How does the type of guidance affect student use of an interactive simulation?

Archie Paulson, Kathy Perkins, and Wendy Adams
Department of Physics, University of Colorado

For students using carefully designed computer simulations, we find that the type of guidance provided to the student during interviews strongly influences the amount of independent exploration and inquiry in which they engage. What would be considered relatively little guidance for typical educational activities appears to limit the benefits of a complex simulated environment. In a study of student interaction with a PhET simulation, we find that when students follow a carefully-designed activity that gently guides their exploration and sense-making of the important concepts, they actually exhibit less engagement with the simulation and do less exploration and discovery than students who receive only minimal guidance.

I. INTRODUCTION

Since the pioneering work of Piaget [1] in the 1970's, constructivist theories of learning have prevailed in educational research. It is generally agreed that students must construct their own understanding as active sense makers within their framework of existing knowledge [2]. To accomplish this, students must be actively engaged with the content and able to learn from that engagement [3].

In this paper we investigate the amount of guidance that should accompany a well-designed, interactive computer simulation to elicit active engagement and learning. In a series of interviews with students using a physics simulation, we find that students learn more through self-guided exploration which only occurs if the interviewer provides minimal guidance. Furthermore, we are able to show that with a carefully designed activity students interact with fewer features of the simulation than in the minimal guidance case. The features that students choose to interact with is directly dependent on and limited by the content of the guidance. While one might ask if we're advocating "pure discovery" learning, we believe that discovery learning is quite different from working with a simulation with minimal guidance. The simulations, by design, provide considerable implicit guidance [4, 5]. The PhET simulations (described below) have been carefully designed and tested with students to included balanced challenges, obvious intriguing phenomena that students will attempt to figure out by engaging with the simulation. In this way students are able to "discover" what it is about the phenomena that is important and begin to create a mental framework of the concept as they acquire new knowledge.

Schwartz et al. [6] have seen a similar phenomena where they compare the tell-and-practice approach with the innovation-then-lecture approach. The tell-and-practice students are provided with a direct explanation of a topic followed by application exercises. The innovation students are first given carefully structured, but unguided, invention tasks followed by direct explanation. The innovation-first students exhibit improved transfer ability, greater appreciation of deep structure in a problem, and improved future learning. Schwartz et al. [7] attribute this improvement to the creation of a mental framework. The creation of this framework is preempted when students first receive the direct explanation.

Another important difference we observe between students who receive more or less guidance is their engagement, or "mode of exploration." Such a difference has been documented in other studies. For example, students can be "learning oriented" or "performance oriented" in their approach to a task, and in the latter case can become too preoccupied with possibly making errors or satisfying the instructor to freely learn [8]. Bing and Redish [9] also discuss how a student's "epistemic frame," or the game they engage in (e.g. math, sense-making), is highly sensitive to the context of the problem. How the problem is written can profoundly affect a student's response to it.

The PhET Interactive Simulations (PhET [10]) project has developed over 80 interactive, research-based computer simulations of physical phenomena that emphasize interactivity, animation, and real world connections. In the course of creating these simulations, the PhET project has conducted hundreds of interviews with students using them. Based on these interviews, Adams et al. [11] compared four different levels of interview guidance and found that engaged exploration of a sim generally occurs only when the student receives minimal guidance.

This paper reports on a controlled study to explore these findings more carefully using two specific types of interview guidance: Driving Questions and Gentle Guidance. "Driving Questions" uses one or two challenging conceptual questions to trigger a self-driven exploration, and "Gentle Guidance" uses a carefully designed activity that asks students to investigate particular controls or features of the simulation. In addition to studying how students interacted differently with the simulation under those two conditions, a third group had two questions omitted from the Gentle Guidance activity to determine if guidance would restrict exploration and affect what students see in the simulation.

II. METHOD

A. The sim

The sim used in this study is called "Faraday's Electromagnetic Lab." It was designed to facilitate learning about current induction in a coil of wire due to changing magnetic fields. The sim has five different configurations, organized in tabs across the top of the window. One of these tabs is shown in Figure 1. Typical of the PhET sims, Faraday's Electromagnetic Lab comes with no instructions but has many inviting, interactive features that provide immediate visual feedback to the user's actions. Even though the PhET sims come with nearly no help or instruction, students spend little time on false starts or fruitless searching. For example, when a magnet is dragged around the screen the magnetic field indicators in the background change direction and brightness; when the magnetic flux through the pickup coil changes, cartoon electrons in the coil move and the light bulb illuminates. Many other features are available in this sim, which may be seen (in color) and experienced by launching it from the PhET [10] website. Investigating the various tabs, the user can induce a current in a variety of ways: by moving a bar magnet, moving the pickup coil through a magnetic field, changing the field generated by an electromagnet, generating an alternating magnetic field by using an AC current source, or by driving a water-wheel which is attached to a bar magnet. After many interviews with students using the sim, it has been repeatedly observed that students who have no initial acquaintance with electromagnetic induction discover that changing magnetic fields can induce electric currents.

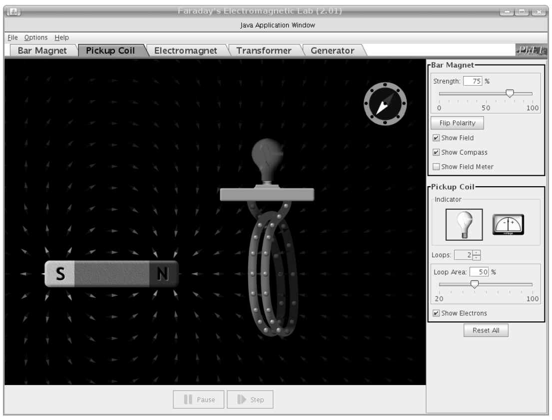

Figure 1: A screenshot of Faraday's Electromagnetic Lab. This sim has five tabs across the top; only the second tab is shown, where motion of the magnet induces a current in the pickup coil, causing the light bulb to illuminate. (The sim is in color.)

B. Interview methodology

For this study, one-on-one interviews were conducted with twelve undergraduate student volunteers at the University of Colorado at Boulder, selected for having little or no physics experience. In order to study the question of how much guidance is required for a student to fully explore the sim, interviews were conducted in the three different styles—Driving Questions (DQ), Gentle Guidance (GG), and Gentle Guidance with Missing Pieces (MP). Four interviews were conducted in each style for a total of twelve interviews.

The "Driving Questions" (DQ) interview style consisted only of asking the student two driving questions at the beginning of each interview but before seeing the sim: "Could a magnet affect an electric current? How or why?" and "What are some of the ways that you could make a magnet?" None of the students could satisfactorily answer these questions before using the sim. Students were then invited to "interact with everything available in the sim, and while you're doing that think out loud." Students in the DQ group explored the sim entirely on their own. They were asked only once at the outset to interact with everything, and after they started were only asked questions like "what are you doing now?" if they became quiet.

The "Gentle Guidance" (GG) interview style was "gently guided", in which the student's interaction with the sim was moderated by a carefully designed activity that was modified through a series of seven prior interviews. The GG activity consisted of a series of 17 questions that ask the student how different features affect each other or what the function of different controls are. It was designed with the intention of allowing exploration while guiding the student to all the clues necessary to state some general principle of Faraday's Law. Before the student began playing, s/he was asked the same two questions used in the DQ interview. The next 15 questions asked the student to use the simulation. Typical GG questions include "What does the Field Meter do?" and "How does motion of the magnet affect the electrons in the coil of wire?" Students in this group were asked to verbally answer the questions in the GG activity while "thinking out loud." (Note that the term "gentle guidance," coined by Adams et al. [11], refers to guidance which invites the user to interrogate certain aspects of the sim but is not "cookbooky".) The entire GG activity is available online[12].

The "Missing Pieces" (MP) interview style is identical to the GG style but with two questions were omitted. The two omitted questions asked the students to interact with three specific elements of the sim—elements that were not mentioned elsewhere in the guidance. The MP style was included in order to observe if students would explore elements of the sim that were not explicitly called for in the activity questions.

In all cases, students used the sim until they decided they were done, usually for 30 to 50 minutes. The interviewer was careful to not intervene, speaking only when necessary to give reminders to "think aloud" or to ask for clarification of a student's comment. When, occasionally, a student in a GG or MP group would ask about the meaning of a question, the interviewer would answer as simply as possible. This first 30-50 minutes was the only period of time in which the groups (DQ, GG, and MP) received different treatment.

In all three groups (DQ, GG, and MP), when the students felt they were done exploring the sim, they were asked the following question: "In all the cases shown in the sim, there is one principle (called *Faraday's Law*) that describes what makes the light bulb turn on. Try to state this principle as generally (and as simply) as possible." This would prompt the student to play a little more with the sim, reviewing what they had seen, as they answered the question to their satisfaction. Student behavior during this time did not vary between groups as much as it did during their exploration.

Finally, putting the computer aside, students were asked to answer several questions in a follow-up activity that used real equipment (a bar magnet, compass, wire coil, battery and galvanometer). Most of the follow-up questions involved directly analogous actions to the sim, such as predicting the motion of the galvanometer needle when the magnet was passed through the wire loop. A couple of questions probed the issue of current induction when the wire loop was parallel to the magnetic field lines (an issue that was not treated in the sim). The final question of the follow-up activity asked the student how a bicycle light could be constructed with a bulb, wire and magnet. The follow-up activity is also available online [13].

C. Data Collection

Video tapes of each student interview were reviewed and coded to count the number of interactions with various elements of the simulation. Verbalizations and actions were also coded that reflected a student's level of engagement with the sim.

Interviews were reviewed and coded to count numbers of student comments and actions that reflected their level of engagement with the sim. Among student verbalizations, the number of questions posed to themselves, and the number posed to the interviewer were coded and counted. Also coded were events that caused the student to transition to consider a different aspect of the sim: was it prompted by the GG/MP questions or was it unprompted (subject only to the student's desires)? And finally, we coded the student actions when they were apparently confused by something they saw; these codes included (1) continued investigation of the confusing feature, (2) transition to investigation of a different but related feature, (3) seeking help from the GG/MP activity, and (4) simply giving up and moving on.

Two interviews were coded separately by another researcher with the same rubric to test the consistency of the coding. The entire sim-interaction portions of the interviews were checked for coding reliability. In a first round of reliability testing, there was some discrepancy in the coding. The coding rubric was subsequently discussed and refined, leading to a new reliability test with 90% agreement between coders. The rubric for the coding is available online [14].

To quantify the amount of interaction each student had with the sim, the components of the sim were broken down into individual elements, where an element was one specific interaction with the sim (for example, moving the voltage slider on the electromagnet). For each sim element, a student's interaction with it was placed in one of three categories: the element was never mentioned, the element was noticed in only a cursory way, or the element was noticed and explored to some extent. For example, if a student noticed that the electromagnet had a slider, but decided not to explore its effects in the sim (perhaps because she was concerned with answering a different question at the time), that element would be considered to have received only cursory notice. If a different student noticed the same slider, and then moved it back and forth and appeared to recognize its effect on other aspects of the sim, his use would be considered full exploration for that sim element.

Twenty nine unique sim elements were coded. (Note that there are more that 29 unique aspects of the sim, but only those that were interacted with by at least one student were coded.) Of these 29 elements, 17 were referred to directly in the GG activity. For example, the GG activity asked the students what effect was seen by changing the strength of the bar magnet so the magnet's strength slider is one of these 17. The remaining 12 elements were not referred to in the GG activity (for example, using the "flip polarity" button for the bar magnet). Of the 17 elements referred to in the GG activity, three were intentionally omitted from the MP activity.

III. RESULTS AND DISCUSSION

This study's results highlight several differences between the DQ and the GG (and MP) groups. The verbalizations made while using the sim show a different "mode of investigation" depending on type of guidance. The DQ students were engaged with the simulation, exploring via their own questioning while the GG and MP interviews tended to lead to less scientific behavior and little independent exploration. The GG and MP activities also limited students' exploration of the sim, as demonstrated by reduced number of sim elements explored by the fact that the MP group almost entirely missed the three sim elements that were not explicitly mentioned in their questions.

A. Differences in Exploration

Before comparing the quantitative measures, we describe the most apparent difference between the groups that received the GG or MP activity and the DQ group: their mode of exploration during the interview, summarized with examples in Table I. The consistent differences between the groups were striking, and provided the original inspiration for this study. The students' mode of engagement is fairly easily seen through their verbalizations while interacting, as they were encouraged to "think aloud" throughout the interview.

There was a clear difference in what might be called the subject's "mode of investigation" of the sim. The DQ group only had the driving questions, but they investigated the sim just as long as the GG and MP groups. They were self-guided while using the sim, posing and answering their own questions in response to sim feedback from their actions. The GG and MP groups' "mode of investigation," on the other hand, was generally driven by the activity, resulting in what might be called a "student mode" in which answering the questions (verbally) takes priority over understanding what's happening.

These differences were primarily discerned from comments made during the interview, comments that showed consistent differences between groups. Comments from the DQ group reflected moments of surprise, curiosity or investigation such as "why did that happen?" The GG and MP groups would make comments such as "is that good enough for question two?" One student even asked "Do I get to play around first" before answering the question, possibly an indication of the student finding the appropriate "mode" for the interview.

Consistent differences were also observed when a student decided to move on to a new sim element—for example, when they decided to change tabs. The DQ students would move on when they had satisfied themselves with the element they had been considering. This would happen when they either answered a question that they had posed to themselves, or when they decided they needed more information to understand the behavior they saw. GG and MP students, on the other hand, would cease exploration when the activity referred them to a new sim component, with no indication that they were considering the need to satisfy his or her own understanding.

A final consistent difference in the mode of exploration was observed in students' responses when they were confused by something they saw in the sim. Students in the GG and MP groups would tend to refer back to the most recent question in the activity, even if it was unrelated to the issue that confronted them in the sim. The GG and MP students would frequently turn their attention away from apparent incongruities in the sim—features which were often included to make students think—and back to the question that they felt it was their task to answer. DQ students appeared content to continue exploring the confusing point until they either came to some conclusion or decided to postpone its resolution if further exploration was required (if, for example, answering a related but smaller question first was helpful).

Table II shows coded interview behavior to support the qualitative observations in Table I. Counts of coded verbalizations show the DQ students asked more questions of themselves (DQ:16.7/40, GG:5.1/27), while the GG and MP students were more often asking the interviewer (DQ:1/40, GG:14/27)—even though the interviewer would refuse to answer questions. There was an average of 40 coded verbalizations per interview in the DQ group and 27 in the GG and MP groups ("other verbalizations" included expressions of surprise and answers to questions—both indicators of engagement and sensemaking). While all sim transitions were unprompted for the DQ group, the GG/MP students primarily transitioned just after answering a question contained in the activity (16.2/21.5 transitions). When students showed confusion with what they were doing in the sim, DQ students almost always continued to explore the sim (18/19.5 times), while GG/MP occasionally looked to the activity for help (4.8/16 times) or chose to give up on the issue's resolution (4.2/16).

B. Exploration of Specific Sim Elements

Using the sim interaction measures described in section II C, the number of sim elements that each student either explored, just noticed without exploration, or did not mention at all were counted. For the 17 sim elements that were included in the GG questions, Figure 2[15] shows the average number of elements that students saw in the DQ and GG groups. While it is not surprising that the GG group saw and explored 90% of the sim elements that were in the GG questions, it is notable that the DQ group saw and explored just as many of these sim elements during their self-driven exploration of the simulation.

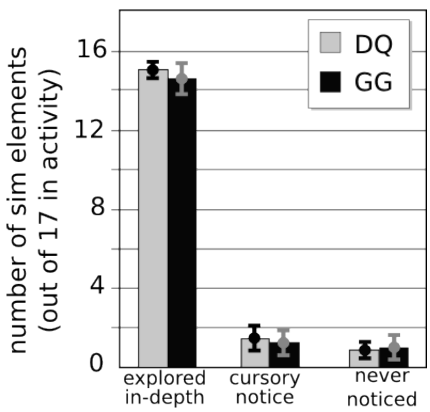

Figure 2: All students (DQ and GG) saw and explored the sim elements mentioned in the written activity (left). Bars in the middle and at the right indicate a just-noticed element or an unmentioned element (respectively). Error bars show the standard error of the mean.

The other 12 sim elements that were coded were not

Driving Questions (DQ)	Gently Guided (GG)
"mode of investigation":	
self-driven engaged exploration	guidance-driven "student mode"
typical student comments:	
"Oh, I wasn't expecting that." "I'm looking for something I can manipulate...." "I was looking around to see if it was an effect of having more wires." "So I bumped the strength of the magnet down, and the first thing that made me think of is 'when would that be desirable?'"	"Ok, continue?" "Is that sufficient for [number] 2?" "Is that enough for that?" "Do I get to play around first, before comparing?"
when do they move on to a new element?	
when they have satisfied themselves	when the guidance refers them to a new element
what do they do when confused?	
continue to explore sim	refer back to question in the guidance

Table I: Differences in exploration.

code	Driving Questions (DQ)	Gently Guided (GG)
▷ *verbalizations*		
number of questions posed to self	16.7 / 40	5.1 / 27
number of questions posed to interviewer	1 / 40	14 / 27
number of other verbalizations	22.3 / 40	7.9 / 27
▷ *number of tansitions to new sim elements*		
number of sim transitions without prompting	19.2 / 19.2	4.8 / 21.5
▷ *actions when confused*		
when confused, explored further	18 / 19.5	7 / 16
when confused, gave up	1.5 / 19.5	4.2 / 16
when confused, sought help from activity	0 / 19.5	4.8 / 16

Table II: Average numbers of times (per interview) that students performed some action. The number after slash is the total number (averaged over interviews) of coded verbalizations, sim transitions, or incidents of confusion.

referred to in the written questions, Figure 3 shows that DQ students consistently fully explored these 12 sim elements, on average exploring 92% of them. In contrast, the exploration by the GG students was much more limited. These students fully explored only 25% of the aspects of the sim which were not explicitly called for in their questions, and noticed but did not explore 50% of these sim elements. The last 25% of these 12 sim elements were not noticed at all by this group.

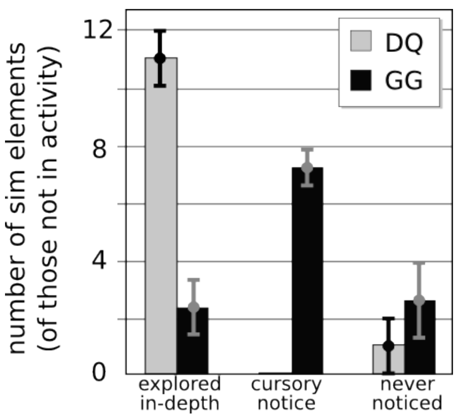

Figure 3: In addition to exploring as many of the sim elements that were specifically in the guidance (as did the GG group), the DQ students explored far more of the sim elements that were not mentioned in the GG guidance. (Color shading is the same as Figure 2; error bars show the standard error of the mean.)

Missing Pieces

Finally, to test the hypothesis that we could "keep students from exploring" by not mentioning something, we developed the MP activity: the GG activity minus two questions. There were three sim elements mentioned in these two questions. Comparing the three groups, Figure 4 shows the DQ and GG groups having substantially more sim interaction with these 3 elements than the MP group. Thus, the students who only had Driving Questions (DQ) did explore these elements, but those with questions used to invite investigation of the sim missed sim elements when they were not explicitly mentioned. It is worth noting that the data in the MP group were significantly altered by one anomalous student in that group; without special instruction, from the beginning one MP student began exploring the sim far beyond what was called for in the guidance, acting more like a DQ student. With her data point removed from the MP average, the results appear as in the column labeled MP*.

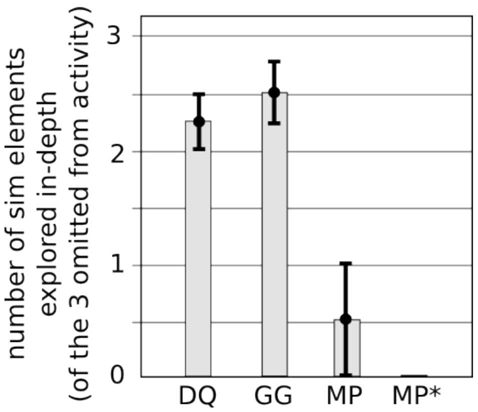

Figure 4: For the three elements omitted from the MP activity, most were explored by the DQ (Driving Questions) and GG (Gentle Guidance) groups. The MP group did not explore much beyond their guidance. All of the signal in the MP column comes from one anomalous student, who is removed in column MP*. (Color shading is the same as Figure 2; error bars show the standard error of the mean.)

C. Post questions

After using the sim, all students used real equipment to answer related questions in a follow-up activity. The grades for each group demonstrate that students with only driving questions learn as much as students with the gentle guidance that was carefully designed through student interviews.

IV. CONCLUSION

In the course of designing and testing the PhET simulations in hundreds of student interviews, consistent differences in student behavior toward guidance have been observed—observations that have been quantified in this study. A series of recorded and coded interviews has shown that students who are only given minimal guidance in the form of two Driving Questions (DQ) before interacting with the sim, engage in a productive self-driven exploration of the simulation much as a scientist engages in research. As demonstrated by the numbers of questions posed to themselves (Table II), these students are looking to themselves rather than to authority to decide how to explore a physical concept.

In contrast, we find that even Gentle Guidance (GG) takes students out of this engaged exploration and encourages a "student mode" (where students take on the role of answering only what they've been asked). Students following "gentle guidance" will not explore as much of the sim, will not engage as much in self-driven investigation, and appear to be effectively restricted to use only those aspects of the sim that are explicitly pointed to in their guidance.

The result that students were found to learn just as much with only two driving questions preceding sim exploration as when they followed the gentle guidance may seem contrary to the results of pure discovery learning. However, exploring sims is very different than pure discovery learning because sims are intrinsically structured. Here, the structure of the sim is providing enough scaffolding that the students can engage in productive exploration where their actions are directed by their own questions as they explore. This is a much more productive mode of exploration than instructor questioning because it allows the students to build a mental framework as they explore.

The measures of comparison used here, level of engagement, amount of sim that was fully explored and direct measures of learning, may not effectively capture all of the differences between the DQ and GG groups. Fully assessing the benefits of student driven exploration may require an assessment that also measures "preparation for future learning." [7] Such a study is being considered as a follow-on to the present work.

Acknowledgements

This work has been supported by the National Science Foundation, the William and Flora Hewlett Foundation, King Saud University, Microsoft Research, C. Wieman and S. Gilbert, the Nobel Foundation, the Kavli Foundation, and the University of Colorado at Boulder. We thank all members of the PhET project and the PER group at CU-Boulder, especially Carl Wieman for useful discussions and Noah Podolefsky for coding some of the interviews providing a test of inter-rater reliability.

[1] J. Piaget, *Science of education and the psychology of the child* (Orion Press, New York, 1970).

[2] J. Bransford, A. Brown, and R. Cocking, eds., *How People Learn: Brain, Mind, Experience, and School* (National Academy Press, Washington, DC, 2004).

[3] R. R. Hake, *Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses*, Science Education **66**, 64 (1998).

[4] W. K. Adams, S. Reid, R. LeMaster, S. B. McKagan, K. K. Perkins, M. Dubson, and C. E. Wieman, *A study of educational simulations part I - engagement and learning*, Journal of Interactive Learning Research **3**, 397 (2008).

[5] W. K. Adams, S. Reid, R. LeMaster, S. B. McKagan,

K. K. Perkins, M. Dubson, and C. E. Wieman, *A study of educational simulations part II - interface design*, Journal of Interactive Learning Research (in press) **4** (2008).

[6] D. L. Schwartz, J. D. Bransford, and D. Sears, in *Transfer of Learning from a Modern Multidisciplinary Perspective*, edited by J. P. Mestre (Information Age Publishing, Greenwich, Connecticut, 2005), pp. 1–51.

[7] D. L. Schwartz, R. Lindgren, and S. Lewis, in *Constructivist Theory Applied to Instruction: Success or Failure?* (Routledge, 2009).

[8] C. S. Dweck, in *Foundations for a Psychology of Learning*, edited by A. Lesgold and R. Glaser (Erlbaum, Hillsdale, NJ, 1989), pp. 87–136.

[9] T. J. Bing and E. F. Redish, *Using warrants as a window to epistemic framing*, PERC Proceedings pp. 71–74 (2008).

[10] PhET, *PhET project*, URL `http://phet.colorado.edu`.

[11] W. K. Adams, A. Paulson, and C. E. Wieman, *What levels of guidance elicit engaged exploration with interactive simulations?*, PERC Proceedings pp. 59–62 (2008).

[12] http://phet.colorado.edu/web-pages/publications/Paulson_etal_2009/gentle_guidance.pdf

[13] http://phet.colorado.edu/web-pages/publications/Paulson_etal_2009/followup_questions.pdf

[14] http://phet.colorado.edu/web-pages/publications/Paulson_etal_2009/coding_rubric.pdf

[15] Figure data available online at http://phet.colorado.edu/web-pages/publications/Paulson_etal_2009/figure_data.pdf