April 2019
Conference: American Educational Research Association
At: Toronto, Canada

Can an Alerting Teacher Dashboard Improve How Teachers Help Their Students Learn Science Inquiry Practices?

Michael Sao Pedro[1], Janice Gobert[2], and Rachel Dickler[2]
[1]Apprendis, [2]Rutgers Graduate School of Education

Introduction

The Next Generation Science Standards (NGSS, 2013) emphasize rich integration of authentic science inquiry practices with science content knowledge to promote deeper understandings and flexible transfer of practices and knowledge (Hilton & Honey, 2011). Meeting this goal poses significant challenges for teachers (Hofstein, & Lunetta, 2004; Shute, 2008), particularly for the science practices (Fadel et al., 2007; Duschl & Gitomer, 1997; Edelson et al., 1999) because there is lack of resources for assessing, monitoring, and scaffolding them. Unless teachers can monitor students individually as they work, it is difficult for them to ascertain where students go awry, particularly if students cannot articulate their difficulties (Aleven & Koedinger, 2000). One way to support teachers is to leverage computer-based learning and assessment environments and augment them with monitoring tools like dashboards (Xhakaj et al., 2018).

While recent developments in teacher dashboards enable real-time monitoring of student performance and actions (i.e. teacher dashboards for mathematics; Holstein et al., 2018), the few dashboards that exist for science (i.e. Acosta & Slotta, 2018) are not specifically designed to support science inquiry practices.

We have recently developed Inq-Blotter, a dashboard that helps teachers with real time formative assessment of inquiry practices, which in turn, helps them support their students on inquiry. Inq-Blotter (Gobert et al., 2018), is a data-driven, web-based real time system that alerts teachers to which students need help and on which inquiry practices. Inq-Blotter is integrated with Inq-ITS (Inquiry Intelligent Tutoring System; Gobert et al., 2012, 2013, 2018), a learning environment aligned to the NGSS in which students "show what they know" by conducting inquiry within virtual labs (Gobert, 2015). In prior work, teachers reported that Inq-Blotter's alerts made formative assessment of science practices possible in a traditional classroom setting. Students also welcomed the proactive help from their teachers and found it useful (Gobert et al., 2018). The present study builds on this work and examines whether help from teachers, triggered by alerts, improves students' competencies on the next task for the inquiry practice on which they were helped. We hypothesized that giving teachers *immediate* and *actionable* information on when and how students were struggling would enable them to provide targeted feedback. This, in turn, would drive students' learning of inquiry, as evidenced by better performance on their next inquiry task. In order to investigate this hypothesis, we conducted a randomized study in which we explored the impact of teachers' help when guided by versions of Inq-Blotter with alerts (Full) versus without alerts (Minimal).

Method

Participants include 8 teachers from public middle schools in the Northwest and Northeast United States and 10 of their class periods, namely, 286 6^{th}-8^{th} grade students. Each class period had between 25 and 32 students. Students completed one of the following labs based on their teacher's preference: Intro to Predation, Plate Tectonics (Divergent), Plate Tectonics (Convergent), Genetics, Ramp, Flower Growth, and Evolution. Teachers were randomly assigned to either the intervention (i.e. Full Blotter) or control condition (i.e. Minimal Blotter).

Materials

Inq-ITS. Inq-ITS (Inquiry Intelligent Tutoring System; Gobert et al., 2013) is an online learning and assessment environment for middle school science. Within Inq-ITS, students navigate virtual labs in which inquiry investigations are carried out based on goals aligned to the NGSS (NGSS, 2013). This performance-based assessment allows us to capture students' competencies for inquiry practices beyond the capabilities of traditional multiple choice tests (Lee et al., 2010). Each Inq-ITS virtual lab has four stages (see Figure 1). In the first stage, students form a question/hypothesis based on a given goal. Students then move on to the data collection stage where they conduct an experiment to investigate the goal. The third stage involves analyzing data collected during the experiment, and the final stage is where students communicate their findings. All student actions across the first three stages of an Inq-ITS activity are automatically stored and evaluated within log files (we are working on integrating automated scoring into the final stage; Li et al., 2017). Specifically, Inq-ITS automatically assesses inquiry practices of forming questions, collecting data, and analyzing data (see measures section for more details). Inq-ITS also provides real-time hints to students based on the automated, real-time assessment of their inquiry practice competencies (Sao Pedro et al., 2013, 2016).

Full Blotter. Teachers in the Full Blotter condition were given the complete version of Inq-Blotter. Inq-Blotter provides alerts to teachers as to which students are struggling in Inq-ITS and *exactly how they are struggling* on the practices, as they work. It works by synthesizing the assessments generated on forming questions, collecting data, and analyzing data from Inq-ITS into meaningful, actionable alerts. Each alert on the Inq-Blotter dashboard for individual students contains: a diagnosis, overall performance, prior alerts, activity progress, and number of hints received from the system. For example, Figure 2 provides a visualization of the Blotter dashboard after a teacher selected student "John Marcone's" alert (center panel), which resulted in a display of detailed information about John's current difficulties with hypothesizing and his overall performance across prior activities (right panel). Teachers had the option to view alerts on their mobile device by recency (i.e. most recent alerts appear first) or by type (i.e. alerts grouped by inquiry practice). All teacher actions (i.e. opening an alert, resolving an alert, etc.) within Full Blotter are automatically logged.

Minimal Blotter. For the control condition, the teachers were given a "featureless" version of Inq-Blotter with only the class list and the list of inquiry practices that Inq-ITS assesses. Teachers did not receive any alerts and did not have access to any information regarding student performance within Inq-ITS through the Minimal Blotter dashboard. This minimal design allows for explicitly comparing whether the contextual, actionable alerts provided in the Full Blotter condition provide any advantage to teachers in terms of helping students versus teachers in the Minimal Blotter condition who would help students based only on raised hands. Teachers checked off which student s/he helped and on which inquiry practices. All teacher actions (i.e. selecting a student name, checking off help for a particular practice for a student, etc.) are automatically logged.

Procedure

All teachers, regardless of condition, were given an introduction to Inq-ITS, our student system. All teachers were then randomly assigned to one of two conditions, minimal Blotter (*control)* or Full Blotter (*intervention)*.

Students. Students (regardless of teacher condition) were introduced to Inq-ITS by their classroom teacher during one of their regular class periods. Students then conducted inquiry with one Inq-ITS virtual lab (e.g. in the Density Virtual Lab, students completed three lab activity investigations to determine how shape of a container, amount of liquid, and type of liquid relate to density of the liquid). Students could receive help from the teacher as they worked.

Full Blotter Teachers. The teachers in the intervention condition were introduced to the full version of Inq-Blotter, our teacher dashboard. Teachers then spent one class period using Full Blotter with their students as students completed their virtual lab in Inq-ITS. The teachers were instructed to help students as they worked in Inq-ITS based on the alerts that were received in Inq-Blotter. Specifically, teachers would walk around the classroom and help students who were identified through Inq-Blotter as requiring assistance.

Minimal Blotter Teachers. Teachers in the control condition were instructed by a researcher on how to track the students they helped using the Minimal Blotter dashboard. Teachers then spent one class period using Minimal Blotter with their students as students completed their virtual lab in Inq-ITS. The teachers were instructed to help students as they worked in Inq-ITS based on students raising their hands or indicating that they required assistance.

Measures and Data triangulation

Student performance on science inquiry practices is automatically assessed within Inq-ITS and evaluations are stored within logs. The key practices scored within the system for each lab activity include:

- Forming questions: scored based on whether students could identify independent variables, dependent variables, and whether their IV and DV matched their experimental goal (Gobert et al., 2012)
- Data collection: scored based on whether students design controlled experiments and collect data to test their hypothesis (Sao Pedro et al., 2013)
- Analyzing data: scored based on whether students' claims contained an IV and DV, if their claim reflected the trends in the data, if their claim correctly supported/refuted their hypothesis, and if they warranted their claims with controlled data as evidence (Moussavi et al., 2016).
- Total Inquiry score: average of the three inquiry practice scores.

We then triangulated the log files generated by teachers' use of Full or Minimal Inq-Blotter with evaluations of students' inquiry within Inq-ITS. Within each classroom period, we examined student performance lab-by-lab, and identified whether or not the student was helped by the teacher (based on the teacher logs). This enabled us to determine whether a teacher's help improved student performance on subsequent labs for the practices on which they were helped.

Results

We conducted an exploratory descriptive analysis that compared whether students who were helped in the *intervention condition* (full Inq-Blotter) performed better on subsequent inquiry tasks on the practice on which they were helped than students who were helped in the *control condition* (minimal Inq-Blotter). We predicted that students who were helped by teachers with Full Blotter would demonstrate greater performance gains because they had access to detailed information on students' specific area of difficulty(ies).

In Figures 3-6 below, we show student performance trajectories across labs both on total inquiry score (aggregated), and on each practice (forming questions, collecting data, and analyzing data). At lab 1, no students received prior help from the teacher, so we treated this as a baseline using all students from each condition. However, for labs 2 and 3**,** we use data *only* from students who were helped by the teacher in each condition. Note, the number of students available for each data point in these figures varies depending on the number of tasks completed by students, and whether students had received help from the teacher in prior labs.

As shown in Figure 3, by the third task students who were helped by the teacher who had access to alerts in the full intervention condition ($M = 0.83$, $SD = 0.20$) performed better than those students who were helped in the control condition (*M* = 0.66, *SD* = 0.37). When we disaggregated by the different inquiry practices, we found that there were no major differences between groups and performance was high for both groups for forming questions (Figure 4) and data collection (Figure 5). For data analysis, those who received help from the teacher in the Full Blotter condition had much higher scores (*M* = 0.83, *SD* = 0.26) than students who received help in the control condition (*M* = 0.22, *SD* = 0.00); Figure 6. These initial analyses suggest that Full Blotter alerts may result in teacher help that benefits student performance for the practice of data analysis, a particularly complex practice (i.e. Gobert et al., 2015).

Scholarly Significance

Our research begins to address the effectiveness of the teacher dashboard, Inq-Blotter that works in conjunction with Inq-ITS virtual labs that auto score science practices. We hypothesized that teacher help guided by the actionable alerts in the Full Blotter condition (versus teacher help based only on raised hands in the Minimal Blotter condition) would better support student inquiry performance. This exploratory study suggests that for the practice of data analysis, students who were helped by a teacher who used Full Blotter demonstrated better performance by their third lab activity compared to students whose teachers used Minimal Blotter. In the future, we will rerun this study and collect more data to see if these effects are found with new groups of students. We will also look into a mixed effects model and clustering analysis to examine if there are significant differences in the effects between teachers. Based on extensive data collection with Inq-ITS on our project writ large, we know that students benefit greatly from 3+ activities (Gobert et al., 2018), thus, it is anticipated that with a similar level of usage, our results regarding Blotter would be similarly strong. We are also currently collecting discourse data between teachers and students (Sao Pedro et al., 2016) in order to better understand the influence of the content of teacher help on student performance. Overall, the Inq-Blotter dashboard demonstrates potential to support teachers' as they guide their students through inquiry investigations.

References

Acosta, A. & Slotta, J. (2018). Orchestration supports for knowledge integration in a blended learning community curriculum for Grade 12 Biology. In the *13th Annual International Conference of the Learning Sciences* (pp. 1262-1263).

Aleven, V., & Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help?. In, *Intelligent tutoring systems* (pp. 292-303).

Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment*, *4*(1), 37-73.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, *8*(3-4), 391-450.

Fadel, C., Honey, M., and Pasnick, S. (2007). Assessment in the Age of Innovation, *Education Week, 26* (38), 34-40.

Gobert, J. (2015). Microworlds. *Encyclopedia of Science Education*, 638-639.

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. Educational Psychologist, 50(1), 43-57.

Gobert, J., Moussavi, R., Li, H., Sao Pedro, M., & Dickler, R. (2018). Real-Time Scaffolding of Students' Online Data Interpretation During Inquiry with Inq-ITS Using Educational Data Mining. Invited book chapter in Abul K.M. Azad, Michael Auer, Arthur Edwards, and Ton de Jong (Eds), *Cyber-Physical Laboratories in Engineering and Science Education*. Springer.

Gobert, J. D., Sao Pedro, M. A., Baker, R. S., Toto, E., & Montalvo, O. (2012). Leveraging educational data mining for real-time performance assessment of scientific inquiry skills within microworlds. *JEDM-Journal of Educational Data Mining*, *4*(1), 111-143.

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 the Learning Sciences*, *22*(4), 521-563.

Hilton, M., & Honey, M. A. (Eds.). (2011). *Learning science through computer games and simulations*. Washington, DC: National Academies Press.

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28-54.

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Informing the design of teacher awareness tools through causal alignment analysis. In the *13th Annual International Conference of the Learning Sciences* (pp. 104-111).

Lee, H. S., Liu, O. L., & Linn, M. C. (2011). Validating measurement of knowledge integration in science using multiple-choice and explanation items. *Applied Measurement in Education*, *24*(2), 115-136.

Li, H., Gobert, J., & Dickler, R. (2017). Automated assessment for scientific explanations in on-line science inquiry. A. Hershkovitz & L. Paquette (Eds.), Proceedings of the 10th International Conference on Educational Data Mining (pp. 214-219). Wuhan, China: EDM Society.

Moussavi, R., Sao Pedro, M., & Gobert, J.D. (2016). Evaluating the Efficacy of Real-Time Scaffolding for Data Interpretation Skills. *Paper presented as part of LaMar, M.'s symposium on Cognitive Models for Assessment at the Annual Meeting of the American Education Research Association*. Washington DC.

Next Generation Science Standards Lead States. (2013). *Next Generation Science Standards: For states, by states.* Washington, DC: The National Academies Press.

Sao Pedro, M.A. (2016). Inq-Blotter: Revolutionizing How Teachers Identify and Support Students Needing Help During Inquiry. Awarded by US DoE Institute of Education Sciences (ED-IES-16-C-0014).

Sao Pedro, M., Baker, R., Gobert, J., Montalvo, O., & Nakama, A. (2013). Leveraging Machine-Learned Detectors of Systematic Inquiry Behavior to Estimate and Predict Transfer of Inquiry Skill. *User Modeling and User-Adapted Interaction*, *23(1)*, 1-39.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78(1)*, 153-189.

Sosnovsky, S., Dietrich, M., Andrès, E., Goguadze, G., & Winterstein, S. (2012). Math-Bridge: adaptive platform for multilingual mathematics courses. *21st Century Learning for 21st Century Skills* (pp. 495-500). Springer Berlin Heidelberg.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing*, 1-16.

Xhakaj, F., Aleven, V., & McLaren, B. M. (2018). Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning. In, *European Conference on Technology Enhanced Learning* (pp. 315-329). Springer, Cham.

Figures

Figure 1. The four stages of an Inq-ITS lab activity: (a) forming questions/hypothesizing, (b) collecting data, (c) analyzing data, and (d) communicating findings.

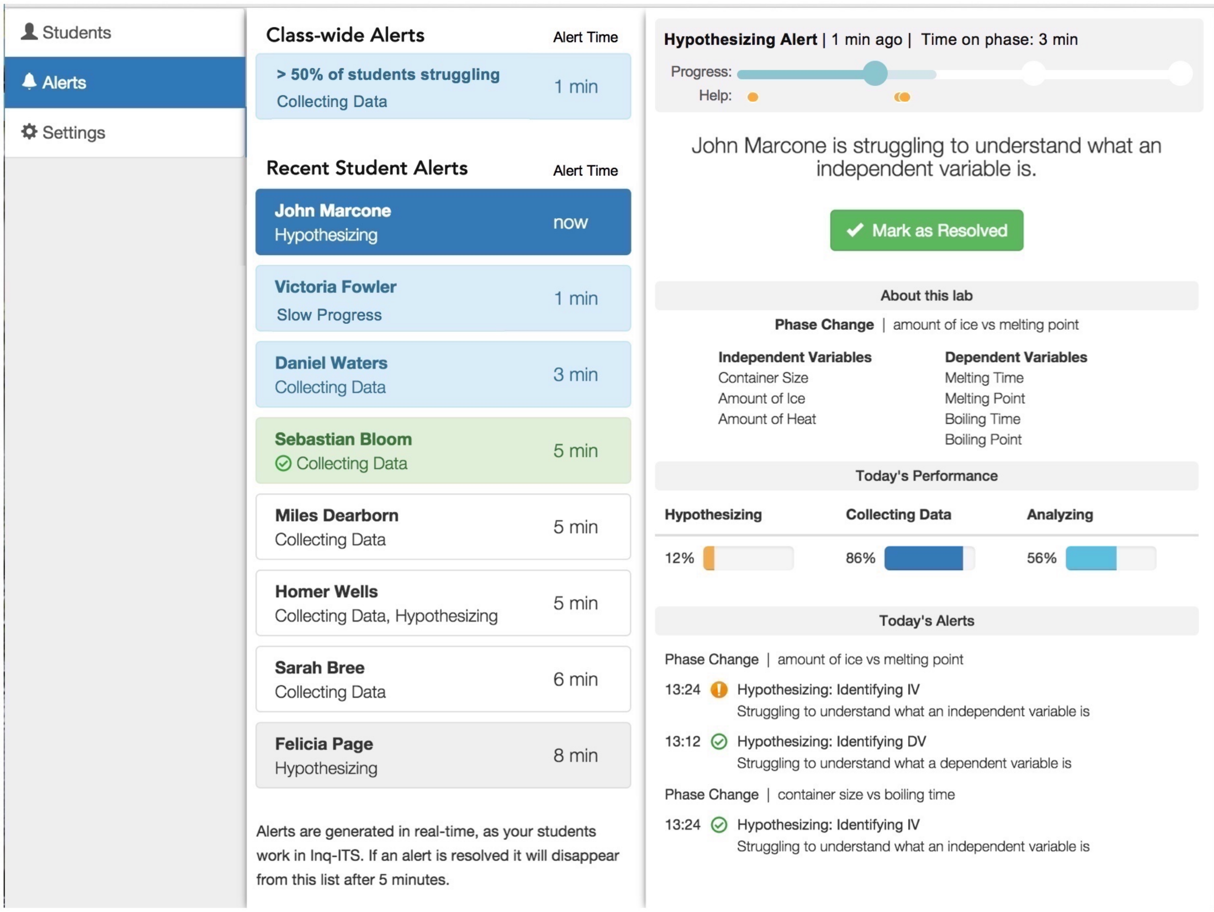

Figure 2. The Inq-Blotter Teacher Alerting Dashboard notifies teachers when students struggle as they conduct inquiry. When a teacher selects an alert (e.g. John Marcone's alert for hypothesizing), contextual information appears which includes a diagnosis of exactly how John is struggling.

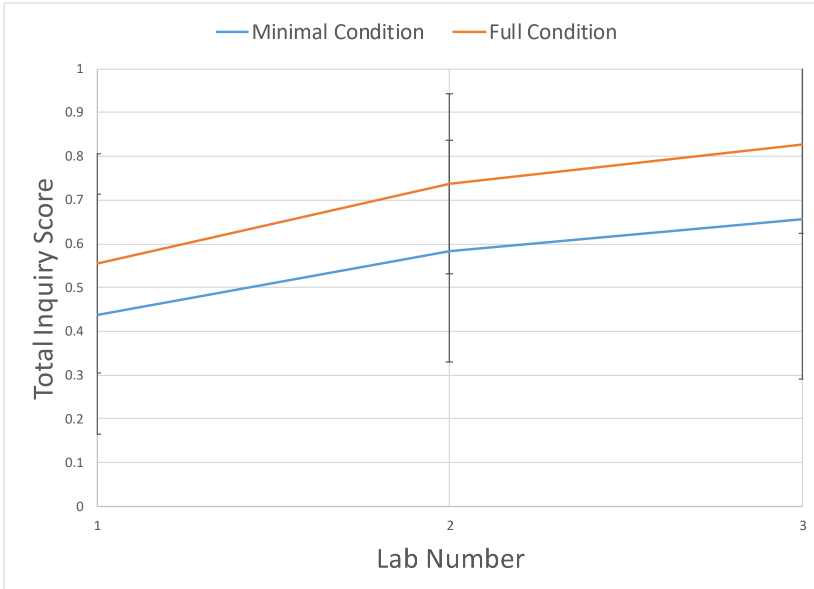

Figure 3. Performance trajectories on overall inquiry score for students in the minimal control and full intervention conditions over 3 lab activities.

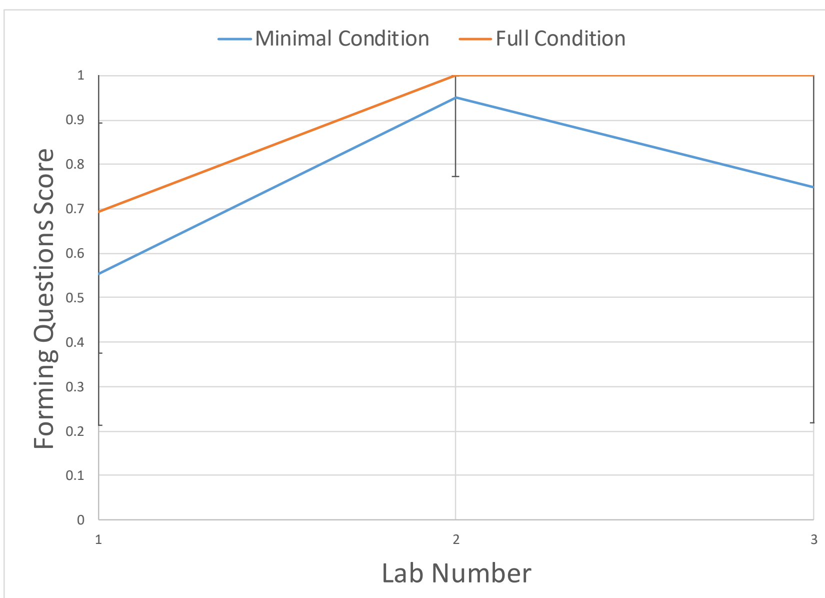

Figure 4. Performance trajectories for the forming questions practice in the minimal control and full intervention conditions over 3 lab activities.

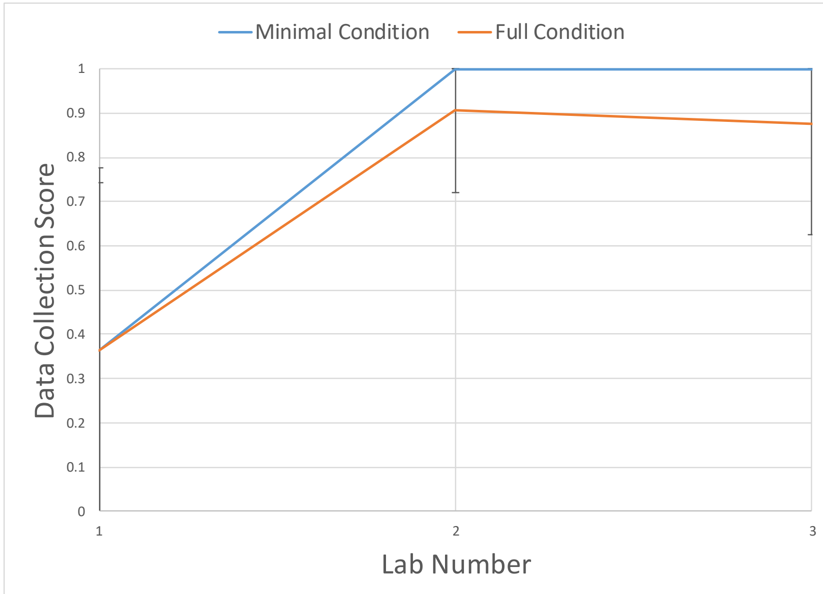

Figure 5. Performance trajectories for the data collection practice in the minimal control and full intervention conditions over 3 lab activities.

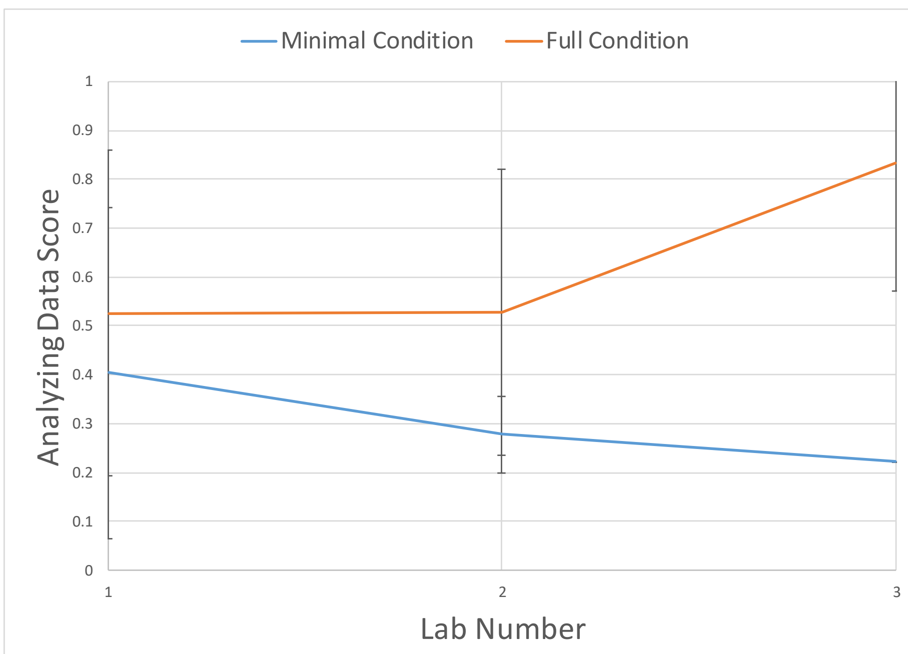

Figure 6. Performance trajectories for the analyzing data practice in the minimal control and full intervention conditions over 3 lab activities.

Acknowledgements
This research is funded by the U.S. Department of Education (R305A120778). Any opinions expressed are those of the authors and do not necessarily reflect those of the funding agencies.