



# The Impact of Multiple Real-Time Scaffolding Experiences on Science Inquiry Practices

Haiying Li<sup>1</sup>✉, Janice Gobert<sup>1</sup>, Rachel Dickler<sup>1</sup>, and Raha Moussavi<sup>2</sup>

<sup>1</sup> Rutgers University, New Brunswick, NJ 08901, USA  
{haiying.li, janice.gobert, rachel.dickler}@gse.rutgers.edu

<sup>2</sup> Massachusetts Institute of Technology, Cambridge, MA 02139, USA  
moussavi@mit.edu

**Abstract.** Computer-assisted assessment environments, such as intelligent tutoring systems, simulations, and virtual environments are now being designed to measure students' science inquiry practices. Some assessment environments not only evaluate students' inquiry practice competencies, but also provide real-time scaffolding in order to help students learn. The present study aims to examine the impact of real-time scaffolding from an animated, pedagogical agent on students' inquiry performance across a number of practices. Participants were randomly assigned to one of two conditions: receiving scaffolding or no scaffolding. All participants completed three virtual labs: Flower (a general pretest), Phase Change, and Density. Results showed that students who received immediate feedback during assessment performed better on subsequent inquiry tasks. These findings have implications for designers and researchers regarding the benefits of including real-time scaffolding within intelligent assessment systems.

**Keywords:** Science inquiry · Educational data mining · Real-Time scaffolding

## 1 Introduction

The Next Generation Science Standards (NGSS) were developed in order to promote high quality science instruction emphasizing the integration of three central dimensions: science inquiry practices, crosscutting concepts, and disciplinary core ideas [1]. For this vision to be realized, valid assessment of these areas is required. The first dimension, science inquiry practices, consists of eight practices that students are expected to engage in and master in grades K through 12. Some practices include: forming testable questions, carrying out experiments, analyzing/interpreting data, warranting claims with evidence [2], and communicating findings. These practices are difficult to capture and assess using traditional forms of assessment [3]. Hands-on science inquiry experiments that can elicit these practices are demanding for teachers to implement in classrooms and are extremely difficult to grade [4] due to teacher-student ratios and lack of rigor on observation-based scoring. Other traditional forms of assessment that involve multiple choice items do not fully capture student competencies at science practices [3]. Additionally, assessments based on open-response items do not fully or accurately capture students' inquiry practice

competencies [5, 6], yielding both false negatives (skilled science learners who cannot articulate what they know in words) and false positives (students who are simply parroting what they have read or heard but do not understand the content or practices, etc.). The challenges involved in effectively and accurately capturing students' science inquiry practice competencies have led to the development of various technological assessment systems, whose goals are to assess students' NGSS competencies. Specifically, researchers have developed assessments using simulations [7], virtual learning environments [8] and intelligent tutoring systems [3] to measure students' science inquiry performance. The designs of these systems vary depending on assumptions held regarding the role of assessments. Some assessment systems have been designed for the sole purpose of evaluating student performance, without providing support or guided feedback to students to promote learning within the system [i.e. 9]. Other researchers [3, 7, 8], however, note the need for and benefits of assessment systems that also promote student learning by providing different forms of scaffolds and feedback.

Scaffolding refers to hints and supports provided to a student as they engage in a task that may otherwise not be possible for that student to complete independently [10]. Students often have difficulty completing science inquiry tasks without guidance [11], which is why it is important to integrate carefully designed scaffolds into science inquiry contexts such as virtual assessments [12]. Scaffolding in online environments may involve providing hints on how to go about completing a task or providing directed feedback based on student performance on the particular task. Several technology-based inquiry assessment systems designed with scaffolded feedback have been found to benefit learning of science content and process skills [3, 7, 8, 13]. For instance, SimScientist evaluates students based on their performance in interactive simulations [7]. Students receive guidance that becomes increasingly informative based on their performance on practices related to conducting experiments and interpreting data. With this system, Quellmaz et al. [7] found that students were deeply engaged and that students, including English Language Learners, demonstrated improved science inquiry performance after completing the assessments. This system, however, covered only two topics and the impact of scaffolding on student performance for specific inquiry practices was not examined.

Another science inquiry system that used scaffolds to support student learning is Co-Lab [8]. Co-Lab virtual environments were designed for four topics related to environmental and physical sciences. The Co-Lab environment provides faded scaffolding, so the system provides gradually less support to students as they become more experienced with the system. Using Co-Lab, van Joolingen et al. [8] indicated that scaffolded support enabled students to engage with increasingly complex models and data. While the scaffolding in Co-Lab was found to be beneficial, it did not address all science inquiry practices and did not provide real-time, directed feedback to students based on their performance.

Inq-ITS, (**Inquiry Intelligent Tutoring System**; inqits.com), is a web-based intelligent tutoring and assessment system with interactive virtual simulations for NGSS science topics in the areas of life, earth, and physical science [3]. Inq-ITS has a pedagogical agent, Rex, who provides real-time scaffolding to students as they engage in

virtual labs. Real-time scaffolding means that feedback is provided immediately following student actions indicating unproductive behavior or lack of skills, as opposed to other kinds of automated feedback that may be provided before or after a student has completed an activity. The real-time scaffolding in Inq-ITS is not faded but is responsive to student performance on a number of science inquiry practices including: forming questions, planning investigations/hypothesizing, conducting experiments, interpreting data, and warranting claims with evidence [13–17]. Prior studies with this system have found the scaffolding to be particularly effective for both students' learning of practices such as hypothesis formation and conducting experiments [14, 15] as well as interpreting data and warranting claims [16, 17]. Researchers have yet to examine, however, the benefits of scaffolding in Inq-ITS across several science topics for multiple practices within the same study. It is important to investigate the influence of scaffolding across topics and practices.

The present study examined the impact of real-time scaffolding on learning of science inquiry practices within Inq-ITS. The following two research questions were addressed:

RQ1: Does students' overall performance across science inquiry practices (i.e. generating a hypothesis, collecting data, interpreting data, *and* warranting a claim) improve in subsequent science topics (i.e. phase change and density) if they receive real-time scaffolding?

RQ2: If so, for which *specific* inquiry practices (i.e. generating a hypothesis, collecting data, interpreting data, *or* warranting a claim) does real-time scaffolding improve performance?

## 2 Method

### 2.1 Materials

This study adopted three virtual labs in Inq-ITS. The Flower virtual lab contains three activities with the aim of fostering understanding about petal loss caused by salt or sugar, and about the changes to the color of a flower caused by adding red dye. The flower virtual lab could be considered the most basic Inq-ITS lab due to the minimal prior knowledge needed to successfully complete the lab and the fact that each independent variable in the simulation has only two levels. The Phase Change virtual lab contains four activities and aims to foster understanding about how the boiling point of water is impacted by a series of independent variables, including: different levels of heat (Low, Medium, and High), different amounts of ice (100 g, 200 g, and 300 g), and different sizes of a container (Small, Medium, and Large). The Density virtual lab contains three activities and aims to foster understanding about the relationship between the density of a liquid and the: type of liquid substance (water, oil, and alcohol), amount of liquid (quarter, half, and full), and shape of a container (narrow, square, and wide). Demos of Inq-ITS activities are available on the website ([inqits.com](http://inqits.com)).

## 2.2 Measures for Inquiry Practices

In each activity, participants complete four stages of inquiry practices, each of which was automatically assessed by our system. The first three inquiry stages constitute the investigative portion of the virtual lab. The last stage involves writing a scientific explanation based on the results of the investigation. The present study examines student performance on inquiry practices during the investigative portion of the virtual lab, as described below. Each practice is measured using our patented algorithms [3, 18]. The first stage of the virtual lab is the Hypothesis/planning investigation stage, where students use a widget (dropdown menu) to formulate a hypothesis based on an activity goal. This practice was measured by correct identification of an independent variable (IV) and a dependent variable (DV). In the Collect Data/conducting experiments phase, students use a widget (clickable buttons) to manipulate the independent variables in a simulation while a data table automatically records and assesses their inquiry for this practice. This practice was measured by testing a hypothesis and conducting a controlled experiment. If the variables were nominal, then data collection was also measured according to whether there was a pair of trials that tested two levels of the target nominal IV. During the Analyze Data/interpret data stage, students use a widget (dropdown menu) to state their claim, identify whether or not their claim supports their hypothesis, and select evidence that supports their claim (clickable). This stage consists two practices. One practice is about data interpretation, which is measured by correctly selecting an IV and DV for claim, interpreting the relationship between the IV and DV, and interpreting the hypothesis/claim relationship. Another practice is about warranting the claim, which is measured by the selection of more than one trial to warrant the claim, selection of controlled trials, providing data for the relationship between the IV and DV, and providing data for the hypothesis/claim relationship.

The sub-components of these practices were automatically scored as 0 or 1 point by educational data mining and knowledge engineering techniques based on whether students demonstrated competency or not [3, 18]. Prior studies have demonstrated the detectors' high performance [3, 14, 15]. Students' total score for each practice was calculated by taking the mean across corresponding sub-practice scores, and the overall inquiry score was calculated by taking the mean across all inquiry sub-practice scores.

## 2.3 Participants and Conditions

48 middle school students in grade 7 were randomly assigned into a scaffolding condition (hereafter called the Rex condition;  $N = 24$ ) or a condition without scaffolding (hereafter called the No Rex condition;  $N = 20$ ). Students were from three different middle schools located in the North Western United States. Two of the middle schools were public (42.3% and 60.0% of students received free and reduced lunch, respectively) and one was an alternative middle school (82.4% of students received free and reduced lunch). All participants completed three Inq-ITS virtual labs during their regular science class time over the course of one month in the order of: Flower, Phase Change, and Density. Students received regular science instruction between the implementation of the labs. The Flower virtual lab was used as a baseline, in which all the participants completed

**Table 1.** Statistics for time  $\times$  condition across three virtual labs.

Time	Condition	Mean	SD	95% C.I.		F	$\eta^2$	Power
				Lower	Upper			
1 (Flower)	No Rex	0.53	0.20	0.45	0.61	0.05	0.001	0.056
	Rex	0.52	0.26	0.45	0.59			
2 (Phase change)	No Rex	0.53	0.28	0.44	0.62	24.41***	0.368	0.998
	Rex	0.83	0.19	0.75	0.91			
3 (Density)	No Rex	0.82	0.26	0.73	0.92	0.09	0.002	0.060
	Rex	0.84	0.21	0.75	0.93			

Note. \*\*\*  $p < .001$ . SD = standard deviation. C.I. = confidence interval.  $N = 44$ ;  $df = 1, 42$ .

three activities without any real-time scaffolding from the animated, pedagogical agent, Rex.

Results of a one-way ANOVA for the Flower virtual lab showed that the total inquiry scores were not significantly different between the two conditions (see Flower in Table 1 for details). Results of a one-way MANOVA (four inquiry practices  $\times$  2 conditions) for Flower further showed that performance on each inquiry sub-practice in Flower was not significantly different between conditions (see Table 2 for details). These results indicated that students in the two conditions (Rex and No Rex) were not significantly different on their competencies at inquiry before real-time scaffolding was provided in the first virtual lab. Thus, to investigate the impact of scaffolding, students in the Rex condition received scaffolding from Rex in the second (Phase Change) and third (Density) virtual labs only when they did not demonstrate the presence of competency. In the No Rex condition, students never received scaffolding and could progress between the stages of an activity even if they demonstrated poor performance on inquiry sub-practices.

Scaffolding in the Rex condition was provided in real time when the system detected that the student needed support on any of the science inquiry sub-practices. If competency on a particular sub-practice was not demonstrated (for example when the student collects data), Rex would pop-up on the screen with a speech bubble providing a general, orienting hint (see Fig. 1) first. For instance, if a student was running multiple trials in an experiment with the wrong independent variable, Rex would remind the student to look at their hypothesis and make sure they were designing an experiment that tested the hypothesis. The student would click an “okay” button when he/she was finished reading the Rex hint and would then continue with the activity. Some scaffolded Rex hints allow students to request additional information, such as the definition of the term “independent variable” [15, 17].

**Table 2.** Statistics for practices × time × condition across three virtual labs.

Practices	Time	Condition	Mean	SD	95% C.I.		F	$\eta^2$	Power
					Lower	Upper			
Generating hypothesis	1	No Rex	0.94	0.16	0.87	1.01	1.88	0.043	0.268
		Rex	0.88	0.16	0.81	0.94			
	2	No Rex	0.83	0.26	0.74	0.91	4.18*	0.090	0.515
		Rex	0.95	0.12	0.87	1.03			
	3	No Rex	0.90	0.17	0.83	0.97	0.26	0.006	0.079
		Rex	0.92	0.13	0.86	0.99			
Collecting data	1	No Rex	0.40	0.30	0.24	0.55	0.11	0.003	0.062
		Rex	0.36	0.38	0.22	0.50			
	2	No Rex	0.50	0.37	0.37	0.63	21.65***	0.340	0.995
		Rex	0.90	0.19	0.79	1.02			
	3	No Rex	0.85	0.22	0.73	0.96	0.17	0.004	0.069
		Rex	0.88	0.27	0.77	0.98			
Interpreting data	1	No Rex	0.62	0.18	0.53	0.70	0.12	0.003	0.063
		Rex	0.60	0.19	0.52	0.67			
	2	No Rex	0.50	0.21	0.41	0.59	16.01***	0.276	0.974
		Rex	0.75	0.21	0.67	0.84			
	3	No Rex	0.79	0.30	0.68	0.91	0.00	0.000	0.050
		Rex	0.79	0.21	0.69	0.90			
Warranting claims	1	No Rex	0.17	0.17	0.06	0.29	0.90	0.021	0.152
		Rex	0.25	0.31	0.14	0.35			
	2	No Rex	0.31	0.31	0.19	0.43	25.05***	0.374	0.998
		Rex	0.71	0.22	0.60	0.82			
	3	No Rex	0.75	0.34	0.62	0.88	0.07	0.002	0.058
		Rex	0.77	0.23	0.65	0.89			

Note. \*\*\*  $p < .001$ . SD = standard deviation. C.I. = confidence interval.  $N = 44$ ;  $df = 1, 42$ .

If a student demonstrated competency on all sub-practices after receiving scaffolding from Rex, Rex would not appear again. If the students' performance on the sub-practice did not improve with subsequent attempts, then Rex would continue to pop-up and provide hints that gradually became more informative. If other sub-practices were not demonstrated, Rex would provide feedback on them. In other words, Rex would not let students progress from one inquiry stage to the next until they had successfully demonstrated all of the inquiry sub-practices related to the particular stage they were on.



Fig. 1. Example of a Rex pop-up hint.

### 3 Analyses, Findings, and Discussion

A repeated measures analysis was performed to investigate whether students' performance on each of the inquiry practices improved with real-time scaffolding provided after completion of the first, baseline virtual lab (i.e. Flower). The two within-subjects factors were the four inquiry practices and time phase of completion (i.e. first, second, or third virtual lab completed). The analyses adopted students' performance on their first attempts before scaffolding was provided by Rex for each inquiry practice, because this performance reflected students' real competency in inquiry practices. The between-subjects factor was the two experimental conditions (Rex versus No Rex). The analyses adopted the mean scores of each inquiry practice across all the activities in each virtual lab.

#### 3.1 Performance on Overall Inquiry Practices Across Time Phases

Results of the repeated measures analysis showed a significant two-way interaction between time and condition,  $F(2, 41) = 16.36, p < .001, \eta^2 = .444$ . Table 1 illustrates the means, standard deviations, and other descriptive statistics. The pairwise comparisons of conditions at each time phase showed that students achieved higher inquiry scores (an aggregated score) in the Rex condition than the No Rex condition in the second virtual lab. There were no significant differences between the two conditions on the third virtual lab. To further explore why there was no difference in performance on the aggregated inquiry score between the two conditions on the third virtual lab, pairwise comparisons of time within each condition were conducted. Results showed that students in the Rex condition achieved significantly higher performance in the second virtual lab relative to the first virtual lab,  $p < .001$ , Cohen's  $d = 1.36$ . This increase was not significant from the second to the third virtual lab, but a significant increase was found from the first to the third virtual lab,  $p < .001, d = 1.35$ . In the No Rex condition, no significant increase in performance was found from the first to the second virtual lab, but in the

third virtual lab students significantly improved relative to the first ( $p < .001$ ,  $d = 1.41$ ) and second virtual lab ( $p < .001$ ,  $d = 1.17$ ).

These findings indicate that students who received scaffolding from Rex significantly improved their performance on inquiry practices in the second virtual lab versus students who did not receive scaffolding from Rex. However, in the third virtual lab, students who never received Rex's scaffolding caught up with those students who received Rex's scaffolding. These findings imply that whether or not students receive scaffolding, their inquiry performance improved with increased use of Inq-ITS virtual labs; three virtual lab activities (each with 3-4 driving questions), however, are required to yield this change. Additionally, the improvement for students who received scaffolding was much faster than those who did not receive scaffolding.

Therefore, the answer to the first research question is that students' performance on overall inquiry practices greatly improved in the subsequent virtual lab if they received real-time scaffolds. Based on our research design, we believe this improvement is due to Rex's scaffolds that provided students with guidance when they needed it in order to support them in conducting inquiry. The series of scaffolds served to elaborate the reasons why a particular practice was important and the steps involved in successfully engaging in the practice. This form of guided discovery facilitates learning and performance on future inquiry tasks [13–17].

### 3.2 Performance on Each Inquiry Practice Across Time Phases

As a follow-up to the analyses above, we examined students' performance on each specific inquiry practice of interest (i.e. generating a hypothesis, collecting data, interpreting data, warranting a claim). Results of the repeated measures analysis showed a significant three-way interaction of practices  $\times$  time  $\times$  condition,  $F(6, 37) = 2.53$ ,  $p = .038$ ,  $\eta^2 = .291$ . Table 2 illustrates the means and standard deviations of inquiry practices and other statistics. The pairwise comparisons for each inquiry practice showed students achieved higher inquiry practice scores in the Rex condition than the No Rex condition in the second virtual lab. There were no significant differences between the two conditions in the third virtual lab at the specific inquiry practice level; similar to results at the overall inquiry performance level.

The pairwise comparisons showed that students in the Rex condition achieved significantly higher performance in the second virtual lab than in the first virtual lab for practices of data collection,  $p < .001$ ,  $d = 1.80$ ; interpreting data,  $p = .004$ ,  $d = 0.75$ ; and warranting claims,  $p < .001$ ,  $d = 1.71$ . This increase was not significant from the second to the third virtual lab, but was significant from the first to the third virtual lab for all three practices, i.e., data collection,  $p < .001$ ,  $d = 1.58$ ; interpreting data  $p = .002$ ,  $d = 0.95$ ; and warranting claims,  $p < .001$ ,  $d = 1.91$ . In the No Rex condition, no significant increase was found from the first to the second virtual lab. However, a significant increase was found from the first virtual lab to third virtual lab for data collection,  $p < .001$ ,  $d = 1.71$ ; data interpretation,  $p = .012$ ,  $d = 0.69$ ; and warranting claims,  $p < .001$ ,  $d = 2.16$ .

These findings are similar to those for overall inquiry practices, except for the practice of generating a hypothesis, on which students scored very high in the first virtual



lab, which suggests that they had already mastered this practice. The answer to the second research question is that real-time scaffolding greatly improved students' performance on data collection, data interpretation, and warranting claims in the second virtual lab. However, students' performance was very similar between the Rex and No Rex conditions in the third virtual lab (i.e. average scores of 0.85–0.88 points for data collection, 0.79 points for interpretation, and 0.75–0.77 points for warranting; see Table 2). This consistent pattern informs us that students' performance on these three inquiry practices does improve with increased use of Inq-ITS virtual labs, but improves faster when real-time scaffolding is provided.

#### **4 Conclusions, Future Directions, and Implications**

In this study we investigated whether real-time scaffolds within an inquiry system improved students' inquiry performance; we also investigated the effects of scaffolds on student performance across multiple activities. We found that students who received scaffolding from Rex significantly improved their performance in the second virtual lab relative to those who did not receive scaffolding from Rex on both overall inquiry and on specific inquiry practices. In the third virtual lab, students who never received Rex's scaffolding eventually reached similar levels of performance relative to those who received Rex's scaffolding. These findings imply that whether or not students received scaffolding, their performance eventually improved. Whether this increase in performance, however, was a demonstration of the benefits of discovery learning or the effects of teacher instruction remains unclear. Future studies are needed to further examine the potential impact of in-class instruction that occurs between student's use of virtual labs.

Additionally, students with scaffolding improved their overall inquiry performance as well as performance on the specific practices of collecting data, interpreting data, and warranting a claim faster than those who did not receive real-time scaffolding. Prior studies have explored the benefits of scaffolding in Inq-ITS through modes such as interviews with students [19]. In the future, it would be valuable to attend to whether the number of Rex hints decreased for students from the first to third virtual lab.

Our study provides empirical evidence that a well-designed computer-assisted science inquiry system alone facilitates learning, but adding scaffolded feedback further accelerates learning of inquiry practices. These findings thus inform assessment designers and researchers that, if technology allows, adding real-time scaffolding can greatly benefit student learning of and performance on inquiry practices. This study contributes to research on the design of assessment systems in the following three ways. First, this study provides empirical evidence that assessment with automated, real-time scaffolding can effectively and efficiently improve students' learning. Second, this study further demonstrates how a science inquiry environment can foster student learning of practices even when scaffolding is not present. While students can learn within carefully designed environments without scaffolding, the rate at which they learn is slower relative to when scaffolding is provided. Lastly, the findings of this study inform designers and researchers of the benefits of adding real-time scaffolding to assessment systems in terms of the rate of student learning.

**Acknowledgements.** This research was supported by the Institute of Education Sciences (R305A120778) and National Science Foundation (1252477, 1629045) to Janice Gobert, Principal Investigator at Rutgers University.

## References

1. Next Generation Science Standards Lead States: Next Generation Science Standards: for States, by States. National Academies Press, Washington (2013)
2. National Research Council: A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. National Academies Press, Washington (2012)
3. Gobert, J.D., Sao Pedro, M., Raziuddin, J., Baker, R.S.: From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *J. Learn. Sci.* **22**, 521–563 (2013)
4. Deters, K.M.: Student opinions regarding inquiry-based labs. *J. Chem. Educ.* **82**(8), 1178–1180 (2005)
5. Li, H., Gobert, J., Dicker, R.: Dusting off the messy middle: assessing students' inquiry skills through doing and writing. In: André, E., Baker, R., Hu, X., Rodrigo, M.M.T., du Boulay, B. (eds.) AIED 2017. LNCS (LNAI), vol. 10331, pp. 175–187. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-61425-0\\_15](https://doi.org/10.1007/978-3-319-61425-0_15)
6. Gobert, J.: US News & World Report, 13 May 2016. <http://www.usnews.com/news/articles/2016-05-13/op-ed-educational-data-mining-can-enhance-science-education>
7. Quellmalz, E.S., Timms, M.J., Silberglitt, M.D., Buckley, B.C.: Science assessments for all: integrating science simulations into balanced state science assessment systems. *J. Res. Sci. Teach.* **49**(3), 363–393 (2012)
8. van Joolingen, W.R., de Jong, T., Lazonder, A.W., Savelsbergh, E.R., Manlove, S.: Co-lab: research and development of an online learning environment for collaborative scientific discovery learning. *Comput. Hum. Behav.* **21**(4), 671–688 (2005)
9. Zapata-Rivera, D., Jackson, T., Liu, L., Bertling, M., Vezzu, M., Katz, Irvin R.: Assessing science inquiry skills using trialogues. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) ITS 2014. LNCS, vol. 8474, pp. 625–626. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-07221-0\\_84](https://doi.org/10.1007/978-3-319-07221-0_84)
10. Vygotsky, L.S.: *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge (1978)
11. Hmelo-Silver, C.E., Duncan, R.G., Chinn, C.A.: Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). *Educ. Psychol.* **42**(2), 99–107 (2007)
12. Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R.G., Kyza, E., Edelson, D., Soloway, E.: A scaffolding design framework for software to support science inquiry. *J. Learn. Sci.* **13**(3), 337–386 (2004)
13. Gobert, J., Moussavi, R., Li, H., Sao Pedro, M., Dickler, R.: Scaffolding students' on-line data interpretation during inquiry with Inq-ITS. In: *Cyber-Physical Laboratories in Engineering and Science Education*. Springer (in press)
14. Sao Pedro, M., Baker, R., Gobert, J.: Incorporating scaffolding and tutor context into Bayesian knowledge tracing to predict inquiry skill acquisition. In: *Proceedings of the 6<sup>th</sup> International Conference on Educational Data Mining*, pp. 185–192. EDM Society (2013)
15. Sao Pedro, M.: *Real-time Assessment, Prediction, and Scaffolding of Middle School Students' Data Collection Skills Within Physical Science Simulations*. Worcester Polytechnic Institute, Worcester (2013)

16. Moussavi, R., Gobert, J., Sao Pedro, M.: The effect of scaffolding on the immediate transfer of students' data interpretation skills within science topics. In: Proceedings of the 12th International Conference of the Learning Sciences, pp. 1002–1005. Scopus, Ipswich (2016)
17. Moussavi, R.: Design, Development, and Evaluation of Scaffolds for Data Interpretation Practices During Inquiry. Worcester Polytechnic Institute, Worcester (2018)
18. Li, H., Gobert, J., Dickler, R.: Automated assessment for scientific explanations in on-line science inquiry. In: Hu, X., Barnes, T., HersHKovitz, A., Paquette, L. (eds.) Proceedings of the 10th International Conference on Educational Data Mining, pp. 214–219. EDM Society, Wuhan (2017)
19. Gobert, J.D., Sao Pedro, M.A.: Digital assessment environments for scientific inquiry practices. In: Rupp, A.A., Leighton, J.P (eds.) The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications, West Sussex, UK, pp. 508–534 (2017)