Student Use of Inquiry Simulations in Middle School Science

Elizabeth McBride
University of California Berkeley
Tolman Hall
Berkeley, CA, USA
bethmcbride@berkeley.edu

ABSTRACT
My research focuses on the integration of science and design through the use of interactive simulations and other scaffolding tools. I specifically look at patterns of use in interactive simulations. To conduct this research, I have developed a curriculum about solar ovens used by middle school students, during which students are guided by an online curriculum to design, build, and test physical solar ovens. This curriculum utilizes interactive simulations as a tool to help students plan the design for their solar ovens. I have evaluated scaffolding for the simulation steps, and plan to evaluate other patterns of student use, based on action log data.

Keywords
Interactive Simulations, Science Education, Inquiry, Log Data

1. RESEARCH TOPIC
My research focuses on the integration of science and design through the use of interactive simulations and other scaffolding tools. I specifically look at patterns of use in interactive simulations. I conduct this research in secondary schools, and work in collaboration with teachers. Through my dissertation work, I aim to answer the following questions:

- What types of use patterns in interactive simulations are beneficial for integrating science and design learning?
- How can we use tools to support integrated understanding in writing activities (e.g.,automated guidance)?

My work is situated in the learning sciences, using techniques from educational data mining and artificial intelligence to understand how students' activities impact their learning and how to improve the learning experience. Recently, I have used natural language processing to develop automated classifiers for multiple short response questions [6]. Using these classifiers, I plan to develop automated guidance for student writing during the curriculum, which will deploy during spring 2017. I have also studied student use of interactive simulations, using log data, feature engineering, and clustering to make sense of patterns (submitted to EDM 2017).

To conduct this research, I have developed a curriculum that is run using an online platform and offers students the opportunity to use interactive simulations while they design a physical artifact. In previous work, I have found that the simulation is beneficial, especially when students use it during the design phase of the curriculum [8]. My work has also been published in a variety of other conference venues [7, 10, 11, 9].

1.1 Curriculum
My research utilizes a curriculum about solar ovens that is run using the Web-based Inquiry Science Environment (WISE). During this curriculum, students design, build, and test a solar oven. They go through the design, build, test process two times to get an idea of how engineers iterate on their designs based on results from testing (Figure 1). This curriculum was designed using the knowledge integration framework [5]. The knowledge integration framework has proven useful for design of instruction featuring dynamic visualizations [14] and engineering design [1, 12]. The framework emphasizes linking of ideas by eliciting all the ideas students think are important and engaging them in testing and refining their ideas [5].

Students are allowed to use only a certain set of materials (e.g., tin foil, black construction paper, plastic wrap, Plexiglas, tape), in addition to a cardboard box they bring from home. Students use an interactive computer simulation to test the different materials in their oven. This simulation helps to elicit student ideas before they get to the building process, consistent with the knowledge integration framework. The testing portion of the project allows students to distinguish their ideas.

Throughout the project, students respond to short response questions about the choices they are making in their design and how their ovens work. This curriculum is unique, since it is guided by an online platform, but students also design, build, and test their solar ovens in a hands on portion of the project.

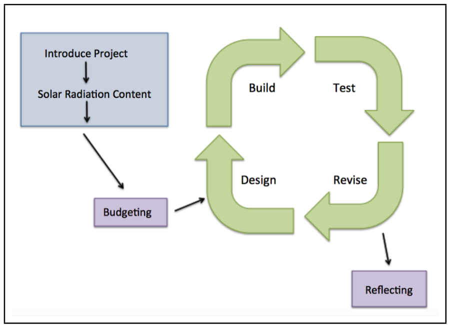

Figure 1: Outline of the solar ovens curriculum

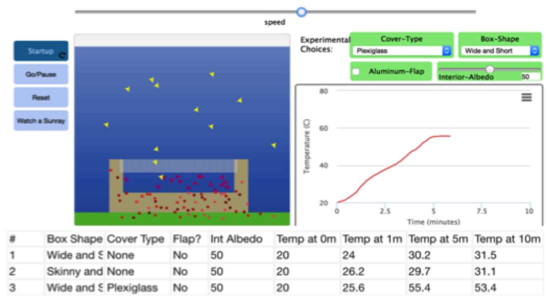

Figure 2: The interactive simulation used by students to test solar ovens and visualize energy transformation; below the table simulation is output from the automatically generated table

The curriculum takes between 10-15 class periods (45 minutes per class period). Students complete this project in groups of 2 or 3 students. Students also complete a pretest the day before the project begins and a posttest the day after completing the solar ovens project. Students do the pretest and posttest individually. The pre-/posttests measure student understanding of science concepts and practices.

1.2 Interactive Computer Simulation

The interactive simulation (figure 2) was built using NetLogo [15]. Students can manipulate the simulation in a number of ways. They can change the cover on top of the oven, whether or not there is a reflective flap on top of the box, the shape of the box (wide and short or skinny and tall), and the albedo (reflectivity) of the inside of the box. Students may also manipulate the speed at which the simulation runs. Once a simulation runs to the end of the graph (10 simulated minutes), a new row is added to the table below the visualization with the settings and results from the trial. If the students do not allow the simulation to run until the simulated 10 minutes finish, nothing is added to the table.

The scaffolds we developed for the interactive simulation are twofold; short response questions direct students to investigate capabilities and limitations of the simulation and an automatically generated table helps students to keep track of trials they have run. The table includes information about all of the settings used in that trial, as well as the results of the trial at certain time points (e.g. 5 minutes, 10 minutes).

2. PROPOSED CONTRIBUTIONS

Making sure students use interactive simulations to aid in learning is a difficult task. To try to encourage students to take advantage of these simulations during learning, various scaffolding methods have been used. Often, these scaffolds are implicit, or built into the system with the simulation [13]. For example, guiding questions are used with inquiry simulations to direct students' attention toward certain features of simulations [4]. Students are also often encouraged in science classes to run multiple trials and control variables between trials (only change one variable between trials). A control of variables strategy can help students to determine the effect of a single variable on a more complex system, although in some cases students may benefit from more exploratory strategies [12].

Using log files from student interactions with the curriculum and output from the automatically generated tables (simulation scaffolding), we use feature engineering to identify how students use the model and whether these uses have an impact on learning. I developed features that have to do with the control of variables strategy, such as the number of trials (rows) a student runs and the percent of those trials that are systematic. These types of techniques have also been used with more complex simulations and microworlds (e.g., [3, 2]). We use results from pre- and posttests to assess student learning in tandem with the log data from the curriculum.

The data in this work comes from 635 students across three schools and five teachers. During this study, students participated in a pretest and posttest (each lasting one class period), as well as the 2-3 week long curriculum. During the curriculum, students worked in teams of 2-3. These 635 students formed 255 teams.

3. RESULTS

I used pretest and posttest scores to understand the effect of actions with the simulation on learning. I then examined the role the number of rows of data a student generated using the table scaffolding on learning. I found that the number of rows generated in iteration 1 of the simulation is a significant predictor of individual posttest scores, when controlling for pretest scores and curriculum group ($b = 0.10$, $t(546) = 2.68$, $p < 0.01$). Next, I examined the impact of controlling variables on learning. I found that the number of *Control Of Variables (COV) Trials* run, however, is not quite a significant predictor of posttest score, when controlling for group and pretest score ($b = 0.06$, $t(546) = 1.63$, $p = 0.10$). In addition, using a dummy variable for conducting any *COV Trials* does not significantly predict posttest scores when controlling for pretest scores and group ($b = 0.005$, $t(546) = 0.13$, $p = 0.90$). Together, these results indicate that the control of variables strategy, while a good practice in science, is not as helpful for developing an understanding of the scientific principles at play in a simulation. More experimentation using the model is beneficial for developing a better understanding of the scientific concepts.

I then split the students up based on their actions during the simulation step (did not generate any rows in table, generated one row, generated 2 or more rows). I found that generating 2 or more rows in the table significantly predicts posttest scores, when controlling for pretest score and working group (b = 0.12, t(546) = 3.11, $p < 0.01$), though generating no rows or 1 row were not significant predictors. I also developed a variable, *Percent Systematic*, that is the percentage of the total rows a group generated that used the control of variables strategy. This variable has the ability to show more nuance in how students were employing the control of variables strategy, but was also not predictive in determining posttest scores, when controlling for pretest and group id (b = 0.05, t(508) = 1.32, $p = 0.188$).

There were also two short response scaffolding questions on the same step as the interactive simulation. I generated a variable based on the number of questions students answered (0, 1, or 2). This was predictive of posttest score, when controlling for pretest score and group id (b = 0.10, t(546) = 2.56, $p = 0.011$).

Overall, evidence suggests that students should be encouraged to experiment with the model and guided to produce at least two rows of data in the table to improve learning outcomes and use the short response questions. Perhaps changing more than one variable at a time in this type of environment indicates that students are spending more time thinking about possible outcomes. I have further examined this data using k-means clustering algorithms.

4. FURTHER QUESTIONS

I have finished the majority of data collection for my dissertation. I will conduct one more study during the spring of 2017, and there will be the potential for a follow-up study later. This is an important time for me to get feedback on my work, especially on the analysis of the action log data I have collected from over a thousand students. I will begin the writing phase of my dissertation work during the summer, and expect to complete my dissertation within the next 12 months.

During the doctoral consortium, I would like to discuss the following:

- How to assess patterns in student actions in interactive simulations (Tools and packages for doing this and assessment of what it means to be a meaningful pattern)
- Designing studies that integrate education theory and data mining
- Assessment of inquiry skills in online environments
- Use of event logs in online curriculum to assess student use of curriculum and how this can be used to assess learning in tandem with other methods

5. REFERENCES

[1] J. Chiu, P. Malcolm, D. Hecht, C. DeJaegher, E. Pan, M. Bradley, and M. Burghardt. Wisengineering: Supporting precollege engineering design and mathematical understanding. *Computers & Education*, 67:142–155, 2013.

[2] C. Conati, L. Fratamico, S. Kardan, and I. Roll. Comparing representations for learner models in interactive simulations. In *Artificial Intelligence in Education*, pages 74–83. Springer, 2015.

[3] J. D. Gobert, Y. J. Kim, M. A. Sao Pedro, M. Kennedy, and C. G. Betts. Using educational data mining to assess studentsâĂŹ skills at designing and conducting experiments within a complex systems microworld. *Thinking Skills and Creativity*, 18:81–90, 2015.

[4] C. Hmelo and R. Day. Contextualized questioning to scaffold learning from simulations. *Computers & Education*, 32(2):151–164, 1999.

[5] M. Linn and B. Eylon. *Science learning and instruction: Taking advantage of technology to promote knowledge integration.* Routledge, 2011.

[6] E. McBride, A. Dixit, and M. Linn. Submitted – using machine learning to automatically score text responses in middle school science projects. In *Proceedings of the 18th International Conference on Artificial Intelligence in Education*, 2017.

[7] E. McBride, M. Martinez-Garza, J. Vitale, and M. Linn. Middle school student use of interactive climate change simulations: Periods of observation and activity. In *Paper presented at the Annual Meeting of the American Educational Research Association, San Antonio, TX*, 2017.

[8] E. McBride, J. Vitale, L. Applebaum, and M. Linn. Use of interactive computer models to promote integration of science concepts through the engineering design process. In *Proceedings of the 12th International Conference of the Learning Sciences*, Singapore, Singapore, June 2016 2016.

[9] E. McBride, J. Vitale, L. Applebaum, and M. Linn. Examining the flow of ideas during critique activities in a design project. In *Proceedings of the 12th International Conference of Computer Supported Collaborative Learning*, June 2017 2017.

[10] E. McBride, J. Vitale, L. Applebaum, J. Madhok, and M. Linn. Using virtual models to improve science understanding in a hands-on solar ovens unit. In *Paper presented at the Annual Meeting of the American Educational Research Association, San Antonio, TX*, 2017.

[11] E. McBride, J. Vitale, H. Gogel, M. Martinez, Z. Pardos, and M. Linn. Predicting student learning using log data from interactive simulations on climate change. In *Learning @ Scale*, Edinburgh, UK, April 2016 2016.

[12] K. McElhaney and M. Linn. Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experimenters. *Journal of Research in Science Teaching*, 48(7):745–770, 2011.

[13] N. Podolefsky, E. Moore, and K. Perkins. Implicit scaffolding in interactive simulations: Design strategies to support multiple educational goals. *Chemistry Education Research and Practice*, 14(3):257–268, 2013.

[14] K. Ryoo and M. Linn. Can dynamic visualizations improve middle school students' understanding of energy in photosynthesis? *Journal of Research in Science Teaching*, 49(2):218–243, 2012.

[15] U. Wilensky. Netlogo. 1999.