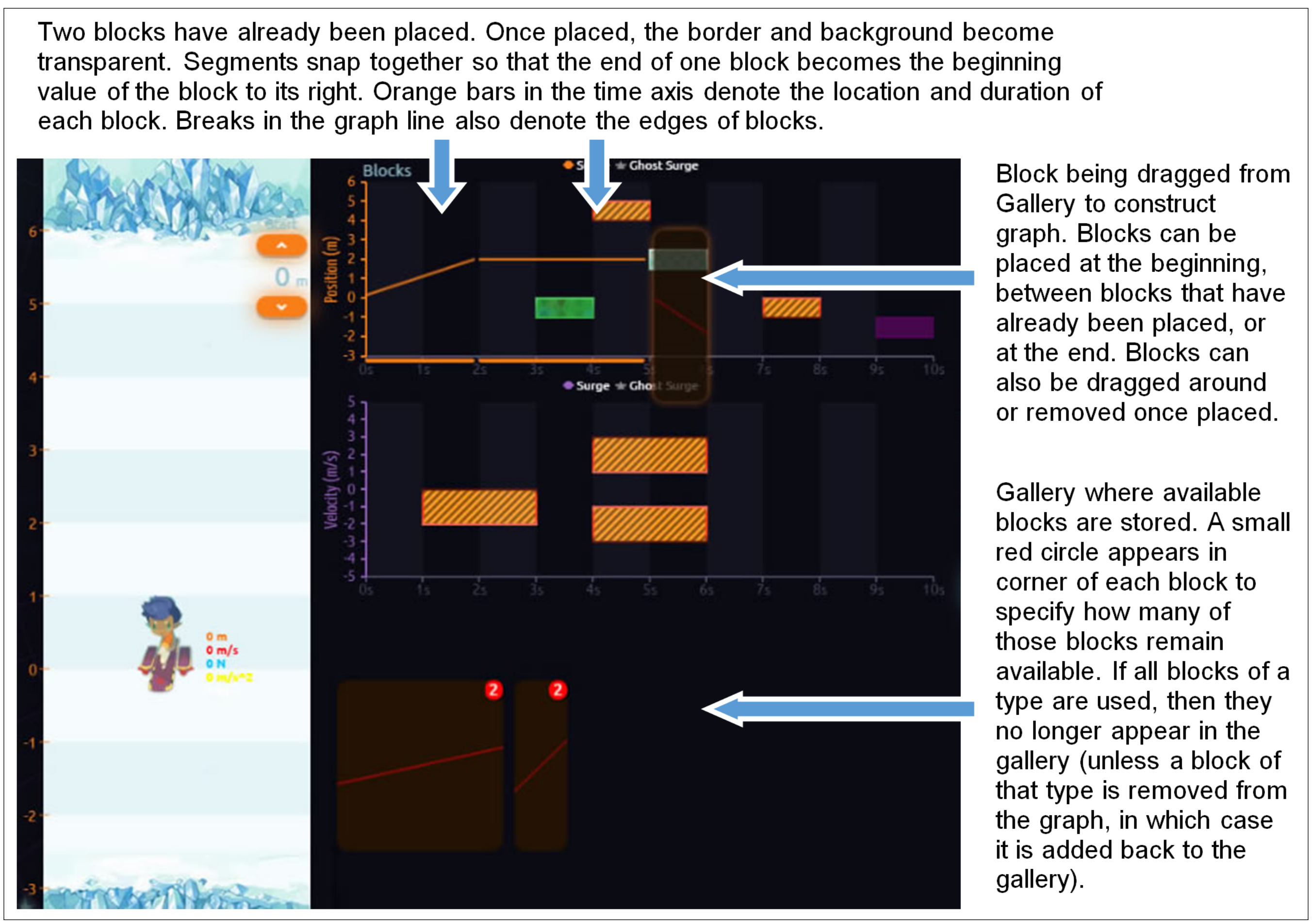


FIGURE 23. Students drag "graph blocks" representing various slopes to assemble graphs directly to control the game.

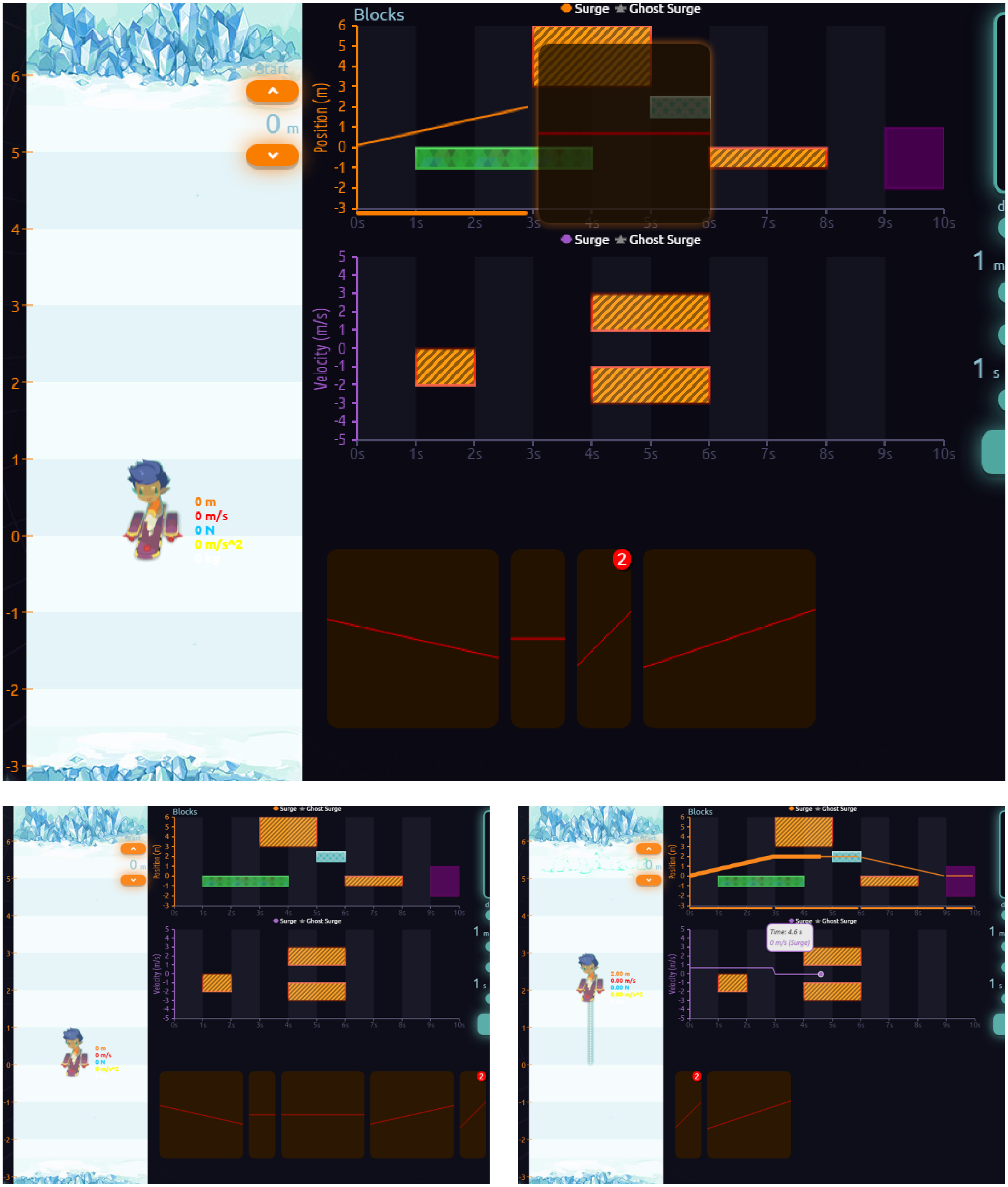


FIGURE 24. Example of student using graphical blocks, including default gallery state (bottom left), dragging blocks onto the graphs (top), and launching the completed plan (bottom right).

FIGURE 25. Example of student using word blocks, including default gallery state (bottom left), dragging blocks underneath a graph (top), and launching the completed plan (bottom right).

we experimented with the real-time throttle slider shown in Figure 17. The intention was to allow players controlling the forces in real time with the throttle to begin to connect the idea of forces to their results in the other graphs and the phenomenological world, but ultimately we realized that the throttle supported only a reactive "twitch" game understanding without strong connections to the formal representations.

This realization led to a huge overhaul in the underlying engine to allow the player to set the initial positions for her ship and for her sidekick's ship (the holographic Ghost Surge) in a manner reminiscent of Parnafes and diSessa's (2004) NumberSpeed (see Figure 15). In Phase 3, the author of a level could specify which of these parameters the player could change, which parameters were fixed, and how many force and throw commands (if any) the player could place in the force graph.

This allowed authors to create levels where players' challenges were located in the position or velocity graphs, which are the formal representations most proximal to the phenomenological representation, the intermediate dot trace representation, and the players' intuitive understandings. More specifically, the dot trace intermediate representation directly bridges the velocity and position graphs to the phenomenological representation and players' intuitive understandings, whereas spanning the force graph to the phenomenological representation requires much longer chains of reasoning. Pilot testing in a school with Phase 3 demonstrated that we were making progress in terms of developing a control scheme that supported players in bridging across representations, but the pilot testing also made clear that (a) we needed to focus early game-play in the position graph, velocity graph, and dot trace representation, (b) we needed to focus game-play even more directly into the heart of the formal representations rather than focusing game-play on initial parameters, and (c) we needed to support players in interpreting and reflecting upon the meanings of the graphs beyond simply encouraging the students to manipulate the shapes of the graphs.

Phase 4 undertakes this next step. We are now engaging the player in manipulating, interpreting, and translating between the meaning and symbolic representation of each piece of each formal representation. Essentially, we are not only allowing the player to drag blocks (see Figures 23 and 24) together to create graphs (which *SimCalc* had also done in one incarnation, Roschelle, Kaput, & Stroup, 2000) but we are also integrating the capability of "word" blocks that specify the block verbally (see Figure 25). This allows players to specify a plan either in word blocks or graph blocks for position, velocity, or acceleration. Graph blocks are dragged directly onto the target graph (see Figure 24) to emphasize the importance of the graph representation in line with Mayer's (2005) spatial contiguity principle. Word blocks are dragged directly underneath the segment of the graph they are controlling in terms of the beginning and ending time of the segment (see Figure 25). Our programmer is currently refining the block technology, and we have thus not piloted the blocks with students, but our intention is to create sequences of levels for the player such that the player begins with controls that are intuitive and conceptually focused on developing familiarity with the position and velocity graphs. These sequences will include levels where players are essentially translating back and forth between verbal and graphical representations of each segment to control the hoverboard. By shifting modes back and forth between verbal and graphical blocks, we hope to incentivize players to make sense of the meanings of the shape of each graph segment rather than simply manipulating the shape of the lines to fit the challenges and goals arrayed within the graphs.

Later design and development of Phase 4 will focus on adding dot trace blocks to control the ship. Dot trace blocks will be dragged directly under the segment of a graph that they control in terms of the beginning and ending time of the segment (just like word blocks). Dot trace blocks will focus the player's strategizing and thinking in the dot trace representations as the organic link between the player's intuitive sense of phenomenological motion and Cartesian representations of that motion. We prototyped the dot trace blocks as part of our planning of the other blocks for Phase 4, but we quickly realized that the dot trace blocks represent a more complex user interface design challenge. In a conceptual sense, the abstract plans for the dot trace blocks fit seamlessly with our word and graphical blocks for position, velocity, and acceleration, but the user interface design of the dot trace blocks will prove more challenging to design in terms of a user interface that is intuitive for players while also capturing the essence of the dot trace representation for the player. We are therefore waiting until we have piloted the word and graphical blocks for position, velocity, and acceleration before continuing with the design and prototyping of the dot trace blocks. Once we have feedback on our current blocks, along with the opportunity to iterate upon the design of those blocks, we will return to the dot trace block design.

THE PATH AND PARALLEL EVOLUTION WITH *SIMCALC*

The design of *SURGE Symbolic* has evolved dramatically over the past seven years from our original conceptions of *Switchball* and *SURGE Classic* in Figure 1. Our initial design perspective focused on highlighting the formal physics ideas within popular game-play mechanics. This perspective prioritized a commitment to the phenomenological representations and controls of recreational games. Formal representations were designed around and over these phenomenological representations. The next seven years involved navigating the tensions between the original recreational genre and creating a new genre within the formal representations themselves.

Our design of *SURGE Symbolic* builds on research on teaching physics using simulations and motion sensors (e.g., Brasell, 1987; diSessa, Hammer, Sherin, & Kolpakowski, 1991; Mokros & Tinker, 1987), research on constructing graphs based on assembling relevant "pieces" of trajectories of motion, and research on *SimCalc* (e.g., Hegedus & Roschelle, 2013; Kaput, 1992; Roschelle, Knudsen, & Hegedus, 2010). In the work of Tinker and colleagues, students have often been provided graphs of position or velocity that they were asked to replicate using the controls of the system, which might involve a motion sensor. Similarly, students have been provided with a dot trace representation overlaid on their phenomenological view that they worked to interpret in terms of a graph (diSessa et al., 1991), and *SimCalc* pioneered scaffolding for students' integration and differentiation between and across Cartesian graphs of position and velocity by dynamically linking across representations (Kaput, 1992; Hegedus & Roschelle, 2013).

Thus we do not claim the design of *SURGE Symbolic* evolved in a vacuum, but Clark, Sengupta, et al. (2015) and Clark, Sengupta, and Virk (2016) argue that *SURGE Symbolic* (and disciplinarily-integrated games more generally) build on these bodies of research by pushing more deeply on approaches for leveraging formal representations as the means of communicating challenges and leveraging abstract formal representations as the players' means of control within the game. Our work also builds upon previous research by pursuing an approach that retains commitments as a game while exploring these commitments in a new genre.

Thus, *SURGE Symbolic* (and all disciplinarily-integrated games) have evolved to have the following design characteristics that emphasize the primacy of the formal representations

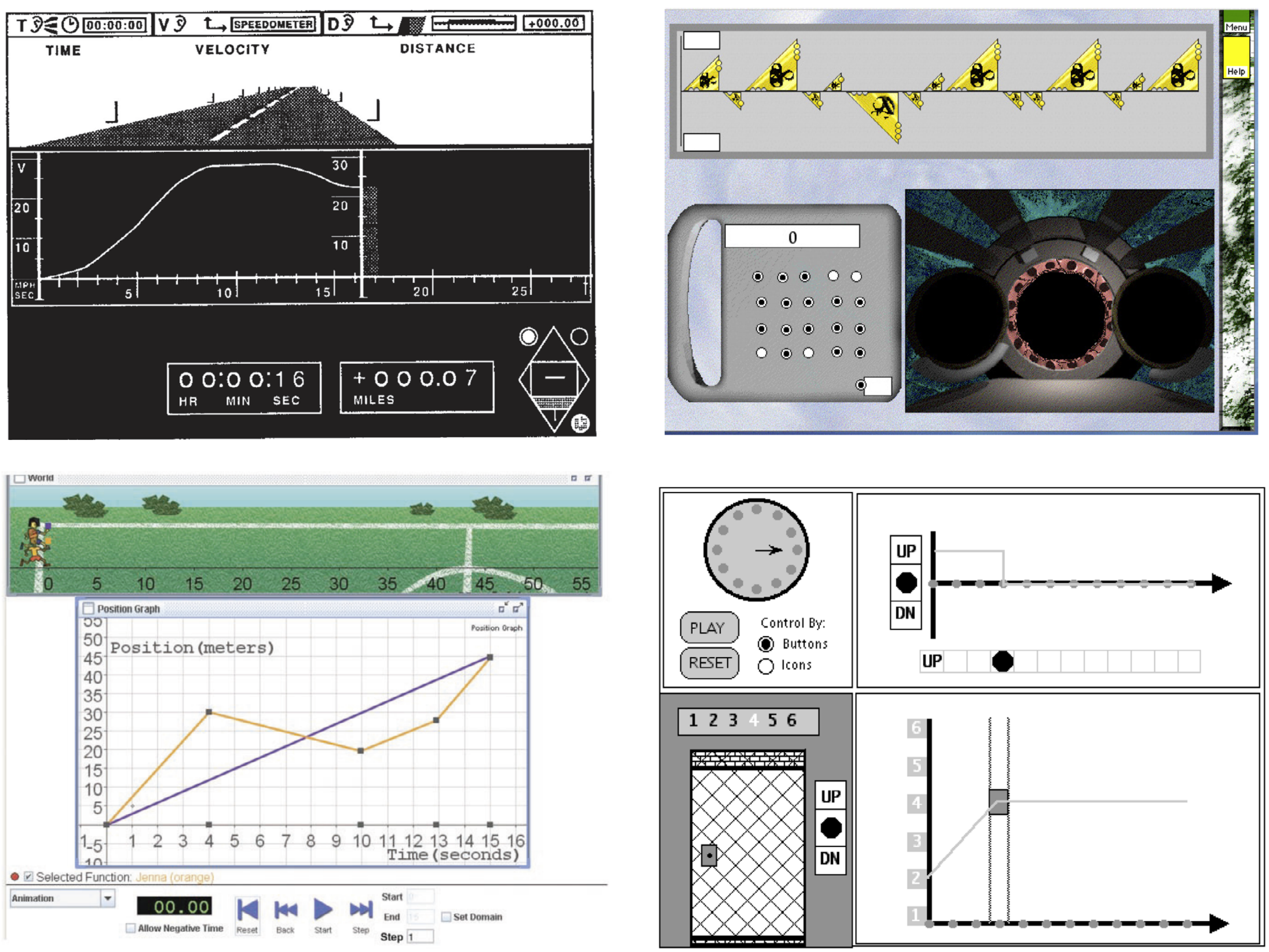


FIGURE 26. The evolution of SimCalc. The original vision with an emphasis on a rich phenomenological perspective (top left). The final version with an emphasis on formal representations as an environment for inquiry and modeling rather than a game (bottom left). The early "Alien Elevators" game version where game mechanics considerations superseded and adulterated the formal representations (top right and bottom right). Figure adapted from figures in Tatar, Roschelle, and Hegedus (2014) with permission from the authors.

as the hub of game-play: (a) formal representations for controlling the game, (b) formal representations for communicating challenges and opportunities, (c) a phenomenological representation presenting the phenomenon being modeled, (d) an intermediate aggregating representation for translating between the formal and phenomenological representations, and (e) game mechanics and goals focused on engaging the player in interpreting, creating, modifying, and translating across these formal and phenomenological representations.

While *SURGE Symbolic* focuses explicitly on Newtonian dynamics, Clark, Sengupta, and Virk (2016) explores the generalizability of the disciplinarily-integrated game template in hypothetical examples of games in physics, biology, chemistry, and the social sciences in terms of: (a) time-series analyses with Cartesian formal representations, (b) constraint-system analyses with Cartesian formal representations, and (c) other model types and non-Cartesian formal representations. Ultimately, we are still committed to designing an environment that is a true game, but which represents a new genre of true game rather than an adaptation of an extant recreational genre to enhance pedagogical affordances.

We have certainly learned a number of lessons through the process, and many of these lessons parallel lessons encountered by the *SimCalc* team as recounted in their retrospective in the recent IJDL special issue on historic design cases (Tatar, Roschelle, and Hegedus, 2014). As the *SimCalc* team recounted, they too began with a commitment to phenomenological representations (see Figure 26 top left) and game mechanics (see Figure 26 top right). They ultimately shifted away from these commitments to focus on the formal representations (see Figure 26 bottom left). In terms of the phenomenological representations, they found that they needed to shift from a rich immersive 3D representation that would "engage students' kinesthetic sense, such as the visual sense of the world 'zooming by' both sides of the perceptual field" (p. 89) to a 2D representation that facilitates mapping between the phenomenological representations and the formal representations. In terms of the focus on game mechanics, the *SimCalc* team found that popular "gameness" could not be allowed to adulterate or dilute the formal representations. According to their retrospective, they themselves had difficulty interpreting the hybrid game/formal representations in the top right image of Figure 26 as they were writing the retrospective. Clearly the screenshots are exciting and game-like, but they occluded the formal relationships and representations rather than accentuating them. In the *SimCalc* team's own words, they had "created an interface that was gamelike, based on a narrative, and involved high-quality graphics but which submerged the mathematics" (p. 89).

Ultimately, the *SimCalc* team shifted away from their emphasis on games and focused entirely on creating an environment for student inquiry and modeling that focused explicitly on the formal representations. We agree entirely with the *SimCalc* team about the need to remain clear and true to the integrity of the formal representations. If "predicting, explaining, and modeling" with the formal representations is the central learning goal of an environment, as we stated in our original 2011 design document for *SURGE Symbolic* Phase 1, we believe from our experience over the past seven years that the learning goals and formal representations need to be the central organizing focus of the game. The *SimCalc* team came to this same conclusion and decided to shift away from an emphasis on game-play. We are still committed to designing true games, but we have realized that this means designing new genres of game-play rather than adapting or augmenting recreational genres.

Our current design work and research explore the degree to which twin emphases on game-play and formal representations can synergistically co-exist without diminishing one another. Figures 19 and 20 show our current thinking in terms of how challenges and goals and game play can be conceptualized as central and true to the formal representations. Our focus on the word and graphical blocks in Figures 24 and 25 as the means of control does not draw on recreational genres, but the focus is certainly true to the formal representations. We hope that, with refinement, we can engage players in interesting game-play that is engagingly puzzle-like in its own right while also focusing players on interpreting, translating, and making sense of the meanings and semiotics of the formal representations.

In terms of the graphics and narrative, we have definitely stepped back from our original recreational models, but we worked carefully with the game company to design graphics and interfaces that are engaging and rich without distorting or occluding the representations and their relationships. That said, we ultimately may find that we need to strip the graphics and narrative further. Work and development on *ThinkerTools* (see Figure 4), for example, followed this path of focusing on very abstract and spare graphical representations with the explicit goal of facilitating transfer to other contexts (White, 1993; White & Frederiksen, 1998). Recreational games like *Tetris*, which is an abstract spatial puzzle game, and *Bejeweled*, which represents a newer Indie genre of puzzle game, have demonstrated that engaging and successful recreational games can have very abstract, spare graphics and narrative that are still visually and viscerally appealing and engaging.

Even the developers of the fabulously successful AAA digital games *Portal I* and *Portal II* made conscious decisions to strip down the detail and richness of the environment to accentuate rather than obscure the salient relationships of their underlying game mechanics. *Portal I* and *II* focus on devilishly engaging physics puzzles. *Portal I* and *II* both won a huge number of "game of the year" awards in the popular press out of the full field of high-budget recreational games developed in their respective years. *Portal I* and *II* do so without diluting the game immersion by choosing the context of a very ominous, post-apocalyptic, training facility run by artificial intelligences and robots. That description does not sound like it could possibly support a highly successful AAA recreational game, but the designers created a pitch-perfect immersive context and narrative through excellent design work. The result is simultaneously dark, suspenseful, and immersive without visually obscuring the underlying mechanics and relationships of the puzzles.

In some ways, this pattern aligns with a finding of the meta-analysis by Clark, Tanner-Smith, and Killingsworth (2015) that demonstrates a slight but significant negative relationship between contextual richness in games and the learning outcomes. As discussed in the meta-analysis, that relationship may have been a function of differences in learning goals assessed across the studies, and the actual relationship may actually be more neutral, but the findings regardless demonstrate that increased contextual richness does not necessarily equate to increased learning.

Perhaps our ongoing research and design and piloting will push us further down the avenue of further stripping away the graphical and narrative elements in service of highlighting, rather than obscuring, the salient relationships of the underlying game mechanics in *SURGE Symbolic*. We certainly do not claim that we will be able to create a game that is as engaging as *Portal*, *Bejeweled*, or *Tetris*, but we do think that we too need to carefully manage graphical richness to highlight rather than obscure underlying puzzles and relationships. The challenge is to do so in a synergistic rather than destructive manner in terms of game-ness, formal representations, and puzzles.

FINAL THOUGHTS AND CONCLUSIONS

Ultimately, we have come a long way from the *SURGE Classic* proposal. We have moved beyond thinking about how can we adapt and overlay formal representations on popular game mechanics. We now explore how we might design a new genre of game mechanics within formal representations while maintaining high fidelity and clarity within those formal representations.

At the highest level, designing games for learning involves integrating learning goals, game design goals, and technology affordances (including budget). Developing a good recreational digital game requires balancing only game design goals, technology affordances, and budget. Designing a recreational game within these boundaries is challenging enough. Developing a good game for learning adds learning goals to the mix, leaving the intersection within the Venn diagram much smaller.

It is not sufficient to design a compelling game that engages students. Games for science learning should also incorporate models and structures that support disciplinary learning goals and optimize players' learning processes in light of those goals. We have observed that this tension is not always solvable. Furthermore, these goals exist in a fluid hierarchy that makes it difficult to prescribe a solution that will work in every case. In terms of examples from *SURGE Classic*, our initial planning explored a wide range of game mechanics from *Tiger Woods PGA Tour*, *Crayon Physics*, shuffle board, *The Incredible Machine*, and other designs before deciding that adopting a "marble" game mechanic would fit best with our learning goals, technology affordances, and budget. Selecting the other genres would have resulted in more problematic design imperatives impinging on our learning or game goals (which would have required compromising either game-play or learning goals). The "marble" mechanic genre seemed to provide the best fit of the extant genres under consideration. We thus needed to approach the design process very flexibly rather than "forcing" a fit between a pre-selected genre and our learning goals.

This need for flexibility continues throughout development. During *SURGE Classic*, for example, many researchers who played versions of our game commented that increasing prediction and explanation would scaffold our learning goals (as described in much psychology and science education research). By this suggestion, however, those reviewers intended prediction and explanation as implemented in many studies (i.e., having students write their predications in a journal before playing and then return to their journals after play). This would have been, from our perspective, a design imperative issuing from our learning goals and another environmental genre that would have dramatically compromised our game goals in terms of immersion and engagement. With *SURGE Next* and *Fuzzy Chronicles*, we essentially stepped back from our design to think more broadly about a design that synthesized the learning goals of prediction and explanation with our game-play goals. We found that even research conducted in other genres of computer mediated environments (including simulation genres with substantial similarities to game genres) did not translate directly into game genres. Essentially, mixing extant games genres and pedagogical genres is a tricky business, often resulting in submerging or adulterating the disciplinary learning goals. The *SimCalc* team's design retrospective highlights these challenges. In our case, potential solutions required willingness to frequently rethink our overall approach rather than insistence on incremental evolution from the *SURGE Classic* vision.

Recreational game design conventions provide powerful affordances for learning and engagement. It is important to remember, however, that these affordances evolved under different pressures and goals than those of formal learning contexts. For formal learning contexts and goals, synergistically rethinking and redesigning recreational conventions is imperative to support players' explicit articulation of the intuitive understandings that the players develop through game-play. In our case, we now think that trying to adapt extant recreational genres is itself too limiting when learning goals focus on students' exploration of formal disciplinary relationships and representations. With *SURGE Symbolic* we are beginning our exploration of a genre that we have called Disciplinarily-Integrated Games (Clark, Sengupta, et al., 2015) to extend beyond Newtonian dynamics and Cartesian graphs to other disciplines and model types (Clark, Sengupta, & Virk, 2016). Clearly our work builds on the research of many others (e.g., Parnafes & diSessa, 2004; Roschelle et al., 2000; White & Frederiksen, 1998), but we are trying to build upon these bodies of foundational research to design a new true game genre within formal representations while maintaining high fidelity and clarity within the formal representations and disciplinary relationships.

Note that we emphatically do not claim that all games for science learning should abandon popular game genres as templates. Informal learning contexts, for example, have very different learning goals, requirements for universal inclusivity, recreational imperatives, and adoption challenges. Additionally, drawing on existing recreation genres allows students who are familiar with those genres easy access and entry (although simultaneously potentially intimidating students who are not familiar with those genres). Thus, as we encountered in terms of broadening the challenge curve for *SURGE Classic*, games designed for informal learning involve a much wider range of potential balances and mixes of goals and genres.

Furthermore, digital games for formal (and informal) contexts that focus on identity formation and perspective-taking (e.g., where goals focus on helping students take on the identity or perspective of a scientist, participate in the practices of a scientist, or view a challenge or conflict from a different perspective) strongly synergize with popular game structures and genres, particularly genres involving 3D immersive worlds and environments. These approaches to games for learning are incredibly powerful and promising as discussed by Gee (2007), Squire (2011), Barab and colleagues (2009), and others.

Thus, we are certainly not arguing that all research and design of games for science learning should focus on new genres. We are arguing, however, that exploring new genres is imperative for learning goals that focus on engaging students in interpreting, translating, and manipulating formal disciplinary representations, relationships, and models.

Is this an example of school transforming games? Or is this an example of games transforming school? We would argue that our earlier versions of SURGE were examples of school transforming games without much transformation in the other direction. While the games were certainly being transformed, the science curriculum was not measurably changed. Students may have found their physics unit to be more novel and engaging, but the learning goals and larger organization of the science curriculum remained the same. Students might have understood the physics ideas more deeply than in the traditional curriculum, but the focus was still on isolated domain knowledge and relationships.

Disciplinarily-integrated games, however, could transform schools in the sense of changing the learning goals and integrating learning across the science curriculum around modeling and scientific practices. The traditional science curriculum (i.e., the curriculum experienced by 98% of students currently in school) focuses almost exclusively on disciplinary knowledge and relationships within individual disciplinary domain units. Each domain unit is instantiated as a two or three week curricular unit that is largely disconnected from all the other domain units.

The Next Generation Science Standards (NGSS) focus on scientific practices, but few vehicles exist for systematically integrating the science curriculum around these practices. Disciplinarily-integrated games could provide a powerful vehicle for integration. Disciplinarily-integrated games could be situated throughout the academic year to focus on meaningful connections between epistemic and representational model types and modeling strategies. Thus, while the science curriculum has certainly transformed disciplinarily-integrated games, our disciplinarily-integrated games are designed and positioned to dialectically transform the science curriculum.

ACKNOWLEDGEMENTS

The research reported here was supported by the National Science Foundation through grants and through the Institute of Education Sciences, U.S. Department of Education, through grant R305A110782 to Vanderbilt University. The opinions expressed are those of the authors and do not represent views of the National Science Foundation, the Institute Education Sciences, or the U.S. Department of Education. We would also like to thank Lee Druce for his work laying out the figures and pages of this manuscript.

REFERENCES

Aarseth, E. (2007). I fought the law: Transgressive play and the implied player. *Proceedings of DiGRA 2007 Conference: Situated Play.* Retrieved from http://www.digra.org/wp-content/uploads/digital-library/07313.03489.pdf

Adams, D. & Clark D. B. (2014). Integrating self-explanation functionality into a complex game environment: Keeping gaming in motion. *Computers and Education, 73*, 149-159.

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education, 53*(1), 74–85.

Barab, S. A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S., & Warren, S. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education and Technology, 18*, 305-320. http://dx.doi.org/10.1007/s10956-009-9171-5

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity: *Journal of Research in Science Teaching, 24*, 385-395.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition and Instruction, 24*(4), 467-563.

Clark, D. B. (2011). "Games and simulations bridging intuitive and formal understandings of physics." Keynote commissioned by the Gordon Research Conference on Visualization, Smithfield, RI.

Clark, D. B., & Martinez-Garza, M. (2012). Prediction and explanation as design mechanics in conceptually-integrated digital games to help players articulate the tacit understandings they build through gameplay. In C. Steinkuhler, K. Squire, & S. Barab (Eds.), *Games, learning, and society: Learning and meaning in the digital age* (pp. 279-305). Cambridge, UK: Cambridge University Press.

Clark, D. B., Martinez-Garza, M., Biswas, G., Luecht, R. M., & Sengupta, P. (2012). Driving assessment of students' explanations in game dialog using computer-adaptive testing and hidden Markov Modeling. In D. Ifenthaler, D. Eseryel, & G. Xun (Eds.), *Game-based learning: Foundations, innovations, and perspectives* (pp. 173-199). New York, NY: Springer.

Clark, D. B., Nelson, B. C., Chang, H.-Y., Martinez-Garza, M., Slack, K., & D'Angelo, C. M. (2011). Exploring Newtonian mechanics in a conceptually-integrated digital game: Comparison of learning and affective outcomes for students in Taiwan and the United States. *Computers & Education, 57*(3), 2178-2195.

Clark, D. B., Sengupta, P. & Virk, S. (2016). Disciplinarily-integrated games: Generalizing across domains and model types. In D. Russell and J. Laffey (Eds.), *Handbook of research on gaming trends in P-12 education*. Hershey, PA: IGI Global. doi:10.4018/978-1-4666-9629-7

Clark, D. B., Sengupta, P., Brady, C., Martinez-Garza, M., & Killingsworth, S. (2015). Disciplinary integration in digital games for science learning. *International STEM Education Journal, 2(2),* 1-21. http://dx.doi.org/10.1186/s40594-014-0014-4

Clark, D. B., Tanner-Smith, E., & Killingsworth, S. (2015). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research.* http://dx.doi.org/10.3102/0034654315582065

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2 & 3), 105-225.

diSessa, A., Hammer, D., Sherin, B. & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior 10*, 117–160.

diSessa A., & Minstrell, J. (1998). Cultivating conceptual change with benchmark lessons. In J. G. Greeno and S. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 155-187). Mahwah, NJ: Erlbaum Associates.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Research Council Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441-467. http://dx.doi.org/10.1177/1046878102238607

Gee, J. P. (2007). *Good video games and good learning: Collected essays on video games, learning and literacy (New Literacies and Digital Epistemologies)*. Peter Lang Publishers.

Hammer, D. (1996). Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions? *The Journal of the Learning Sciences, 5*(2), 97-127.

Hegedus, S., & Roschelle, J. (Eds.). (2013). *The SimCalc vision and contributions: Democratizing access to important mathematics*. New York, NY: Springer.

Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory: A response to March 1995 critique by Huffman and Heller. *The Physics Teacher, 33*(8), 502–506.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*(3), 141–158.

Kafai, Y. B. (2008). *Beyond Barbie and Mortal Kombat: New perspectives on gender and gaming*. Cambridge, MA: MIT Press.

Kaput, J. (1992). Technology and mathematics education. In D. Grouws (Ed.) *Handbook on research in mathematics teaching and learning* (pp. 515–556). New York, NY: Macmillan.

Killingsworth, S., Clark, D. B., & Adams, D. (2015). Self-explanation and explanatory feedback in games: individual differences, gameplay, and learning. International *Journal of Education in Mathematics, Science and Technology. 3*(3), 162-186. Retrieved from http://ijemst.com/issues/3_3_1_Killingsworth_Clark_Adams.pdf

Kinnebrew, J., Killingsworth, S., Clark, D. B., Biswas, G., Sengupta, P., Minstrell, J., Martinez-Garza, M., & Krinks, K., (in press). Contextual Markup and Mining in Digital Games for Science Learning: Connecting Player Behaviors to Learning Goals. *IEEE Transactions on Learning Technologies.*

Koster, R. (2004). *A Theory of fun for game design* (1st ed.). Paraglyph Press.

Krinks, K. D., Sengupta, P., Clark, D. B. (2013, April). *Conceptual change in physics through use of digital games*. Paper presented at the annual conference of the National Association for Research in Science Teaching, Rio Mar, Puerto Rico.

Krinks, K., Sengupta, P., & Clark, D. (2015, April). *Benchmark lessons: Integrating modeling with games for learning physics.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Kuo, M.-J. (2007). How does an online game based learning environment promote students' intrinsic motivation for learning natural science and how does it affect their learning outcomes? *Digital Game and Intelligent Toy Enhanced Learning, 2007. DIGITEL '07. The First IEEE International Workshop on* (pp. 135-142). http://dx.doi.org/10.1109/DIGITEL.2007.28

Lehrer, R, & Schauble, L. (2006a). Cultivating model-based reasoning in science education. In RK Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 371–388). Cambridge, UK: Cambridge University Press.

Lehrer, R, & Schauble, L. (2006b). Scientific thinking and science literacy: Supporting development in learning contexts. In W Damon, RM Lerner, KA Renninger, & IE Sigel (Eds.), *Handbook of child psychology* (6th ed., Vol. 4). Hoboken, NJ: John Wiley and Sons.

Martinez-Garza, M., Clark, D. B. & Nelson, B. (2013). Digital games and the US National Research Council's science proficiency goals. *Studies in Science Education, 49,* 170-208. http://dx.doi.org/10.1080/03057267.2013.839372

Mayer, R. E. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity and temporal contiguity principles. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 183–200). New York, NY: Cambridge University Press.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world.* New York, NY: Penguin Press.

Minstrell, J. (2000). Student thinking and related assessment: Creating a facet assessment-based learning environment. In N. Raju, Pellegrino, J., Jones, L., & Mitchell, K. (Eds.), *Grading the Nation's report card: Research from the evaluation of NAEP*. Washington, DC: National Academy Press.

Minstrell, J. Anderson, R. & Li, M. (in press). Diagnostic instruction: Toward an integrated system for classroom assessment. In R. Duschl & A. Bismack (Eds.*), Reconceptualizing STEM education: The central role of practices*. New York, NY: Routledge Taylor & Francis Group.

Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*, 369–383. http://dx.doi.org/10.1002/tea.3660240408

Munz, U., Schumm, P., Wiesebrock, A., & Allgower, F. (2007). Motivation and learning progress through educational games. *Industrial Electronics, IEEE Transactions on, 54*(6), 3141-3144. http://dx.doi.org/10.1109/TIE.2007.907030

Parnafes, O., & Disessa, A. (2004). Relations between types of reasoning and computational representations. *International Journal of Computers for Mathematical Learning, 9*, 251-280. NL: Kluwer Academic Publishers.

Pelletier, C. (2008). Gaming in context: How young people construct their gendered identities in playing and making games. In Y. B. Kafai, C. Heeter, J. Denner, & J. Y. Sun (Eds.), *Beyond Barbie and Mortal Kombat: New perspectives on gender and gaming*. Cambridge, MA: MIT Press.

Pickering, A. (1995). *The Mangle of Practice: Time, Agency, and Science*. Chicago, IL: University of Chicago Press.

Roschelle, J., Kaput, J., & Stroup, W. (2000). SimCalc: Accelerating student engagement with the mathematics of change. In M.J. Jacobsen & R.B. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning* (pp. 47-75). Hillsdale, NJ: Earlbaum.

Roschelle, J., Knudsen, J., & Hegedus, S. (2010). From new technological infrastructures to curricular activity systems: Advanced designs for teaching and learning. In M. J. Jacobson & P. Reimann (Eds.), *Designs for learning environments of the future: International perspectives from the learning sciences* (pp. 233-262). New York, NY: Springer.

Sengupta, P. & Clark, D. B. (in press). Playing modeling games in the science classroom: The case for disciplinary integration. *Educational Technology*.

Sherin, B. L. (2001). How students understand physics equations. *Cognition and Instruction*, *19*(4), 479-541.

Squire, K. (2006). From content to context: Videogames as designed experience. *Educational Researcher, 35*(8), 19.

Squire, K. (2011). *Video games and learning: Teaching and participatory culture in the digital age*. New York, NY: Teachers College Press.

Squire, K., Barnett, M., Grant, J. M., & Higginbotham, T. (2004). Electromagnetism supercharged!: Learning physics with digital simulation games. *Proceedings of the 6th International Conference of the Learning Sciences* (pp. 513–520).

Tatar, D., Roschelle, J. and Hegedus, S. (2014). SimCalc: Democratizing access to advanced mathematics, *International Journal of Designs for Learning*, *5*(2), 83-100.

White, B. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10*(1), 1-100.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.