A Study of Educational Simulations
Part I - Engagement and Learning

W. K. Adams, S. Reid, R. LeMaster, S. B. McKagan, K. K. Perkins, M. Dubson and C. E. Wieman

Abstract

Interactive computer simulations with complex representations and sophisticated graphics are a relatively new addition to the classroom, and research in this area is limited. We have conducted over 200 individual student interviews during which the students described what they were thinking as they interacted with simulations. These interviews were conducted as part of the research and design of simulations for the Physics Education Technology (PhET) project. PhET is an ongoing project that has developed over 60 simulations for use in teaching physics, chemistry, and physical science. These interviews are a rich source of information about how students interact with computer simulations and what makes an educationally effective simulation. We have observed that simulations can be highly engaging and educationally effective, but only if the student's interaction with the simulation is directed by the student's own questioning. Here we describe our design process, what features are effective for engaging students in educationally productive interactions and the underlying principles which support our empirically developed guidelines. In a companion paper we describe in detail the design features used to create an intuitive simulation for students to use.

Table of Contents

Introduction

Technology is becoming increasingly important in today's classroom and has been integrated in a variety of ways; however, computer animations and interactive simulations are among the most common. This popularity is partly due to the fact that simulations are quite easy to introduce into a curriculum. Such simulations have been developed on a large scale by a group of educators working together – e.g. Physlets (Christian & Belloni, 2001) – and on a small scale by individual educators who would simply like to communicate an idea visually to their students. Textbooks now regularly include DVDs or a URL to websites with a library of various simulations. While many educators find it appealing to use simulations in their classroom, very little research has been done to determine if simulations improve a student's understanding of or enthusiasm for science and how simulations can be designed and used most effectively.

Available simulations use a wide variety of appearances, controls, graphics, interactivity, and design principles, often guided only by the designers' preferences or ease of coding. Little is known, however, about design principles and features that are important for optimal student use and understanding. In this paper we present an extensive analysis of student use of simulations, including comparisons of multiple incarnations of a single simulation. This analysis has led to an empirically determined and tested set of design principles based on our observations of student use. This work also provides a rich body of data for the study of student thinking and learning while using simulations, and it has clearly demonstrated that a carefully designed and tested simulation can be a very powerful educational tool (Finkelstein, Adams, Keller, Perkins, Wieman, and the PhET Team, 2006; Finkelstein, Adams, Keller, Kohl, et. al, 2005; Finkelstein, Perkins, Adams, Kohl, and Podolefsky, 2005)

This research focuses on identifying which characteristics make a simulation effective or ineffective through the use of extensive think-aloud student interviews using simulations. This paper is part I of a two part series. This paper will focus on the simulation design process; what are desirable features – those that are found to be important for encouraging students to discover and understand physical relationships- which include and specific methods to provide engaging ways to help students 'discover' the desired learning goals of the simulation; how our design guidelines were developed; and, the underlying principles that support the guidelines. The second paper (Adams, Reid, LeMaster, McKagan, Perkins, Dubson and Wieman, 2008) describes more specific details on interface design, specifically features that make a simulation engaging and easy to use, types of controls that are intuitive for the student, effective use of representations, the impact of different types of "help" and the impact of even small amounts of irrelevant information.

Background

The context of this research is the PhET project (Perkins et al., 2006; The PhET Team, 2006a), an ongoing program to develop an extensive suite of freely available online simulations for teaching and learning physics, chemistry and physical science. These simulations create animated, interactive, game-like environments that emphasize the connections between real life phenomena and the underlying science while making the visual and conceptual models of expert scientists accessible to students. Currently there are about 60 PhET simulations.

The primary target for these simulations was originally college undergraduates with a wide range of science backgrounds and interests, and this is the population that has been studied in our research. However, these simulations appear to be useful for a surprisingly large range of students and are now extensively used in many high schools as well as some middle schools. In addition, we have received numerous anecdotal reports of grade school students finding them highly engaging and have observed physics graduate students learning new physics by playing with them. An interesting area of future research would be the study of how the findings we report here might depend on the age and background of the student beyond the levels explored in this work.

Simulation Design Process

To understand how our studies have been carried out, it is first necessary to understand the PhET development process. Our process for creating and evaluating a simulation begins with the selection of the simulation design team consisting of between three or four individuals including a programmer, at least one content expert, and at least one student interface expert. The design cycle (Figure I) starts with the content and student interface experts creating a detailed initial layout for the simulation based on the learning goals of the simulation and the research base – research in education and cognitive science relating to the topic plus the current PhET design guidelines. The first set of stu dent interviews are conducted, once all team members feel the simulation is clear, accurate and engaging. These interviews always reveal interface weaknesses, resolve interface questions that were not agreed upon by the team, and often reveal pedagogically undesirable (and occasionally unexpected desirable) features and subtle programming bugs. Subsequent revisions are made, and if they are extensive, a further set of interviews are conducted. These interviews are not only used to improve the particular simulation but continue to improve our research base. More recent interview results are finding much smaller problems than the interviews conducted on simulations that were written two years ago, indicating that our empirically developed design principles are working. After interviews establish that the desired engagement and learning is being achieved, the simulation is used in a classroom setting where student use is observed and informally evaluated.

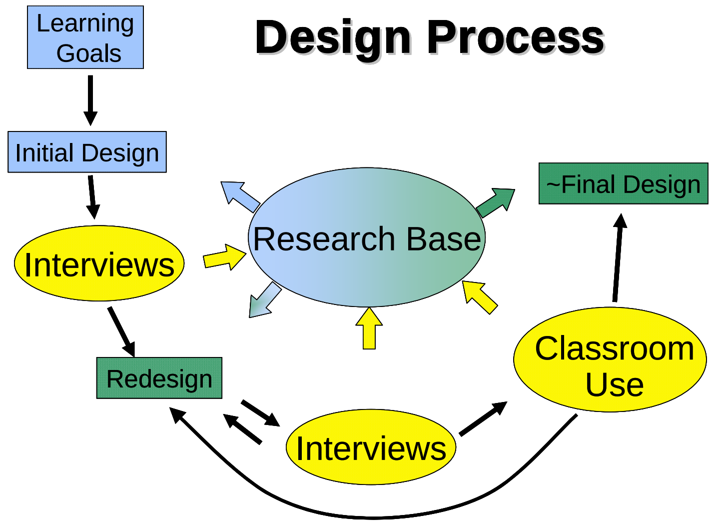


Figure I: Design Process – Flow chart showing the simulation design process.

Interview Methodology

Over the past three years we have conducted more than 200 simulation interviews with 89 different students covering 52 of 60 simulations. Student interviewees are volunteers that are typically non-science majors. For the more advanced quantum simulations, we also interview physics majors. For each simulation, we typically interview a diverse group of four to six students consisting of equal numbers of male and female students, and a representative share of minority students. Care is taken to acquire a selection of students with a wide range of academic performance. We also attempt to interview students who have not yet received formal instruction on the ideas covered by the simulation.

When we began this work, we were unsure if representative information could be gained from the observation of such a small number of students per simulation; however, in the sorts of issues explored here, we have found a high level of consistency. For example, the interface problems that arose in interviews were problems for most if not all of the interview subjects. In fact, when six students were interviewed on a single simulation, the last two interviews very rarely provided new useful information regarding interface design. Responses related to physics conceptual issues, which are not the primary focus of this paper, were more varied but still show considerable consistency. In this regard, these interviews are rather different from typical educational or psychological research. Because the results are so consistent, ever such small sample numbers produce quantitative results in that they allow one to make accurate predictions. For example, in addition to these formal interviews, we have also observed numerous groups using the simulations for the first time including students in both physics and chemistry courses, physics graduate students, and high school and college teachers. The observations of use in those settings have been quite consistent with the predictions from the corresponding student interview results; the rare exceptions are noted in the appropriate sections below and in Part II.

The PhET interviews are typically conducted with the same set of students during a given semester. If major revisions are required for a particular simulation and multiple iterations of interviews are needed, we find additional volunteers so that we can observe students' first encounter with the simulation. This type of protocol is required because we observe profound differences in how students interact with a simulation once they have been instructed on its use or have had opportunities to use it on their own, compared to seeing it for the first time.

Our standard interview protocol includes the following: in the first interview with a particular student, the interviewer begins by getting to know the student, asking about their background, career and major choices, and courses as necessary to break the ice. Once the student relaxes, and in all subsequent interviews with that student, the simulations are explored in a think-aloud style format. With this approach, the students are asked to talk out-loud while they investigate the simulation. The simulation explorations are structured one of two ways: 1) The student is asked prediction-type conceptual questions (where the student describes their understanding of an idea/concept before seeing the simulation) to guide their interactions. Then, after, or more often while, interacting with the simulation, they are allowed to revise their answer; or 2) The student is simply asked to explore the simulation freely without a guiding question.

In all cases, interview results were useful for determining: the level of student engagement promoted by the simulation; if controls are intuitive and easy to use; if any definitions or ideas are misunderstood or missed altogether; and if there is any extra information that is distracting the student from the simulation's learning goals. Using the prediction-type questions is useful in evaluating the simulation's ability to help students learn particular concepts. Additionally, these questions focus the students on the particular aspect of the simulation that we are currently interested in evaluating. These questions are imperative for evaluating the more involved simulations, because these simulations are sufficiently complex, with multiple levels of controls and presentations, that fully exploring the simulation could take hours. The unguided explorations are useful for determining how people interact with the simulations on their first encounter and for evaluating how students explore and understand the less involved simulations.

All interviews are video-taped and detailed summaries are prepared for each interview, describing the student's interactions with the simulation. These summaries identify any interface difficulties encountered during exploration as well as indicate what concepts were understood/misunderstood and at what level. When studying simulation design, these summaries are more meaningful (as well as much shorter), than detailed transcripts, because the manipulation of and references to the simulation plays such a large role in the communication between the student and interviewer that it is not possible to fully understand the interview simply from a transcript. A short section of an interview transcript and an individual summary for the same interview can be found online (Adams, 2003). After interviews on all subjects have been completed, a detailed summary of the individual summaries is compiled and distributed to the design team. The research results described in this paper draw largely from these detailed summaries. However, seven hours of interviews have been transcribed and coded for research questions (Perkins, Adams, Finkelstein and Wieman, 2004) that require this level of qualitative analysis. To ensure the interpretations and summaries are robust and not subject to interviewer bias, a number of tapes were observed, coded and interpreted independently. For a short section of coded transcription we determined the inter-rater reliability initially to be 95%, but after discussion and revision of the coding scheme, it increased to nearly 100%.

Some interviews were conducted with both an interviewer and an observer or the tapes were independently observed. Interview summaries were then completed independently by each and checked for consistency. This was done with a total of six different interviewers/observers and forty-six hours of interviews. These independent evaluations showed high levels of consistency except when there was a lack of advanced physics mastery by the interviewer or observer. In these cases, less expert interviewers/observers incorrectly interpreted some subtle misconceptions by the student being interviewed as correct physics learning. We found that a mastery of physics at the master's level, preferably with teaching experience, was necessary for interviewing on beginning and intermediate level simulations, while Ph.D. level mastery was desirable for interviewing on student learning and understanding with the more advanced simulations, such as quantum mechanics.

Although it is not the purpose of this paper, the fact that it is necessary for interviewers to have a very high level of content mastery illustrates a general feature that we have observed for sophisticated simulations of the type discussed here, where there are complex behaviors that depend on multiple variables. These simulations will routinely engage students to raise questions and explore the underlying science topic of the simulation in great depth, and it is this depth of understanding and exploration that requires interviewers with expert knowledge. Similarly, designers also need to have expert content knowledge for the same reason.

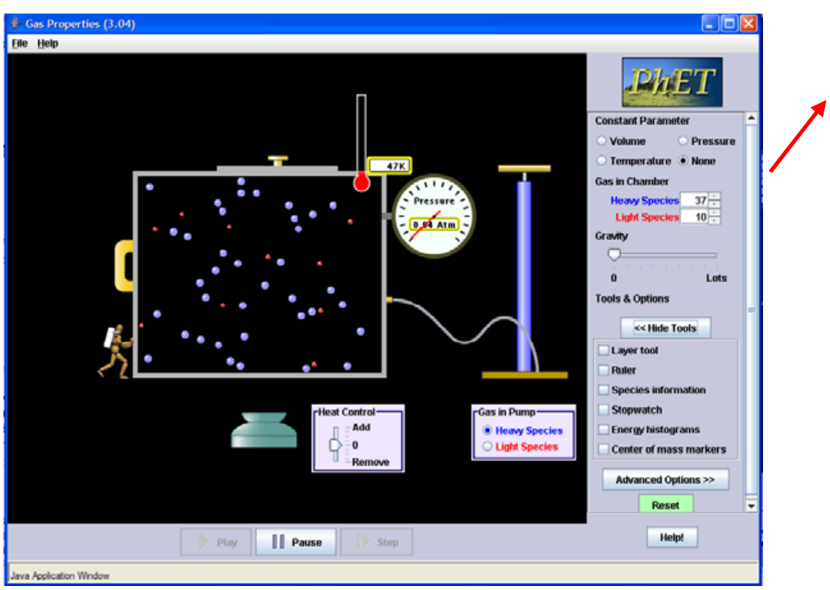
Interview Results

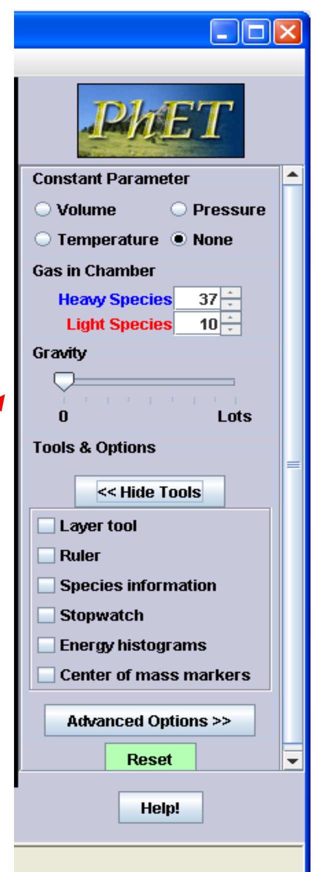
The following discussions of design features focus on the specific simulations and interviews where the problems were discovered, the potential solutions were explored, and the desirable design features first confirmed. We have checked the validity of these design features and principles in subsequent interviews with new simulations; however, in the interest of brevity, discussions of these follow up interviews will not usually be provided in these papers when the interviews merely confirmed the previously observed results. All general conclusions presented here have been confirmed with interviews on at least several simulations.

Encourage Exploration

We consistently observe that engaging students in thoughtful exploration of the simulation is necessary for improving students' understanding of the concepts. When in *engaged exploration*, students are posing questions and seeking answers by observing the results of their own interactions with the simulation and making sense of what they see. In this section we focus on the interface design (Figure II) aspects that enhance educational effectiveness. Engaging the students can be accomplished by having the students use the simulation in the appropriate context, such as with a well designed homework assignment or laboratory activity. However, we

Figure II – Interface Design: The black region is the play area containing the representations of physical objects that students can manipulate themselves and observe the effects of their actions instantly. The grey area on the right is the control panel which contains radio buttons, sliders and text boxes for adjusting various parameters and in the lower half of the control panel there are several tools for the students to use while working in the play area.





also strive to encourage the students to spontaneously ask themselves questions ("why does that happen?") that they can subsequently answer by exploring with the simulation. We see a variety of factors that influence students' engagement with and learning from the simulations, including: the *interactivity* of the simulation; the presence of *little puzzles;* strategically placed but limited text such as *legends and labels;* and features that make the simulations *fun* to play with. We have also found surprising negative influences from prior "understanding" of the topic.

Our work relating to effective engagement techniques is consistent with and builds on previous research of video games. Work done by Malone (1981) has found that video games are intrinsically motivating because they include balanced challenges, fantasy and an optimal level of informational complexity to create curiosity. Malone (1981) found that challenge is created by including personally meaningful goals and uncertain outcomes. All challenges must be attainable to foster self-esteem rather than discouraging users. His research also found that while fantasy was required, it is difficult to create fantasy that is appealing to a wide range of users. For example, most of the videogames that he studied had a scenario that appealed to only one gender. He defined a fantasy-inducing environment as one that evokes "mental images of things not present to the senses or within the actual experience of the person involved" (Malone, 1981 pg 360). Mental images can be either of physical objects or social situations. Finally, curiosity is evoked by an environment that is novel and surprising, but not completely incomprehensible.

It is well established that clear goals are important for motivation. Our designs only deal with this indirectly, by attempting to make the primary goal/challenge that of being able to understand the phenomenon portrayed by the simulation. We have seen that by relating to the real world and using suitable animation and interactivity, the desired curiosity is encouraged. In the simulations that students investigate on their own time, as described below in the *Fun* section, there are fairly clear goals such as navigating a maze or creating novel circuits and exploring their behavior. These goals obviously contribute to the attraction. During our interviews, we have not found these goals or the simulations themselves to be gender biased (possibly due to the balance of men and women on the PhET team). However, we are implicitly assuming that most simulations will be used in the context of an educational setting where teachers will primarily provide the scaffolding and goals for the simulation use. In the interviews, the guiding question or the interview itself provides this structure. Because these goals and uses will vary widely with the teacher and level of student, we have, in most cases, avoided constraining their use by not building highly specific tasks or goals into the simulation. For examples of activities created by teachers for use with the PhET simulations please see the PhET Activity Database (The PhET Team, 2006b).

Animation and Interactivity

- *Students notice animated features first; however, when only observing and not interacting, students do not ask questions or make new connections.*
- *User control of every perceived potentially significant parameter is valuable.*
- *Limiting students control over certain items must be done carefully.*

One of the most obvious benefits of presenting a concept using a simulation is that the simulation is animated. Interviews show that anything in motion draws the student's attention first; but, if the simulation simply demonstrates the motion of an object, students rarely develop new ideas or insights. In these cases, students seem to accept what they are seeing as a fact, but very rarely engage in understanding the meaning of the animation. In contrast, when students see an animated motion instantly change in response to their self-directed interaction with the simulation, new ideas form and they begin to make connections. Students create their own questions based on what they see the simulation do. With these questions in mind, they begin to investigate the simulation in an attempt to make sense of the information it provides. In this way, students answer their own questions and create connections between the information provided by the simulation and their previous knowledge.

A series of interviews on ‘Radio Waves’ illustrates the value of interactivity coupled with animation. The initial version of the simulation began with the full oscillating electric field emanating out from the transmitting antenna (see Figure III). At the beginning of these interviews students had very negative reactions to this mode that they would tend to watch passively. Students commented: “*Full field view doesn’t make sense to me*” or “*I don’t like this view*”. Students then watched the simulation and attempted to correct the predictions they had made before opening the simulation, without any interaction with the simulation. Their descriptions were incorrect, very superficial, and/or simply based on bits of prior knowledge. For example, one student said that electric fields move in a circular direction. To answer the question of how a radio signal is transmitted students said: “*by radio waves*” or “*I don’t know, I never thought about it*”. Once the students began interacting with the simulation and switching views a few times, they all began to appreciate the full field view and made comments such as “*this makes sense, the wave has to go out in all directions or my radio would only*

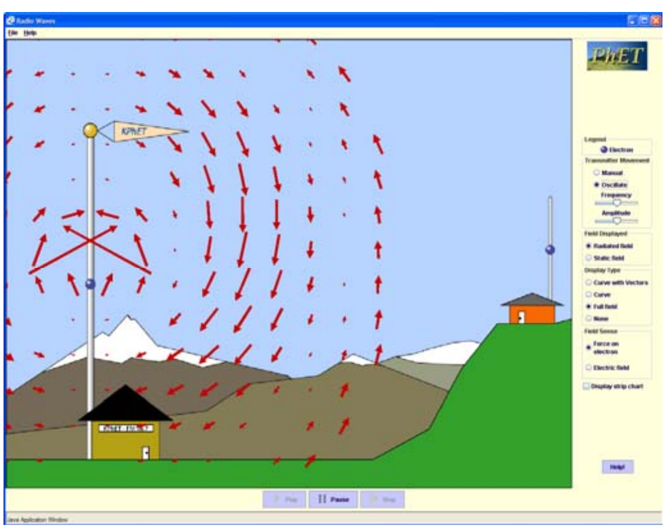


Figure III – An early version of ‘Radio Waves’. When the simulation first opens, the transmitting electron moves up and down along the antenna producing an electromagnetic wave that radiates out filling the screen with oscillating red arrows.

work in one spot” or “*this is my favorite view*”. In all of the interviews, we’ve seen that interactions, guided by the student’s personal questioning, are what make simulations an effective learning tool. Students engage in exploration and sense-making only *after* they begin to interact with the simulation. This finding suggests that the educational value of animations without interactivity is quite limited.

When making the simulation interactive, the choice of parameters that can be manipulated is important and several factors must be taken into account. By limiting the parameters that can be changed and by emphasizing particular controls, a simulation scaffolds and guides student thinking. While it is useful to provide scaffolding by allowing only relevant parameters to be adjusted, we find that it is sometimes also valuable to allow adjustment of parameters that students commonly think might have an effect on the phenomena, even if they do not. If students are limited to interacting with only the features that have an effect, their misconceptions about which parameters actually will/will not change a situation cannot be addressed. For example, ‘Projectile Motion’ allows students to manipulate many parameters including air resistance, mass and surface area. Many students believe a heavier object will have more air resistance. Since the parameter is available to change, even though they ‘know’ the answer, students try the parameter and are surprised by the result – learning from this control.

Because students learn that PhET simulations allow them to interact with the important objects on the screen, not allowing an object to be manipulated by the user also creates questioning and ideas. In ‘Radio Waves’, after users played with the transmitting electron, several tried to move the receiving electron and realized they could not directly manipulate its motion. See Figure III. Many asked, “*why doesn’t this one move?*” They investigated further and found that the only way to move it was to send a radio wave from the transmitting antenna. This lack of control sparked questioning that led to a better understanding of the effect a radio wave has on an electron.

On the other hand, disabling controls for non-physical reasons can lead to incorrect ideas because students attribute meaning to the ability to manipulate controls. We have seen many examples of this behavior. In 'Quantum Tunneling', for instance, the radio button that allows the user to view the incoming and reflected waves separately was initially disabled for wave packets and enabled for plane waves – implemented by graying out the radio button in wave packet mode. This restriction was not for any physical reason, but because it would have been difficult to program for wave packets and would have relatively little pedagogical value. In interviews, students became very frustrated that they could not use this control and tried to figure out the reason that it was grayed out for wave packets. In the current version, rather than graying out the control, it simply disappears in wave packet mode. Later interviews showed no problems with this implementation.

Little Puzzles/Clues (questions/answers that stimulate the student to explore and learn)

One effective way we've found to encourage exploration is to include little puzzles or tantalizing clues that stimulate the user to form questions that relate to the learning goals of the simulation. Many of these questions are easily answered by interacting briefly with the simulation and not only create understanding but increase confidence and motivation. Other questions are more involved and take some time to answer but are answerable by interacting with the simulation.

- *When students encounter small features that they do not understand, they will explore how interacting with that feature changes the simulation until they can create a working definition of the feature.*
- *Legends and control labels help students build connections, and then when they interact with the simulation, they learn a working definition of the term on the label.*
- *Multiple Representations - Simulations that have multiple views of the same item, such as beam view and photon view, facilitate further understanding and connections about the idea.*
- *Exploration is not always productive – elements that distract students' exploration in irrelevant directions must be avoided.*

Students quite often encounter a word in the simulation that they don't know. Typically when this happens, students play with the control that is labeled with the unknown word and subsequently create a working definition for the word. Frequency and amplitude were words students were unable to clearly describe before exploring the 'Sound Waves' simulation. After playing with the simulation, students correctly described the meaning of these words using visuals from the simulation. A few weeks later, during interviews on 'Radio Waves', the same students used the visual descriptions from 'Sound Waves' to describe frequency and amplitude. These non-science majors then used 'Radio Waves' to create an accurate working definition of an electric field. (See Figure III)

When using 'Nuclear Physics', students did not know what the abbreviations on the nuclei such as ^{235}U meant. In response, a small legend that included a thumbnail of the nuclei with the label Uranium 235 beside it was added to the top of the control panel. After this simple addition, further interviews with new students were conducted. All of these new students found the legend and used the correct terms to describe the nuclei from that point forward. In 'Signal Circuit' interviews, students were asked what was moving around the circuit. Only one student correctly identified the little blue dots as electrons. Once the other three students discovered that un-checking a box that said "show electrons" made the blue dots disappear, they corrected their responses given about 10 to 15 minutes earlier, to identity that it was electrons that were moving around the circuit. In each of these examples the text is very limited. We've found, as described in the *Help* section found in Part II, that legends and control labels can become useless if they contain too many words.

Multiple representations that can be clearly and easily switched between, are also an effective way to get the students to ask questions about what they are seeing and to interact with the simulation. For example, in 'Color Vision' both beam view and photon view are offered for the light going from a lamp to Howie Hue's eye. During interviews, students were unsure about the photon view until they switched to beam view. Once they explored these two views, all students stated with confidence that they are the same thing. A student exploring these views for the white light said: "*One just shows the tiny little photons so you can see the separate colors.*"

Although encouraging exploration is necessary for learning, it is also possible to create features in the simulations that encourage exploration and student thought that is not productive. As an example, in an earlier version of 'Color Vision' a pulsing brain inside of Howie Hue's head was used to represent that Howie's brain was interpreting colors that entered his eyes. This was displayed when a "Show Inside" checkbox was checked. Every student who was interviewed on this simulation spent a fair amount of time playing with the check box and looking at the brain carefully while changing the other parameters of the simulation. All students were looking for some feature of the pulsing brain to change if the appropriate parameters were selected. Some students quickly determined that there was no conceptual value to the pulsing brain feature *"Obviously this guy has a brain."*, and others had to be told by the interviewer that there was no significance to the brain *"K, the, well the brain doesn't seem to be doing anything when I show the color, so I don't know if….really why it's there"*. This pulsing brain feature encouraged exploration and thought from all students interviewed; however, no further understanding of the concepts was garnered from this exploration.

Fun

- *When the simulations are fun, students enjoy playing with them. The Flash simulations, and Java simulations with similar characteristics, draw students to them.*
- *When simulations look boring or intimidating, students are not drawn to playing or they are afraid they will break them.*
- *Features can be so much fun to play with that students are distracted from learning.*

To engage students in exploration, students should want to play with the simulations. Every feature adds to a student's cognitive load and so needs to have educational purpose. The example of the pulsing brain is one of a number of examples we have seen where features violated this rule. This point must also be considered in how one designs fun into simulations. If a feature is fun, it must also create learning. There seem to be two levels of fun. The first level is the surface appearance; if the simulation is fun-looking (game like, colorful and cartoon-like, interesting graphics, non-threatening…) students want to try it out. When student users browse the PhET website, they consistently choose Flash simulations over Java simulations. Extensive discussions with users have provided vague answers such as, *"they look more fun"*. We hypothesize that the bright colors, 3-d look of the controls, and simple cartoon-like features are what attract users to the Flash simulations. Too crude and simplistic graphics, or an overly complex appearance, are both perceived as less fun. We've seen a positive response to subsequent Java simulations that incorporate many of the same characteristics of the Flash simulations, supporting our hypothesis.

We've also seen in interviews that when a simulation is first opened up, if it appears too complicated or has unfamiliar features, students are less likely to engage without interviewer intervention. If the simulation has the look of a lab workbook – meaning lots of numbers and detail such as closely spaced graph lines and abstract representations of the physical features – then students are not only less interested but actually uncomfortable about using such simulations. They are afraid they will break them and make comments about "*…[not knowing] how to use stuff like that*." If they don't know what physical item is being depicted on the screen, they are very uncomfortable manipulating that item.

The next level of fun moves beyond merely stimulating initial interest to repeated voluntary use of the simulation. There are several simulations that students regularly say they play with during their leisure time, including 'Electric Field Hockey', 'Circuit Construction Kit (CCK)', 'The Maze Game', 'Travoltage', 'Energy

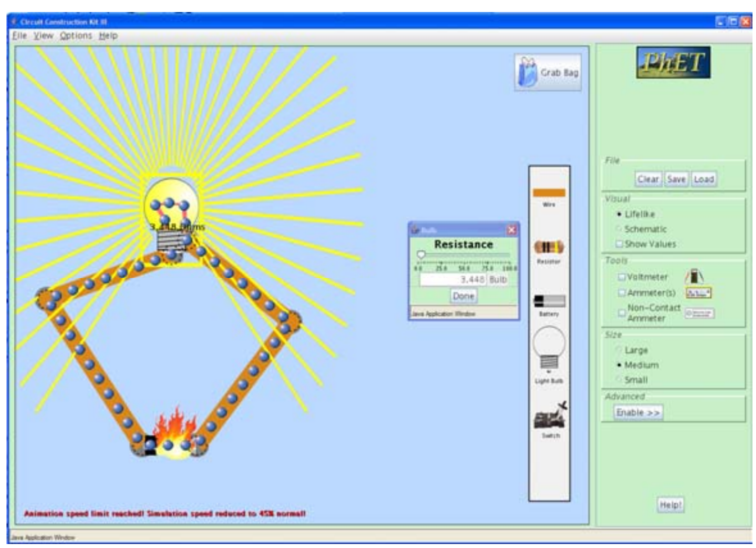


Figure IV – 'CCK'. Fun engaging features are included such as the brightness of the bulb changes as students adjust resistance and voltage or the battery can catch on fire.

Skate Park' and 'Ramps'. In each of these simulations we've worked to successfully add game-like features that create a fun environment for exploration. Interviews show that the addictive features of these simulations now focus on the central physics concept of the simulation. For example in 'CCK' as current is increased through a light bulb, it becomes brighter and when too much current runs through a battery, it catches on fire (Figure IV). In 'The Maze Game' a student can adjust one of three parameters (position, velocity or acceleration) while attempting to direct a ball through a maze. An annoying pop sounds if a barrier is hit and a satisfying music clip is played when the goal is reached. These little features create environments where students spend their free time becoming familiar with the concept of electric charge or the differences between velocity and acceleration.

However, there is a fine line between a fun simulation that stimulates learning and fun features of a simulation that distract the student from learning. 'Ramps' provided an example of the latter. In this simulation, bar graphs represent different forms of energy including kinetic, potential and thermal. With continued friction, the thermal energy bar increases and eventually extends off the screen. For this reason, we added a way to reset the thermal energy. When the user clicks "Cool Ramp" a firefighting dog comes out and sprays water from a fire hose on the ramp to cool it off. Originally, each time the button was clicked, a new dog appeared. Students reacted by seeing how many firefighting dogs can fit on their screen at once – a fun, but unproductive, game. Even teachers who were in a workshop learning about the simulations engaged in the same unproductive behavior of adding as many firefighting dogs as possible. Interviews showed that a suitable balance was achieved by allowing only a single dog to appear. This approach preserved the pedagogical value of using the firefighting dog to stimulate the students to think about how the ramp was heating up and connect that to the physics of the conversion of mechanical energy to thermal energy, while avoiding the danger that simply creating more firefighting dogs became the focus of attention.

Credibility of Simulations

- *For engaged exploration to occur, students must believe the simulation.*
- *Student's level of skepticism is related to their level in school.*

One important question is: How skeptical are students about the correctness of the simulations? The answer is particularly relevant when the simulation gives results that students do not expect and hence have the most to learn from. We have found students to be quite trusting of the simulations, e.g. *"These are really smart people. I'm sure they don't make mistakes."* However, our observations have found that students' level of skepticism is related to their level in school. Non-science majors are very trusting while students in quantum mechanics are quite skeptical. There have been a few cases where the quantum mechanics instructor points out a bug in the simulation during class. Afterwards students were observed to typically take the simulation less seriously. Similar reactions were encountered during quantum mechanics interviews. If the interviewer said that a simulation was still under development or might have bugs, students were much more likely to attribute what they did not understand to programming bugs. On the other hand, introductory students have been disturbingly trusting of simulations, even to the point of attributing significance to behaviors observed under conditions where they were explicitly told the simulation did not function properly.

This high level of trust is demonstrated by a task associated with the first version of 'Energy Skate Park' (formally 'Energy Conservation Kit'). During the first semester of physics for non-science majors, we added short simulation questions to the end of the student's weekly homework assignments. The questions covered material that the students had not yet been introduced to in class. One such task asked the students "If a person wanted to lift a 1 kg rock to a height of 20 meters on Earth or to the same height on the moon, will it require more work (Energy input) on the moon or on Earth? 91% of students correctly predicted that it requires more work to lift the rock on the Earth. After playing with the simulation only 17% of the students believed it took more work on the Earth. Upon close inspection of the simulation we discovered that the default mass for the object on Earth was 1 kg and on the moon it was 1650 kg. After finding the opposite result from what they expected, students trusted the simulation (or at least believed this was the answer we were looking for) and answered accordingly.

Performance Mode

- *Students who do not believe they already know the relevant ideas, are more likely to explore a simulation and use it to learn. Students who think they should understand the topic of a simulation often use it much less effectively and learn much less from it.*

The profound effect of students' self-expectations is illustrated by the multiple interviews that have been done on the 'Radio Waves' simulation. This topic is not important for simulation design, but it is very important for simulation use and testing. These and similar interviews revealed that if students think they understand material prior to the interview and in this case, have previous experience with the simulation, they lapse into what we call "performance mode" – equivalent to behavior associated with performance goals as described by Dweck (1989). In this mode students have difficulty exploring and learning effectively from the simulation. They try to recall what they know and make excuses for their lack of answers. Students who have not covered the simulation in class have very different expectations and are much better at exploring the simulation to develop understanding.

In the fall of 2003, we conducted two sets of interviews on 'Radio Waves' with four students from the first semester of physics for non-science majors. The following semester, we interviewed on 'Radio Waves' again using students enrolled in the second half of this two course sequence. Three of the spring interviewees had taken the first semester of the sequence (one had also been interviewed in the fall), while the fourth student had enrolled in the second semester of the sequence without taking the first semester. The first set of interviews in the fall showed the simulation to be quite successful. These non-science majors gained an impressive conceptual understanding of an electric field from the simulation, before they had ever encountered the term "electric field" in class. Later in the fall semester the concept of an electric field and the 'Radio Waves' simulation were covered as part of the course.

During the spring interviews, a very different pattern was observed. Three of the students interviewed struggled with the simulation, rushed through it, and never really effectively engaged in learning from the simulation. The two students who had taken the first semester course but had not participated in the fall interviews reacted similarly to the 'Radio Waves' simulation. In one case, once the interviewer started asking questions about radio waves, the student quickly decided he didn't understand, and rather than exploring with the simulation to find answers, he responded that he'd aced the homework in the fall and couldn't understand why he didn't get it now. In the other case, as soon as the student was asked the first question, she responded that she had missed a lot of class during this section. Every time she was asked a question, she said, "*I haven't had lecture on this*". When asked further questions, she simply said, "*I just don't understand this stuff*". She kept apologizing, gave fast answers, and the interviewer was quite unsuccessful getting her to look at the simulation and think about what it was depicting. When talking about other simulations in previous interviews, this student appeared to be one of the most intelligent and resourceful.

The third student was an interview subject both during the fall and spring semesters. She was able to work out a reasonable definition of an electric field during her fall interview, but in the spring she responded differently. When the spring interview began, she said she liked this simulation and that it was one of her favorites as she opened it. By the end she said she didn't like it anymore. She was confused and couldn't believe she didn't remember all of it. When attempts were made to guide her, she'd just say, *"I should know this*" and didn't appear to really think it through. She just kept trying to remember and became increasingly frustrated. At times during the interviews, these three students would begin to engage with the simulation, but as soon as they'd make a connection with something in their memory, they'd slip back into unproductive *performance mode*.

In contrast, the fourth interview student in the spring, who had appeared to be the weakest during all previous interviews that semester, performed as well or better than the students had in the fall 'Radio Waves' interviews. This student had not taken the first semester of the course sequence, and so had never seen the 'Radio Waves' simulation nor had formal instruction on electric fields. This student began by saying he knew nothing about radio waves and was more relaxed than the others. When he started with the simulation he wiggled the electron and said "*it appears to be some sort of wave simulation but I haven't had lecture on this stuff so don't understand it*". He proceeded to carefully explore the simulation with only very minor encouragement from the interviewer. In fact, this interview was the first where he actually slowed down and explored. In prior interviews on other simulations, if he'd used the ideas in homework, he would generally rush through the simulation. It typically required a lot of intervention from the interviewer to get him to slow down, reflect, and explain in these previous interviews. When he didn't know something previously, he had tended to become frustrated and annoyed (more so than the other three). However, now working with 'Radio Waves' he took his time, didn't seem bothered if he didn't know something, and worked through most of the concepts very successfully. This level of engagement and learning was similar to the 'Radio Waves' interviews during the previous fall semester, before students had seen the topic in class.

Students often begin any interview that involves some familiar ideas in *performance mode,* explaining what they know. The more the students believe they know, the less they engage with the simulation and the greater their tendency to become tense and frustrated when asked questions they don't quite understand. When in *performance mode,* they move too quickly through the simulation for it to help them clarify their thoughts. The above 'Radio Waves' interviews are an extreme example of this problem since not only had the students had instruction on this topic; but, they also had experience with this simulation and thought they should know everything. They did remember a lot of useful information, but anything that was not completely clear frustrated them, and they were reluctant to slow down and learn from the simulation. In all other simulation interviews, it took only a short amount of time and occasionally a little prompting before students started exploring the simulation and making sense of the presentation provided by the simulation. During the quantum mechanics interviews with upper-level students, this transition into engaged exploration occurred quickly and without prompting. These students seem to realize that they are far from mastering quantum mechanics and in general have stronger meta-cognitive skills than the non-science majors who typically interview on the introductory simulations.

Discussion

In these interviews we find that nearly all the simulations, after suitable testing and revision, consistently result in a high level of learning in our diverse group of interview subjects. After a simulation interview, most students understand the concepts covered in the simulation well enough to explain them accurately and to use them to make accurate predictions about behaviors in the simulation. Students also often volunteer correct predictions or explanations about related real world phenomena. This level of understanding is far beyond what we have observed is typically obtained from the coverage of these concepts in a physics course. There are some reasons why simulations help student learning that are very obvious from our interviews and so shape our design characteristics – e.g. the ability to provide visual models. These reasons were described above or can be found in Part II, in the relevant sections. However, in this work we primarily focus on the somewhat simpler problem, namely *what* characteristics a simulation should have to achieve this impressive level of learning that we have observed. A detailed analysis of how and why simulations result in such learning will be the focus of future work.

The PhET Look and Feel

From these interviews we created the "PhET Look and Feel" (Adams, Perkins & Wieman, 2006), which the design teams now follow while creating a new simulation. During the first year of interviews, when the look and feel was still in the early development stages, student difficulties ranged from simulation usability to conceptual problems. These difficulties included problems such as interface design, help functions, tool placement, effective types of representations, and what types of features encouraged students to interact with and think about the simulation. Many interface problems and successes were found to be consistent from simulation to simulation, and thus informed our simulation design guidelines contained in the PhET Look and Feel. We would typically research particular aspects of the interface design in depth using multiple versions of the same simulation, and then utilize those results in designing subsequent simulations. Results from interviews on the subsequent simulations would then confirm or refine the design guidelines.

Interviews have also revealed three different levels of usability:

1. Non-intuitive –difficult to use even with instruction.
2. Semi-intuitive – easy to use after instruction and demonstration; and
3. Intuitive – easy to use with no instruction.

It is relatively easy to create a simulation that will be easy for a student to use after observing a demonstration. It is more difficult to create an intuitive simulation that requires no instructions; but, we have found that an intuitive simulation can be designed rather routinely (even for rather complex simulations) by following the now highly-refined PhET Look and Feel guidelines derived from our interview studies. Thus, our new simulations rarely have usability issues, and our current interviews focus primarily on a simulation's ability to engage the student and achieve the desired learning goals.

In this paper we described the *Encourage Exploration* section of the PhET Look and Feel, while the second paper, Part II, contains the larger part of the PhET Look and Feel that focuses on the features we have found to be successful at creating an *intuitive* interface as defined above. This second paper also contains extensive interview results to support each feature of the PhET Look and Feel.

Underlying Principles

Three major principles support nearly all of the desirable design features identified through our interview studies and are consistent with the literature. These include *Engaged Exploration,* the *Coherence Principle* (Clark & Mayer, 2003) and *Consistency*.

Engaged Exploration

- *When in engaged exploration, students are actively working to make sense of the information before them.*
- *Students are more easily engaged in the exploration of topics that include relatively unfamiliar science.*

We have found it particularly important to get the students involved in what we have labeled as *engaged exploration*. When in *engaged exploration*, students are posing questions and seeking answers by observing the results of their own interactions with the simulation and making sense of what they see. We have seen various reasons for students not to engage in exploring a simulation. A short, but far from exhaustive list includes: they have been interacting with the simulation for a very short time; they are unable to successfully figure out how to use the simulation; they are overwhelmed by the simulation and do not know where to start; or they believe that they are familiar with the content and attempt to quickly explain the scientific concepts to the interviewer simply using the simulation as a demonstration tool, rather than as a learning tool. The idea of *engaged exploration* is consistent with work by Minstrell and Kraus (2005) and Dweck (1989).

Coherence Principle

- *Adding interesting but unnecessary material to simulations can harm the learning process in several ways.*

Clark and Mayer's (2003) Coherence principle describes many of the simulation features that our interviews have shown are important. The empirically-based Coherence principle emphasizes the importance of having all elements (controls and visual cues) directly related to the learning goals of the simulation and excluding extraneous information. Clark and Mayer (2003) discuss how unnecessary information can interfere with learning in three ways: "*distraction* – by guiding the learner's limited attention away from the relevant material and towards the irrelevant material; *disruption* – by preventing the learner from building appropriate links among pieces of relevant material because pieces of irrelevant material are in the way; *seduction* – by priming inappropriate existing knowledge (suggested by added visual cues, sounds, or words), which is then used to organize the incoming material." Our research has repeatedly confirmed the need to limit simulation features to only those items that are directly necessary to convey the learning goals of the simulation.

Consistency

- *Users' interpretation and use of simulations depends heavily on their prior experiences.*

As described in the *Interview Methodology* section, interviews were conducted with students who had various levels of experience with PhET simulations. Users experienced with one or more simulations were able to start using a new simulation more quickly than completely inexperienced users. Experienced PhET users also 'know' what a particular representation should look like and bring what they've discovered from one simulation to the next. However, experienced users were bothered by seemingly minor inconsistencies between simulations, even if the subject of the simulation was quite different.

Further Work

The PhET interviews have provided a rich source of ideas for further studies of student thinking and learning with interactive simulations. We see students clearly achieving impressive levels of mastery on a variety of difficult topics in physics. It will be interesting to study in more detail what are the topic specific questions they formulate in working with the simulations, how do students address these questions, and how does that result in their understanding? By exploring these issues with a number of students, it will provide a greater understanding of topic specific learning and how better to teach these subjects, with or without the use of simulations.

A second area of potential research is based on the observations of how students used the ideas they developed using 'Sound Waves' to understand 'Radio Waves'. We are currently building on this to explore the broader issue of analogical scaffolding in creating understanding (Podolefsky & Finkelstein, 2006). A third interesting area is the use of gesture by the students while using and discussing simulations. The use of gesture while interacting with simulations was analyzed and coded in order to help interpret the interviews (Adams, 2004). It was seen that there was a decrease in rate of gesture while using simulations, and that students generally use deictic gesture (indicating an object or person by pointing to where they are or have been) while using the simulations. Instances where students use lexical forms of gesture (smooth, continuous shapes in space indicating places, objects or ideas) are indicative of either students drawing on prior knowledge, or if the gesture mimics the simulation, the simulation is not quick enough in demonstrating the necessary animation. These observations support the notion that the simulations can be considered an extension of gesture, and suggest that analysis of gesture can be a useful tool for analyzing student interactions with simulations, and how they are using simulations to construct meaning.

Conclusion

We have carried out extensive interview studies on the student use and learning from interactive simulations for teaching physics. We find overwhelming evidence that simulations that suitably incorporate interactivity, animation, and context can provide a powerful learning environment where the students productively engage with and master physics content. However, we find that this can only be achieved by following an extensive set of principles for design and layout as contained in the PhET Look and Feel. Here we have included only one section of the PhET Look and Feel, *Encourage Exploration*, while the sequel to this paper contains the detailed specific design guidelines along with relevant interview results for creating an intuitive simulation including layout, representations, tool use and help functions. This work reveals many design pitfalls that can result in simulations not achieving the desired educational effectiveness. Finally, this work demonstrates the importance of testing educational simulations carefully with the desired target users.

Acknowledgements

We would like to thank Danielle Harlow, Noah Podolefsky, and Stephanie Fonda who conducted some of the interviews whose results are incorporated in this paper. We also thank Noah Finkelstein and the other members of the Colorado Physics Education Research group for many useful discussions. We are pleased to acknowledge support of this work by the University of Colorado, the National Science Foundation, the Kavli Operating Institute, and the Hewlett Foundation.

References

Adams, W.K., (2004). Gesture with Interactive Computer Simulations. [On-line] Available: http://phet.colorado.edu/web-pages/publications/Gesture.pdf

Adams, W.K. (2003). Sample transcript and summary from a 'Radio Waves' interview. http://phet.colorado.edu/web-pages/publications/Int Paper Transcipt_Summary.pdf

Adams, W.K., Perkins, K.K. and Wieman, C.E., (2006). *PhET Look and Feel* Retrieved November 23, 2006, from University of Colorado, Physics Education Technology Web site: http://phet.colorado.edu/web-pages/publications/PhET Look and Feel.pdf

Adams, W.K., Reid, S., LeMaster, R., McKagan, S.B., Perkins, K.K., Dubson, M. and Wieman, C.E. (2008). A Study of Educational Simulations Part II – Interface Design. *Journal of Interactive Learning Research, 12.*

Christian, W. and Belloni, M., (2001). *Physlets: Teaching Physics with Interactive Curricular Material.* New Jersey: Prentice Hall, Inc.

Clark, C. and Mayer, R., (2003). *E-learning and the Science of Instruction* (pp 111-129). San Francisco, California: Pfeiffer.

Dweck, C., (1989). Motivation. In Lesgold, A. and Glaser, R. (Eds.), *Foundations for a Psychology of Education* (pp 87-136). New Jersey: Lawrence Erlbaum Associates.

Finkelstein, N., Adams, W., Keller, C., Perkins, K., Wieman, C. and the PhET Team, (2006). High-Tech Tools for Teaching Physics: the Physics Education Technology Project. *Journal of Online Learning and Teaching, 2,* 109-121.

Finkelstein, N. D, Adams, W. K., Keller, C. J., Kohl, P. B., Perkisn, K. K., Podolefsky, N., Reid, S. and LeMaster, R., (2005). When learning about the real world is better done virtually: a study of substituting computer simulations for laboratory equipment. *Physical Review, Special Topics: Physics Education Research, 1,* 010103.

Finkelstein, N.D., Perkins, K. K., Adams, W., Kohl, P., Podolefsky, N., (2005) Can Computer Simulations Replace Real Lab Equipment? In Heron, P., Marx J. and Franklin, S. (Eds.), *2004 Physics Education Research Conference* (pp 101-104). New York: American Institute of Physics Conference Proceedings.

Malone, T., (1981). Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science 4,* 333-369.

Minstrell, J. and Kraus, P., (2005). Guided Inquiry in the Science Classroom. In Donovan, M. S. and Bransford, J. D. (Eds.), *How Students Learn: History, Mathematics, and Science in the Classroom* (pp 475-513). Washington, D.C.: The National Academies Press.

Perkins, K. K., Adams, W., Finkelstein, N. D., Dubson, M., LeMaster, R., Reid, S. and Wieman, C.E., (2006). PhET: Interactive Simulations for Teaching and Learning Physics. *The Physics Teacher 44,* 18-23.

Perkins, K.P., Adams, W.K., Finkelstein, N.D., and Wieman, C.E., (2004). Learning Physics with Simulations: The Role of Interactivity, Animation and Context. *American Association of Physics Teachers Summer Meeting: Sacramento, CA.* Retrieved November 23, 2006, from University of Colorado, Physics Education Technology Web site: http://phet.colorado.edu/web-pages/research.html

The PhET Team, (2006a). PhET Activity Database [Computer database] Available at: http://phet.colorado.edu/new/teacher_ideas/browse.php Colorado: Physics Education Technology Project.

The PhET Team, (2006b). PhET Interactive Computer Simulations [Computer software] Available at: http://phet.colorado.edu Colorado: Physics Education Technology Project.

Podolefsky, N. S. and Finkelstein, N. D., (2006). Use of analogy in learning physics: The role of representations. *Physical Review, Special Topics: Physics Education Research*. 2, 020101