International Society of the Learning Sciences

Supporting Students Remotely: Integrating Mathematics and Science in Virtual Labs

Rachel Dickler, Rutgers University, rachel.dickler@gse.rutgers.edu
Michael Sao Pedro, Apprendis, mikesp@apprendis.com
Amy Adair, Rutgers University, amy.adair@gse.rutgers.edu
Janice Gobert, Rutgers University & Apprendis, janice.gobert@gse.rutgers.edu
Joseph Olsen, Rutgers University, joseph.olsen@rutgers.edu
Jason Kleban, Apprendis, jkleban@apprendis.com
Cameron Betts, Apprendis, cam@apprendis.com
Charity Staudenraus, Apprendis, charity@apprendis.com
Patrick Roughan, Apprendis, paroughan@gmail.com

Abstract: Tools that automatically assess and support students are important during remote instruction due to COVID-19 because students do not have direct access to teacher support. We present results of the remote use of virtual labs in an Inquiry Intelligent Tutoring System (Inq-ITS), which captures student performance on practices at the intersection of mathematics and science. Implications are discussed for the development of scaffolding and design of labs to support remote instruction.

Introduction

During the COVID-19 pandemic, teachers faced many challenges with supporting their students' learning in *STEM* classrooms in particular (Reimers & Schleicher, 2020). Specifically, providing support to students remotely requires technologies that can assess students' STEM practice competencies (NGSS, 2013) and difficulties in real time. There are some technologies with this goal for science (i.e., WISE, Vitale et al., 2019) and for mathematics (PAT, Koedinger & Anderson, 1998), but few support the full range of STEM practices. The Inquiry Intelligent Tutoring System (Inq-ITS) is a technology that assesses and supports students on inquiry practices in real time as they complete authentic virtual lab investigations (Gobert et al., 2013). Inq-ITS is currently being expanded to assess and support the practices involved in using math in science (Sao Pedro & Betts, 2019). In this study, we examined data from students who completed Inq-ITS labs with mathematics during remote instruction to address the following research questions: (1) Are math practices more challenging for students relative to other inquiry practices in Inq-ITS?, (2) What influence does the type of math relationship (e.g., inverse square) have on the difficulty of a math practice?, (3) Which specific math sub-practices are most difficult for students?

Methods

Participants in the present study included 4 teachers of eighth grade science courses and their students (N = 74 total students) from middle schools across the United States. Students completed at least one tutorial Inq-ITS lab prior to completing the lab that involved using mathematics in science (i.e., the Gravity and Mass Lab) during remote learning between May to June of 2020. In the Gravity and Mass lab, students use a simulation to investigate how certain variables relate to the force of gravity on a pile of gold on a spaceship. Specifically, students complete three investigations to identify the mathematical relationships between: the amount of gold and force of gravity on the gold (linear), the gold's distance from a planet's center and force of gravity on the gold (inverse square), and the planet's mass and the force of gravity on the gold (linear). Informed by the NGSS (2013) practices, each investigation in the lab consists of six stages: 1) **asking questions**/hypothesizing about the relationship between variables based on a goal; 2) **carrying out an investigation**/collecting data using a simulation; 3) **setting up graphs** by selecting the variables to place on each axis and data to plot; 4) **constructing graphs and equations** by determining the type of mathematical relationship between variables and creating a best fit line/curve; 5) **analyzing and interpreting data** to draw a final conclusion; and 6) **explaining findings** in writing.

To identify the practices that were most challenging for students, we used the validated automated scoring in Inq-ITS (Gobert et al., 2013; Sao Pedro & Betts, 2019) within stages 1–4 (automated scoring for stages 5–6 is in development). Each practice (in bold) is scored according to student performance on sub-practices involved in completing the practices: **asking questions** (identifying the target independent and dependent variable), **carrying out investigations** (running sufficient, controlled trials that target the variables of interest), **setting up graphs** (labeling axes and selecting sufficient data to plot from controlled, targeted trials), and **constructing graphs and equations** (identifying the mathematical relationship in the graph, adjusting the equation to create a best fit curve). The final score for each practice is the average of the sub-practices, scored as binary. We analyzed student performance on practices and sub-practices across the entire lab.

Results

First a Mixed Model ANOVA was conducted to investigate whether there were differences in student performance between teachers and across practices. Results showed that the overall model was significant, $F(3, 68) = 13.46$, $p < .001$, $n^2 = .37$. The between-subjects effect ($F(3, 70) = 1.50$, $p = .223$, $n^2 = .06$) and interaction between teacher and practice was not significant ($F(3, 70) = 1.27$, $p = .290$, $n^2 = .05$), which indicated that there were no significant differences in student performance between teachers. There was, however, a significant within-subjects effect of inquiry practices, $F(1, 70) = 21.44$, $p < .001$, $n^2 = .23$, which means that there were significant differences in student performance on practices. Follow-up analyses using within-subjects t-tests revealed that students performed significantly higher on carrying out investigations ($M = .97$, $SD = .09$) than on asking questions ($M = .92$, $SD = .16$, $t(73) = -2.99$, $p = .004$), students performed well on both asking questions ($M = .92$, $SD = .16$) and setting up graphs ($M = .88$, $SD = .18$, $t(73) = -1.62$, $p = .109$), and students performed significantly higher on setting up graphs ($M = .88$, $SD = .18$) than on constructing graphs/equations ($M = .77$, $SD = .27$, $t(73) = 5.08$, $p < .001$). This suggests that the practice of constructing graphs/equations in science was most difficult for students.

Next, we used a Repeated Measures ANOVA to determine whether there were any differences in student performance on the practice of constructing graphs/equations across the three investigations in the Gravity & Mass lab in relation to the type of mathematical relationship. We found that there was a significant main effect of the investigation, $F(2, 146) = 21.71$, $p < .001$, $n^2 = .23$. Follow-up analyses using within-subjects t-tests revealed that students performed significantly better on the first investigation (linear; $M = .84$, $SD = .33$) relative to the second investigation (inverse square; $M = .61$, $SD = .41$, $t(73) = 5.04$, $p < .001$), and third investigation (linear; $M = .88$, $SD = .28$) relative to the second investigation (inverse square; $M = .61$, $SD = .41$, $t(73) = -5.68$, $p < .001$). There were no significant differences between the first and third investigations (both linear; $t(73) = 1.06$, $p = .292$), but students did perform slightly higher on the third investigation (meaning that students did maintain performance with some improvement from the first to third investigation on this practice). These results support prior findings indicating that students experience challenges with non-linear relationships (De Bock et al., 2017).

Finally, we examined whether there were any significant differences in student performance between the sub-practices involved in constructing graphs/equations using a within-subjects t-test. We found that students performed significantly better on the sub-practice of adjusting an equation to create a best fit line/curve ($M = .83$, $SD = .28$) relative to the sub-practice of identifying the type of mathematical relationship ($M = .73$, $SD = .32$, $t(73) = -3.62$, $p < .001$, $d = .33$). This finding aligns with earlier research findings that students have difficulty determining the type of functional relationships between data (i.e., linear, inverse, etc.; De Bock et al., 2017).

Conclusions and future work

As a result of the fine-grained automated scoring in the Inq-ITS virtual lab, we were able to determine that students had difficulties with the math practice of constructing graphs and equations relative to other inquiry practices, particularly in the context of graphing an inverse square relationship. Additionally, we found that the sub-practice of identifying the type of math relationship in the graph was particularly challenging for students. This level of assessment is critical as we develop auto-scaffolding to ensure students receive support on these difficult practices.

References

De Bock, D., Neyens, D., & Van Dooren, W. (2017). Students' ability to connect function properties to different types of elementary functions: An empirical study on the role of external representations. *International Journal of Science and Mathematics Education, 15*(5), 939-955.

Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of LS*, *22*(4), 521-563.

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments, 5*(1), 161-179.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: NAP.

Reimers, F. M., & Schleicher, A. (2020). *A framework to guide an education response to the COVID-19 Pandemic of 2020*. OECD.

Sao Pedro, M. A., & Betts, C. (2019). *Real-Time formative assessment of NGSS mathematics practices for high school physical science (Report).* Berlin, MA: Apprendis.

Vitale, J. M., Applebaum, L., & Linn, M. C. (2019). Coordinating between graphs and science concepts: Density and buoyancy. *Cognition and Instruction, 37*(1), 38-72.