



Evaluating the Transfer of Scaffolded Inquiry: What Sticks and Does It Last?

Haiying Li^(✉), Janice Gobert, and Rachel Dickler

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

Abstract. The Next Generation Science Standards [1] expect students to master disciplinary core ideas, crosscutting concepts, and scientific practice. In prior work, we showed that students benefited from real time scaffolding of science practices such that students' inquiry competencies both improved over time and transferred to new science topics. The present study examines the robustness of adaptive scaffolding by evaluating students' inquiry performances at a very fine-grained level in order to investigate *what* aspects of inquiry are robust over time once scaffolding was removed. 108 middle school students in grade 6 used Inq-ITS and received adaptive scaffolding for three lab activities in the first inquiry topic they completed (i.e. Animal Cell); they then completed 10 activities without scaffolding across three new topics. Results showed that after removing scaffolding, student's inquiry performance generally improved with slight variations in performance across driving questions and over time. Overall, these findings suggest that adaptive scaffolding may support students' inquiry learning and transfer of inquiry practices over time and across topics.

Keywords: Science inquiry · Growth in inquiry performance · Scaffolding

1 Introduction

In science inquiry contexts, students require support in order to effectively engage in inquiry investigations [2–4]. Supports provided to students can be in the form of scaffolds designed to help students reach a level of performance that would not be possible if they were to do a task independently [5, 6]. The types of scaffolds students receive within online science environments may vary from fixed [7] to faded [8] to adaptive scaffolds [9]. Fixed scaffolds are supports that are provided to all students consistently, regardless of student performance [7, 10]. Faded scaffolds, on the other hand, are supports that are gradually removed with increasing use of a particular system [8, 10, 11]. Another form of scaffolds are adaptive scaffolds, which are supports that are provided to students in real-time based on students' performance in a system [12, 13]. While fixed [7], faded [8], and adaptive scaffolds [12] have benefited student learning in science environments to some extent, adaptive scaffolds show the greatest promise in terms of promoting transfer of inquiry practices [13, 14] because they provide students with the information they need when they need it most [15].

In the context of science inquiry, transfer of inquiry practices may be assessed in terms of near transfer (i.e. transfer to similar inquiry tasks presented briefly after the initial inquiry task; [16]) or far transfer (i.e. transfer to inquiry tasks in different contexts and after extended periods of time; [16]). Studies have demonstrated how engagement in computer-supported learning environments can promote transfer of science content understandings [17, 18] and practices such as scientific reasoning [16]. In the intelligent tutoring system, Inq-ITS [9], researchers have demonstrated transfer of multiple scientific practices across topics and over time [14] including: hypothesizing [12, 19], collecting data [20, 21], and interpreting data/warranting claims with evidence [22, 23]. Each of these practices can be operationalized into different finer-grained sub-practices. Studies have yet to investigate the transfer of inquiry at the sub-practice level over time and across topics. The present study examines whether adaptive scaffolding of inquiry practices in the first three Inq-ITS activities (i.e. driving questions) leads to transfer of inquiry practices across topics at varying time intervals at the sub-practice level.

2 Method

2.1 Participants and Materials

The participants in the present study were 108 6th grade students from a middle school in the northeastern United States who completed the following Inq-ITS [9] lab activities: *Animal Cell* (three driving questions: (1) how can you increase the transfer or protein in an animal cell?, (2) how you can decrease the production of ribosomes?, and (3) how you can reduce the production of protein?), *Plant Cell* (three driving questions: (1) how can you increase the transfer or protein in a plant cell?, (2) how you can decrease the production of ribosomes?, and (3) how you can reduce the production of protein?), *Genetics* (three driving question activities: how does changing a mother monster's (1) F, (2) L, and (3) H alleles impact the traits of the babies?), and *Natural Selection* (four driving questions: what is the optimal foliage for (1) the green, long furred and (2) the red, short furred monsters?, what is the optimal temperature for (3) the green, short furred and (4) the red, long furred monsters?).

Each of these Inq-ITS activities contained four stages where students first formed a question/hypothesis, carried out an investigation/collected data, analyzed and interpreted data, and finally communicated their findings [9, 10]. Currently, adaptive, real-time scaffolding is available within the first three stages of the microworlds [19–23] (scaffolding is being developed for communicating findings [24]) based on automated scoring in Inq-ITS ([25]; see Measures section). The only difference between adaptive scaffolded and unscaffolded Inq-ITS activities is the presence of the pedagogical agent, Rex. For example, in the scaffolded animal cell activities in the present study, if a student was evaluated as having difficulty on a particular practice, then Rex would pop up on the student's screen with different types of information depending on the student's specific difficulty [26, 27]. Rex would first provide students with an orienting hint reminding the students of the inquiry practice/sub-practice that they were engaging in [28]. If the students continued to have difficulty with the practice, Rex would provide

a procedural hint (explaining the steps involved in the practice/sub-practice) followed by a conceptual hint (explaining the inquiry practice/sub-practice) and finally an instrumental hint (explaining the exact steps).

2.2 Measures

In the present study, the dependent variables were four inquiry practices. Each inquiry practice in Inq-ITS is operationalized at a fine-grained level (i.e., broken down into different sub-practices/sub-components). The *hypothesizing practice* was measured by: identifying an independent variable (IV) and dependent variable (DV). The *collecting data practice* was measured by: testing the hypothesis and running targeted and controlled trials. The *interpreting data practice* was measured by: correctly selecting the IV and DV for a claim, correctly interpreting the relationship between the IV and DV, and correctly interpreting the hypothesis/claim relationship. The *warranting claims practice* was measured by: warranting the claim with more than one trial, warranting with controlled trials, correctly warranting the relationship between the IV and DV, and correctly warranting the hypothesis/claim relationship. Each inquiry sub-practice was automatically scored as 0 points if incorrect or 1 point if correct using the knowledge engineering and educational data mining techniques in Inq-ITS, validated in prior studies [9].

This study had a time variable with four levels: Time 1 (i.e., Animal Cell in month 0), Time 2 (i.e., Plant Cell in month 1.3), Time 3 (i.e., Genetics in month 2.7), and Time 4 (i.e., Natural Selection in month 5.7). Moreover, this study had a variable of the number of driving questions that students completed over time: driving questions 1 to 3 in month 0 (i.e., Animal Cell), 4 to 6 in month 1.3 (i.e. Plant Cell), 7 to 9 in month 2.7 (i.e., Genetics), and 10 to 13 in month 5.7 (i.e., Natural Selection).

3 Results and Discussion

We used linear mixed models (LMMs) to investigate whether there was evidence of transfer by evaluating students' inquiry competencies across driving questions over time after removing the adaptive scaffolding. We performed four sets of LMM analyses where we focused on the pattern within each inquiry practice.

3.1 Model Selection

For the analysis of the data, we followed the “top-down” modeling strategy and selected the models that best fit the data. We ran an unconditional model with intercepts only, and then added each variable independently as well as in combination. Each type of added variable(s) generated three models based on the variation of random effects: subjects only (Intercept), the number of driving questions and/or time variable(s) only (Slope), or both subjects and the number of driving questions and/or time variable(s). We compared the models using the -2 Restricted Log Likelihood ($-2RLL$) [29] and selected the full models in this study due to their best fit for a greater number of practices (namely, hypothesis, data collection, and warranting claims).

3.2 Performance Across Driving Questions and Over Time

We then examined inquiry scores across driving questions and time for each practice. Results showed that the fixed effects for the hypothesizing practice were significant, $F(1, 108.25) = 24.39, p < .001$ for driving questions and $F(1, 107.25) = 11.32, p = .001$ for time. Fixed-effects parameters were significant for hypothesizing ($\beta = 0.03, p < .001$ for driving question; $\beta = -0.04, p = .001$ for time), collecting data ($\beta = 0.05, p < .001$ for driving question; $\beta = -0.06, p < .001$ for time), and warranting claims practices ($\beta = 0.04, p < .001$ for driving question; $\beta = -0.05, p < .001$ for time). These results indicate that students improved their performance on these three inquiry practices with the increasing use of Inq-ITS, but that the long-time intervals between usage resulted in a slight decrease in performance. This pattern was not found for the practice of interpreting data, potentially due to students starting with relatively high performance (Mean = 0.79) or interactions with topic complexity [30].

The random effects showed a significant intercept for the hypothesizing ($\beta = 0.03, Z = 3.17, p < .01$), collecting data ($\beta = 0.05, Z = 3.87, p < .001$), and warranting claims practice ($\beta = 0.06, Z = 4.18, p < .001$). Results also showed a significant driving question random effect for hypothesizing ($\beta = 0.001, Z = 1.97, p < .05$) and collecting data ($\beta = 0.002, Z = 2.00, p < .05$). Additionally, in hypothesizing, we found a significant driving question and time random effect ($\beta = -0.002, Z = -2.03, p < .05$) and significant time effect ($\beta = 0.004, Z = 2.26, p < .05$). We also found a significant covariance between the intercept and the driving question coefficient for collecting data ($\beta = -0.01, Z = -2.01, p < .05$). The findings of these random effects confirmed a fair amount of student-to-student variation in the starting performance for practices of hypothesizing, collecting data, and warranting claims, but varied patterns for driving question, time, and both driving question and time effects. This demonstrates that transfer of learning was different for different inquiry practices for students.

4 Conclusions, Future Directions, and Implications

In this study we investigated the robustness of our scaffolding using students' performances on various inquiry practices across driving questions at different time intervals, thereby addressing near (across driving questions at each time) and far transfer (over time). Our results showed, in general, that our scaffolding was robust for practices of hypothesizing, collecting data, and warranting claims. A limitation of the present study is that there was no control condition, which makes it challenging to distinguish between effects of external factors such as teacher instruction between usage of the system. In the future it will be valuable to examine differences between students in a scaffolded and unscaffolded condition to more fully understand the influence of the adaptive scaffolds in Inq-ITS on students' inquiry performance.

Overall, the findings in the present study inform assessment designers and researchers that, if properly designed, scaffolding aimed at supporting students' competencies at various inquiry practices can greatly benefit students' deep learning of, transfer of, and performance on inquiry practices over time.

References

1. Next Generation Science Standards Lead States: Next Generation Science Standards: For States, by States. National Academies Press, Washington (2013)
2. 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**, 99–107 (2007)
3. Kang, H., Thompson, J., Windschitl, M.: Creating opportunities for students to show what they know: the role of scaffolding in assessment tasks. *Sci. Educ.* **98**, 674–704 (2014)
4. McNeill, K.L., Krajcik, J.S.: Supporting grade 5–8 students in constructing explanations in science: the claim, evidence, and reasoning framework for talk and writing. Pearson (2011)
5. Vygotsky, L.S.: *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge (1978)
6. Quintana, C., et al.: A scaffolding design framework for software to support science inquiry. *J. Learn. Sci.* **13**, 337–386 (2004)
7. Tabak, I., Reiser, B.J.: Software-realized inquiry support for cultivating a disciplinary stance. *Pragmat. Cogn.* **16**, 307–355 (2008)
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**, 671–688 (2005)
9. 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)
10. McNeill, K.L., Lizotte, D.J., Krajcik, J., Marx, R.W.: Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *J. Learn. Sci.* **15**, 153–191 (2006)
11. Martin, N.D., Tissenbaum, C.D., Gnesdilow, D., Puntambekar, S.: Fading distributed scaffolds: the importance of complementarity between teacher and material scaffolds. *Instr. Sci.* **47**, 1–30 (2018)
12. Gobert, J.D., Moussavi, R., Li, H., Sao Pedro, M., Dickler, R.: Real-time scaffolding of students' online data interpretation during inquiry with Inq-ITS using educational data mining. In: Auer, M.E., Azad, A.K.M., Edwards, A., de Jong, T. (eds.) *Cyber-Physical Laboratories in Engineering and Science Education*, pp. 191–217. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-76935-6_8
13. Noroozi, O., Kirschner, P.A., Biemans, H.J., Mulder, M.: Promoting argumentation competence: extending from first- to second-order scaffolding through adaptive fading. *Educ. Psychol. Rev.* **30**, 1–24 (2017)
14. Li, H., Gobert, J., Dickler, R.: Testing the robustness of inquiry practices once scaffolding is removed. Submitted to: *Intelligent Tutoring Systems* (submitted)
15. Koedinger, K.R., Anderson, J.R.: Illustrating principled design: the early evolution of a cognitive tutor for algebra symbolization. *Interact. Learn. Environ.* **5**, 161–180 (1998)
16. Chen, Z., Klahr, D.: Remote transfer of scientific-reasoning and problem-solving strategies in children. In: *Advances in Child Development and Behavior*, pp. 419–470. JAI (2008)
17. Borek, A., McLaren, B.M., Karabinos, M., Yaron, D.: How much assistance is helpful to students in discovery learning? In: Cress, U., Dimitrova, V., Specht, M. (eds.) *EC-TEL 2009*. LNCS, vol. 5794, pp. 391–404. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04636-0_38
18. Tao, P.K., Gunstone, R.F.: The process of conceptual change in force and motion during computer-supported physics instruction. *J. Res. Sci. Teach.* **36**, 859–882 (1999)

19. Gobert, J.D., Sao Pedro, M.A., Baker, R.S., Toto, E., Montalvo, O.: Leveraging educational data mining for real-time performance assessment of scientific inquiry skills within microworlds. *J. Educ. Data Min.* **4**, 111–143 (2012)
20. 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)
21. 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 6th International Conference on Educational Data Mining, pp. 185–192. EDM Society (2013)
22. Moussavi, R.: Design, Development, and Evaluation of Scaffolds for Data Interpretation Practices During Inquiry. Worcester Polytechnic Institute, Worcester (2018)
23. 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 International Conference of the Learning Sciences, pp. 1002–1005. Scopus, Ipswich (2016)
24. 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 Conference on Educational Data Mining, pp. 214–219. EDM Society, Wuhan (2017)
25. Gobert, J.D., Baker, R.S., Sao Pedro, M.A.: Inquiry skills tutoring system. U.S. Patent No. 9,373,082. U.S. Patent and Trademark Office, Washington, DC (2016)
26. Anderson, J.R., Corbett, A.T., Koedinger, K.R., Pelletier, R.: Cognitive tutors: lessons learned. *J. Learn. Sci.* **4**(2), 167–207 (1995)
27. Koedinger, K.R., Corbett, A.: Cognitive tutors: technology bringing learning sciences to the classroom. In: *The Cambridge Handbook of the Learning Sciences*, pp. 61–77. Cambridge University Press, New York (2006)
28. de Jong, T.: Computer simulations: technological advances in inquiry learning. *Science* **312**, 532–533 (2006)
29. West, B.T., Welch, K.B., Galecki, A.T.: Linear mixed model. Chapman Hall/CRC, Boca Raton (2007)
30. Li, H., Gobert, J., Dickler, R., Moussavi, R.: The impact of multiple real-time scaffolding experiences on science inquiry practices. In: Nkambou, R., Azevedo, R., Vassileva, J. (eds.) ITS 2018. LNCS, vol. 10858, pp. 99–109. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91464-0_10