AN INTELLIGENT TUTORING SYSTEM AND TEACHER DASHBOARD TO
SUPPORT STUDENTS ON MATHEMATICS IN SCIENCE INQUIRY

By

RACHEL DICKLER

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Education

Written under the direction of

Janice Gobert

And approved by

_____________________________________

_____________________________________

_____________________________________

_____________________________________

New Brunswick, New Jersey

October 2021
New educational technologies present an opportunity to help teachers monitor and support their students remotely in Science, Technology, Engineering, and Mathematics (STEM). However, most technologies do not have the capacity to comprehensively assess and report on students’ critical 21st century practice competencies (such as those described in the Next Generation Science Standards including Using Mathematics). The science inquiry dashboard, Inq-Blotter, has recently been extended to provide alerts to teachers on high school students’ performance on practices involving mathematical competencies as students complete investigations in the Inquiry Intelligent Tutoring System, Inq-ITS. Inquiry competencies are scored at a fine-grained, sub-component level by Inq-ITS’ underlying algorithms. The logging infrastructure in both the teacher dashboard and student platform allows for investigating how these tools are implemented across contexts, especially when triangulated with additional data including teacher-student discourse, as in the present work.

In this dissertation, a design-based research project was conducted over two phases to characterize and evaluate the remote use of Inq-Blotter’s alerting feature for
math practices to inform future iterations of the dashboard alert design (i.e., embedding Teacher Inquiry Practice Supports (TIPS) for Using Mathematics within the dashboard’s alerts). Specifically, the present work: 1) examined the use of Inq-Blotter in remote contexts and students’ corresponding performance on math practices (which informed the design of Inq-Blotter alerts with TIPS for math practices), 2) examined if remote support elicited by alerts with TIPS was associated with student improvement on mathematical practices, and 3) compared the remote teacher discourse elicited by alerts with TIPS relative to the teacher discourse with basic alerts using Epistemic Network Analyses (Shaffer et al., 2016). These studies provide valuable data on how technologies can help realize the vision of the NGSS in remote settings by supporting teachers’ pedagogical practices for inquiry to promote students’ learning of complex practices such as Using Mathematics in science inquiry.
ACKNOWLEDGEMENTS

The present dissertation was funded by a P.E.O Scholar Award. The material in this dissertation is also based upon work supported by the National Science Foundation under grant numbers 1629045 and 1902647, and Institute of Education Sciences under grant numbers EDIES16C0014, EDIES15C0018, 91990018C0022, and 91990019C0037. Any opinions expressed are those of the author and do not necessarily reflect those of the funding agencies.

Thank you so much to my Ph.D. advisor, Dr. Janice Gobert, for your guidance and feedback throughout the dissertation process. You have spent endless hours reviewing and providing support on materials from conference and funding submissions to the final dissertation papers. I truly appreciate all of your time and encouragement, as well as the opportunities that you have provided to help me further develop as a scholar. Completing a dissertation during a pandemic would not have been possible without your advice, expertise, and encouragement.

I would also like to thank my committee members, Dr. Drew Gitomer, Dr. Juan Pablo Mejia-Ramos, and Dr. Arthur Graesser. Thank you for taking the time to review my work from the proposal stage through to the data defense and final defense. Your suggestions at each stage of the dissertation have been key to strengthening the analyses and contributions of the present work. I have learned so much from your valuable insights and expertise.

I have also received extensive support with navigating professional components of the Ph.D. process from faculty and staff at the Rutgers Graduate School of Education including Dr. Angela O’Donnell, Colleen McDermott, and Nicole Symonds.
From technological support to moral support, I am so appreciative of the members of the Rutgers Inq-ITS team and Apprendis team. Many technical components of the present dissertation would not have been possible without Dr. Michael Sao Pedro and your amazing ideas behind the Inq-ITS high school math project. I would also like to thank Charity Staudenraus, who patiently walked through so many tutorials when I first joined the team and has provided invaluable support on a number of research projects. I also learned so much from Dr. Haiying Li regarding automated scoring of student writing during her time as a post-doctoral researcher with the team. The development of the technologies used within this dissertation and corresponding graphics would not have been possible without the help of Cameron Betts, Michael Brigham, Jason Kleban, and Patrick Rowan.

The graduate students on the Rutgers Inq-ITS team have been amazing research partners. It has been incredible working with Amy on elements of this dissertation as well as several other research projects! Thank you to Joe who also helped with coding discourse in the present dissertation and provided support on many conference submissions. It has also been a pleasure working with Mariel, Huma, and Jeremy as well as our undergraduate research assistants on additional research projects.

There are so many fellow Rutgers GSE students who have supported me along the way. A huge thank you to Kristen who spent hours in virtual study sessions writing with me and talking through research ideas as well as incredible GSE colleagues including Christian, Danielle, Brittany, Marina, Shari, Na’ama, Olivia, and Emily. Thank you to the members of my Ph.D. cohort - including Molly, Ruby, Martha, Ana - who have been supportive through our classes together and each step of the program.
I have also learned about navigating the Ph.D. process from amazing GSE alumni and senior students. Thank you to Roberta for all of your advice and support! I also appreciate all the guidance from Hebbah, Veronica, Randi, Carol, Moraima, and Jeanette.

Thank you to my undergraduate thesis advisor, Dr. Lisa Grimm, for your encouragement and guidance in taking the first steps to pursue this degree. You taught me so many important skills that I have used throughout the dissertation process!

Thank you so much to the P.E.O. Chapter AD for your support on my dissertation, especially Carrie and Carrol for checking in throughout the final year of writing.

The work in this dissertation would also not have been possible without the support I received from friends and family. Thank you to the best friends a person could ask for, Michelle and Danny, for listening to presentations, running mock-interviews, and making me laugh through everything, you are truly incredible! I have been so fortunate to have friends like Michelle, Ashli, Alissa, and Tara, as well as my TCNJ Unified Sports team family and the Livingston Apartments team who have been an amazing support system!

There are not enough words to describe how appreciative I am of the family support I have received! My Grandmother, Esther, has truly been one of my greatest cheerleaders and been so encouraging throughout the process! Thank you to the Dickler family crew, Shawn, Mary Beth, Logan, Andrew, and Colleen who have provided wonderful advice at each step in the program and always put a smile on my face! Also Shari, Jessica, and Lauren - thank you so much for all of the love and support!
And last, but far from least, my parents - Ellen and Doug - thank you for everything. Thanks for listening to countless research ideas and practice talks, embracing new challenges (especially during quarantine), and celebrating each of the milestones along the way. You have been the best moral support I could ask for and words cannot express how appreciative I am of everything you have done to help me get to this point!
# TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION ................................................................. ii

ACKNOWLEDGEMENTS .............................................................................. iv

LIST OF TABLES ......................................................................................... xi

LIST OF ILLUSTRATIONS .......................................................................... xii

Introduction ....................................................................................................... 1

Theoretical Framework .................................................................................... 4

STEM Education & Science Inquiry ............................................................... 4

Mathematics Use in Science Inquiry ............................................................... 6

Scaffolding in Science Inquiry ........................................................................ 8

Automated Scaffolding ................................................................................... 10

Teacher Scaffolding ...................................................................................... 11

Teacher Dashboards ...................................................................................... 16

Dashboard Evaluation Studies ....................................................................... 19

Inq-ITS and Inq-Blotter ................................................................................... 22

Present Dissertation ....................................................................................... 27

Phase 1: Inq-Blotter Alerts for Using Mathematics ........................................ 27

Phase 2: Inq-Blotter Alerts with TIPS for Using Mathematics ....................... 29

Phase 1: Study 1 ............................................................................................ 32

Abstract .......................................................................................................... 32

Introduction to Phase 1 Study 1 .................................................................. 32

Methods ......................................................................................................... 37

Participants .................................................................................................. 37

Materials ..................................................................................................... 38

Procedure .................................................................................................... 46

Measures ..................................................................................................... 48

Analyses ...................................................................................................... 53

Results ......................................................................................................... 55

Research Question 1 Results ...................................................................... 56
Research Question 2 Results ....................................................................................... 57
Research Question 3 Results ....................................................................................... 60
Research Question 4 Results ....................................................................................... 65
Discussion .................................................................................................................. 67

Phase 2: Study 2 ........................................................................................................ 71
Abstract .................................................................................................................... 71
Introduction to Phase 2 Study 2 .................................................................................. 71
  Dashboard Evaluations ............................................................................................ 72
  Remote Dashboard Use ............................................................................................ 73
  Prior work Inq-Blotter and Inq-ITS ........................................................................ 74
Development of TIPS for Using Mathematics ............................................................ 76
Methods ..................................................................................................................... 80
  Participants ............................................................................................................... 80
  Materials .................................................................................................................. 81
  Procedure ............................................................................................................... 89
  Measures ............................................................................................................... 90
  Analyses ................................................................................................................ 92
Results ....................................................................................................................... 93
  Research Question 1 Results ................................................................................. 93
  Research Question 2 Results ............................................................................... 95
  Research Question 3 Results ................................................................................. 102
Discussion .............................................................................................................. 105

Phase 2: Study 3 ....................................................................................................... 110
Abstract ................................................................................................................... 110
Introduction to Phase 2 Study 3 ............................................................................... 111
Methods ................................................................................................................... 114
  Participants ............................................................................................................. 114
  Materials ............................................................................................................... 115
  Procedure ............................................................................................................. 120
  Measures ............................................................................................................. 121
  Analyses ............................................................................................................. 122
Results ..................................................................................................................... 127
  Research Question 1 Results .............................................................................. 127
  Research Question 2 Results .............................................................................. 131
  Research Question 3 Results .............................................................................. 134
Discussion ............................................................................................................. 138

Conclusion ............................................................................................................. 142
LIST OF TABLES

Table 1. Demographics of participating schools.......................................................... 38
Table 2. Mathematics practices and sub-practices automatically scored in Inq-ITS........ 49
Table 3. Challenges experienced by teachers when supporting math practices remotely 52
Table 4. Proportion of students with improvement on mathematical practices............. 58
Table 5. Paired samples t-tests for math performance prior to and after teacher help .... 58
Table 6. Proportion of interactions with each support type ........................................... 62
Table 7. Proportion of interactions with each support type in relation to student prior knowledge .......................................................... 63
Table 8. Frequency of teacher challenges with remote dashboard use ......................... 66
Table 9. Inq-Blotter alerts for mathematical practice stages ....................................... 76
Table 10. Table with TIPS for applying equations alerts ............................................. 77
Table 11. Demographics of participating schools .......................................................... 80
Table 12. Participant conditions and details ................................................................. 81
Table 13. Proportion of students with improvement on mathematical practices .......... 94
Table 14. Paired samples t-tests for math performance prior to after teacher help ........ 95
Table 15. Average scores on math practices across activities in each condition ............ 98
Table 16. Comparisons of student performance across activities on math practices ...... 99
Table 17. Inq-Blotter usage data (with and without TIPS) .......................................... 103
Table 18. Examples of interactions elicited by Inq-Blotter alerts with and without TIPS .......................................................... 104
Table 19. Inq-Blotter with TIPS usage data ................................................................. 105
Table 20. Demographics for participating high schools ................................................ 115
Table 21. Teacher discursive support codes ................................................................. 124
Table 22. Example of teacher supports in response to TIPS for Applying Equations.... 130
Table 23. Average frequency of supports when students improved versus did not improve .......................................................... 135
Table 24. Average frequency of supports when students improved versus did not improve (for subset of the data) .......................................................... 136
Table 25. Linear regression model coefficients (frequency to predict improvement) .... 136
Table 26. Linear regression model coefficients (ENA connection weight to predict improvement) .......................................................... 138
LIST OF ILLUSTRATIONS

Figure 1. Conjecture Map for Phase 1 ................................................................. 28
Figure 2. Conjecture Map for Phase 2 ................................................................. 30
Figure 3. Stages of the Inq-ITS Tutorial Lab ....................................................... 39
Figure 4. Stages of the Inq-ITS Ramp with Graphing Lab .................................... 44
Figure 5. Inq-Blotter Teacher Dashboard ........................................................... 46
Figure 6. Student performance on math practices across opportunities ............... 59
Figure 7. Inq-ITS Tutorial lab stages ................................................................. 82
Figure 8. Constructing Graphs stage of the Inq-ITS lab ...................................... 83
Figure 9. Applying Equations stage of the Inq-ITS lab ....................................... 84
Figure 10. Inq-Blotter Teacher Dashboard (without TIPS) .................................... 85
Figure 11. Inq-Blotter with TIPS initial screen ................................................... 86
Figure 12. Inq-Blotter TIPS Orienting Support ................................................... 87
Figure 13. Inq-Blotter TIPS Conceptual Support ............................................... 87
Figure 14. Inq-Blotter TIPS Procedural Support ................................................ 88
Figure 15. Inq-Blotter TIPS Instrumental Support .............................................. 88
Figure 16. Student performance on mathematical practices across activities ........ 99
Figure 17. Student performance on constructing graphs across activities .......... 100
Figure 18. Student performance on applying equations across activities ............ 101
Figure 19. Constructing Graphs (left) and Applying Equations (right) stages of the Inq-ITS Ramp Lab ...................................................................................... 116
Figure 20. Inq-Blotter Dashboard with a Constructing Graphs Alert ................... 117
Figure 21. Inq-Blotter Dashboard with TIPS for Constructing Graphs Alert ........ 119
Figure 22. Four levels of Inq-Blotter TIPS for a Constructing Graphs Alerts ....... 119
Figure 23. Epistemic network of support patterns for teachers with alerts (left) versus alerts with TIPS (right) ................................................................. 127
Figure 24. Subtracted network comparing teacher support in response to alerts (orange) versus alerts with TIPS (blue) .................................................. 128
Figure 25. Epistemic network of support patterns when students improved (left) versus did not improve (right) ................................................................. 131
Figure 26. Subtracted network comparing teacher support when students improved (blue) versus did not improve (red) .................................................. 132
Figure 27. Phase 1 Conjecture Map Results ......................................................... 144
Figure 28. Phase 2 Conjecture Map Results ......................................................... 145
Introduction

The global economy is centered around STEM fields (Noonan, 2017), which has led to a significant increase in the number of STEM-related jobs in the United States (National Science Board, 2018). Students in the United States, however, have performed poorly on international benchmarks for both science (i.e., ranking 18th) and math (i.e., ranking 37th; OECD, 2018), and graduating high school students have lacked the competencies needed to fill jobs in STEM (Noonan, 2017; OECD, 2016b). Therefore, attention must be directed towards supporting students on developing core competencies that are needed for future success in STEM (National Research Council, 2018).

National reform efforts in the United States, such as the Next Generation Science Standards (NGSS, 2013), describe and seek to improve students’ STEM competencies, but it is extremely challenging for teachers to meet the expectations outlined in the standards (Fulmer et al., 2018). In particular, it can be difficult for educators to operationalize and assess complex dimensions such as the science inquiry practices (Pruitt, 2014). Traditional assessments such as multiple-choice tests are often less effective in capturing students’ inquiry practice competencies relative to innovative technologies such as simulations with automated assessment (e.g., Quellmalz et al., 2012, 2013). Many online simulations and resources designed to support teachers in assessing science inquiry, however, do not address the full range of inquiry practices (Davenport et al., 2018; Quellmalz et al., 2013). For instance, there are few resources that support the practice of using mathematics in science contexts (NGSS, 2013; Wilkerson & Fenwick, 2016), which is concerning given that: 1) the competency of using mathematics is particularly relevant to students’ success in high school science coursework (i.e., Physics;
Basson, 2002; Schuchardt & Schunn, 2016), and 2) using mathematics is also one of the most difficult STEM competencies for students to master (Lai et al., 2016; LópezLeiva et al., 2016; McDermott et al., 1987; Potgieter et al., 2008). Even when students’ inquiry practice competencies have been assessed, it can be challenging for educators to determine how to scaffold their students based on students’ specific inquiry difficulties (Luna, 2018; Pruitt, 2014; Talanquer et al., 2013).

The challenges experienced by educators in assessing and scaffolding students on inquiry practice competencies have been further magnified during remote instruction due to the COVID-19 pandemic (Arnett, 2021; Kaden, 2020). In particular, teachers have fewer opportunities for assessing and supporting their students in online settings (Archambault, 2010; Marshall et al., 2020; Means et al., 2021). Therefore, technological resources are needed that can support teachers in comprehensively assessing and scaffolding students on science inquiry practices including using mathematics across learning contexts (Arnett et al., 2021). Teacher dashboard technologies (i.e., teacher monitoring tools; Verbert et al., 2014) have features such as real-time alerts that can guide teacher scaffolding remotely in STEM, but few dashboards and their corresponding student environments have the capacity to fully assess students’ inquiry practice competencies at a fine-grained level (Lajoie et al., 2020; Martinez-Maldonado et al., 2015; Matuk et al., 2016; Tissenbaum & Slotta 2019; VanLehn et al., 2019).

The inquiry intelligent tutoring system, Inq-ITS (i.e., online student learning and assessment system; Gobert et al., 2013; Gobert et al., 2014), assesses the core competencies outlined in the NGSS (2013) and has recently been extended to assess high school students’ use of mathematics during science inquiry (Dickler et al., 2021a; Sao
Pedro, 2018; Sao Pedro & Betts, 2019; Mislevy et al., 2020). Inq-ITS has a corresponding teacher dashboard system called Inq-Blotter (i.e., online teacher monitoring tool; Gobert et al., 2016, 2018; Gobert & Sao Pedro 2016) that provides actionable, fine-grained data within real-time alerts to teachers on students’ difficulties on specific practices and their respective sub-practices (i.e., sub-components). The alerts within the Inq-Blotter dashboard were also recently updated at the middle school level to include Teacher Inquiry Practice Supports (TIPS) that provide instructional supports to guide teacher scaffolding on science practices (i.e., asking questions/hypothesizing, carrying out investigations, analyzing and interpreting data; Adair et al., 2020). TIPS have yet to be developed, however, for the newly added mathematical practice stages in the Inq-ITS high school level labs (i.e., constructing graphs and applying equations).

In the present dissertation, a two-phase design-based research project was conducted to first test the use of Inq-Blotter alerts at the high school level for supporting mathematical practices (Phase 1), and then correspondingly develop and test Inq-Blotter alerts with TIPS for mathematical practices (Phase 2). This work builds on prior literature involving how technologies, such as real-time alerting dashboards, can be used within remote learning contexts to promote students’ science inquiry competencies and prepare them for future success in STEM. In particular, the real-time alerts in Inq-Blotter with fine-grained assessment data (and TIPS instructional supports) at the sub-practice level meet an essential need for supporting STEM educators’ remote assessment and scaffolding of inquiry practices, as detailed in the Theoretical Framework.
Theoretical Framework

The following sub-sections present an overview of the literature regarding: the general movement in STEM education towards assessing and supporting critical science inquiry practices, the operationalization of the practice of using mathematics in science inquiry and corresponding student difficulties, scaffolding students on inquiry difficulties, challenges with scaffolding students’ inquiry difficulties during remote learning, teacher dashboard technologies in STEM and evaluations of dashboard features that address teacher challenges, and prior work on the Inq-Blotter dashboard (and corresponding Inq-ITS student environment). The review demonstrates the need for and potential of technologies, such as teacher alerting dashboards that accompany intelligent tutoring systems, to support students on using mathematics in science inquiry during remote learning.

STEM Education & Science Inquiry

National reform efforts, such as the NGSS (2013), were implemented to create more consistent standards for science with higher expectations for students in response to low student performance in STEM in the United States (NAEP, 2011; OECD, 2016a, 2018). The NGSS (2013) emphasized a multi-dimensional approach to science education that emphasized learning important scientific ideas through engagement in authentic science inquiry (National Research Council, 2012). Specifically, science educators were expected to incorporate the three dimensions of 1) Science (and Engineering) Practices, 2) Crosscutting Concepts, and 3) Disciplinary Core Ideas. The dimension of science practices, however, has introduced a number of challenges for educators including
operationalizing the practices, finding authentic assessments, and supporting the full range of practices (Pruitt, 2014).

One of the primary challenges with the science inquiry practices is that they are not well-defined within the NGSS or associated frameworks (Kuhn, 2005; Pruitt, 2014). Some expectations for inquiry practices (e.g., NGSS Practice 1: Asking Questions and Defining Problems) are more clearly outlined relative to other practices that require further interpretation (e.g., NGSS Practice 5: Using Mathematics and Computational Thinking; Wilkerson & Fenwick, 2016). As a result, there can be differences in definitions of the competencies involved in more complex practices and, correspondingly, variations in alignment of curricular materials (Fulmer et al., 2018). The lack of concrete operationalization of the practices has also made them very challenging to assess (Pruitt, 2014).

It can be difficult to find assessments that capture the inquiry practices because they can be challenging to define (as stated above), and are also complex (i.e., practices involve multiple components and there is not one single approach to successfully engaging in components of an inquiry practice; Baker et al., 2016a). Typical traditional static assessments, such as multiple-choice tests, have been found to be less successful in capturing the intricacies in student performance on science practices relative to other forms of assessments (i.e., interactive simulations; Lee et al., 2011; Quellmalz et al., 2013). Researchers have since developed innovative, three-dimensional assessments that align to the NGSS and capture inquiry practice performance in more authentic ways (Baker et al., 2016a; Clarke-Midura et al., 2012; Harris et al., 2016; Linn et al., 2018; Pellegrino et al., 2014; Rupp et al., 2012). In particular, researchers have begun to unpack
the science practices and apply methodologies such as Evidence-Centered Design (ECD; Mislevy et al., 2012) to inform the design of assessments that measure aspects of practices in addition to disciplinary core ideas and cross cutting concepts (Harris et al., 2016; Mislevy et al., 2017; Quellmalz et al., 2013). There are some inquiry practices, however, that continue to be overlooked across science inquiry assessments.

Assessments for science inquiry often focus on specific practices (e.g., Practice 7: Engaging in Argument from Reasoning; Osborne et al., 2016) rather than the full range of practices outlined in the NGSS and there are certain practices that are rarely captured within assessments. Until recently, there has been minimal attention to operationalizing and developing authentic assessments for NGSS Practice 5: Using Mathematics and Computational Thinking. Researchers have begun to attend to the “Computational Thinking” component of the practice (Basu et al., 2016; Hutchins et al., 2019; Wilkerson-Jerde, 204; Wilkerson & Fenwick, 2016), but attention is needed towards defining and integrating the “Using Mathematics” component into authentic inquiry assessments. This need is heightened by the importance (Basson, 2002; Schuchardt & Schunn, 2016) and difficulty of mathematical skills in science as detailed in the following sections (Lai et al., 2016; LópezLeiva et al., 2016; McDermott et al., 1987; Potgieter et al., 2008).

**Mathematics Use in Science Inquiry**

Prior to the NGSS, the National Science Education Standards (NSES; National Research Council, 1996) noted that there was an important relationship between mathematics and science, but did not explicitly require that mathematics be assessed and supported within science contexts. The NGSS, however, recognized the importance of the skill and included “Using Mathematics and Computational Thinking” as one of the eight
science inquiry practices that students should master in K-12th grade (National Research Council, 2012). While “Using Mathematics” is critical to students’ success in high school level science courses such as Physics and Chemistry (Basson, 2002; Hoban et al., 2013; Sadler & Tai, 2007; Sadler et al., 2014), there is a lack of consensus regarding the exact definition and operationalization of this practice that is needed for assessment purposes, as previously stated. Assessment of using mathematics is necessary to identify student difficulties and provide necessary support. If unassessed and unsupported, mathematics in science can be a barrier to students’ performance in high school science courses as well as college level courses and professional careers in STEM (Gottfried & Bozick, 2016).

There are several mathematical understandings and skills that are essential to science inquiry at the high school level (National Research Council, 2012), including: covariational reasoning (Carlson et al., 2010; Casey, 2015), interpreting graphs and data tables (Lai et al., 2016; Nixon et al., 2016; Planinic et al., 2012; Shah & Hoeffner, 2002), applying ratios and proportions (Nixon et al., 2016; Planinic et al., 2012; Wollman & Lawson, 1978), making predictions using graphs and equations (Casey, 2015; Nixon et al., 2016), constructing graphs (Lai et al., 2016; LópezLeiva et al., 2016; Strobel et al., 2018; Tairab & Khalaf Al-Naqqbi, 2004), and applying equations (Casey, 2015; De Bock et al., 2017; Nixon et al., 2016; Planinic et al., 2012). The present work focuses on the mathematical practices of Constructing Graphs and Applying Equations in science, as these are core competencies that appear within performance expectations across the scientific domains of Physical Science, Life Science, and Earth Science (i.e., HS-LS2-4/HS-PS1-7/HS-ESS1-4: “Use mathematical representations of phenomena to support claims…using algebraic thinking and analysis, a range of linear and non-linear
functions”; NGSS, 2013, pp. 246-289). Constructing Graphs and Applying Equations, however, are both extremely challenging practices for students when applied in science contexts (Lai et al., 2016; LópezLeiva et al., 2016; McDermott et al., 1987; Potgieter et al., 2008).

In terms of challenges related to Constructing Graphs in science, students have difficulty: setting up the axes of their graphs with appropriate variables reflecting the scientific phenomenon under investigation (Lai et al., 2016; Nixon et al., 2016), determining the data to include in their graphs (Lai et al., 2016; Tairab & Khalaf Al-Naqbi, 2004), and understanding the connections between data in their graph and data table (LópezLeiva et al., 2016; Strobel et al., 2018). When Applying Equations, students struggle with: determining the type of mathematical relationships (i.e., functional relationships) that best represent data (e.g., linear, curved, etc.; De Bock et al., 2017; Lai et al., 2016; Shah & Hoeffner, 2002), understanding the role of constants versus coefficients (i.e., slope) in an equation (Lai et al., 2016; Nixon et al., 2016; Planinic et al., 2012), and using best-fit lines to summarize data (Casey, 2015; Nixon et al., 2016).

Assessments are needed that can capture students’ mathematical practice competencies at fine-grained level to target their specific difficulties and correspondingly support, or scaffold, them in the context of science inquiry. The following section provides an overview of scaffolding as well as how it has been applied to address student difficulties with science inquiry practices.

**Scaffolding in Science Inquiry**

Scaffolding is assistance provided to a student to help them complete a task that would otherwise be too difficult to complete independently (Quintana et al., 2004; Tabak
& Kyza, 2018; Vygotsky, 1978). For example, the practices involved in using mathematics in science inquiry (i.e., constructing graphs and applying equations) are often too challenging for high school students to complete independently without external support (Lai et al., 2016; Lin et al., 2013; Nixon et al., 2016). Ideally, scaffolds will result in learning where students are able to complete tasks with increasing independence, and the types of tasks that can be completed with assistance will become more complex (i.e., students will be able to complete tasks with assistance that previously would have been too difficult even with assistance; Vygotsky, 1978). Scaffolding is a critical component of many approaches to education including: cognitive apprenticeship (Collins, 2006; Collins & Kapur, 2014), reciprocal teaching (Palincsar & Brown, 1984), and inquiry learning (Hmelo-Silver et al., 2007). The way scaffolds are implemented and the types of scaffolds provided to students may vary extensively both within and across these methodologies (National Research Council, 2000). For example, scaffolding in the context of science inquiry (e.g., Hmelo-Silver et al., 2007) can take on many forms.

Common approaches to scaffolding students within science inquiry contexts include adding explicit structures to inquiry task materials (Quintana et al., 2004; White & Frederiks, 1998), example-based supports (Bransford & Schwartz, 1999; van Gog & Rummel, 2018), and personalized guidance (Tabak & Kyza, 2018). Additionally, scaffolding can be presented consistently to all students, faded over time, or adapted to the needs of individual students (Puntambekar & Hubscher, 2005; Tabak & Kyza, 2018). Personalized, adaptive guidance based on continuous assessment of student competencies allows for providing support at the point when students need help most (Anderson et al., 1995; Tabak & Kyza, 2018), which is particularly beneficial in complex
settings such as science inquiry where each student may experience unique difficulties that require support (Hmelo-Silver et al., 2007; Puntambekar & Hubscher, 2005). Personalized, adaptive guidance can be provided to students within online environments that give real-time supports based on automated assessment (National Research Council, 2018; Quintana et al., 2004) as well as by teachers based on their assessment of students’ needs (Roschelle et al., 2017; Scott, 1998).

**Automated Scaffolding**

In terms of adaptive personalized scaffolding in online environments, computer-based training programs (or computer-based instructional software; National Research Council, 2018) scaffold students in response to automated assessment of independent problems presented during or at the end of an instructional unit. Additionally, online environments such as intelligent tutoring systems (Corbett & Anderson, 1995; Koedinger et al., 1997; Graesser et al., 2018), provide adaptive scaffolding based on students’ performance on particular knowledge components that are tracked in relation to an ideal model (i.e., model tracing; Anderson et al., 1995) and/or using bayesian techniques (i.e., knowledge tracing; Baker et al., 2016b; Corbett & Anderson, 1995). Scaffolds in intelligent tutoring systems are then provided in the form of personalized supports (i.e., pop-up messages, feedback from a pedagogical agent, dialogue with a conversational agent; Graesser et al., 1999) as well as through varying the presentation of tasks based on students’ level of mastery (Graesser et al., 2018).

While there are several online environments for science inquiry (e.g., PhET, Wieman et al., 2010; Virtual Physics Labs, Darrah et al., 2014; Co-Lab/Go-Lab, de Jong et al., 2013, 2014; BGuILE, Tabak & Reiser, 2008; InquirySpace, Lee et al., 2014;
Thinker Tools, White & Frederiksen, 1998; NEWTON, Lynch & Ghergulescu, 2017), many of these tools lack automated scoring or *personalized adaptive scaffolding* of student inquiry practice performance. When available in online environments such as SimScientist (Quellmalz et al., 2013) and ChemVLab+ (Davenport et al., 2018), studies have shown the potential of personalized scaffolds in supporting students on science inquiry difficulties. These systems present pop-up messages with increased specificity when a student has been assessed as completing an element of the environment incorrectly (e.g., making an inaccurate prediction about a scientific phenomenon; Quellmalz et al., 2013). The challenge with automated scaffolds is that students will sometimes “game the system” (i.e., click through the pop-up messages until they receive the correct answer) or ignore scaffolded feedback (e.g., not read the text in a pop-up message; Baker et al., 2008). Additionally, the text within the automated scaffolds is predetermined, so there are limitations to the extent with which automated scaffolding can be personalized relative to supports provided by teachers (Merrill et al., 1992; Puntambekar & Hubscher, 2005).

Teacher scaffolding can supplement automated scaffolding to allow for further flexibility in how students are supported. Studies have shown that combining teacher scaffolding with automated scaffolding in online inquiry environments can result in reduced student frustration and promote student learning gains (Montrieux et al., 2017; Raes & Schellens, 2016; Wu & Pederson, 2011). Additionally, teacher personalized scaffolding *alone* (Scott, 1998) can be effective in guiding student learning in science inquiry contexts (Hmelo-Silver et al., 2007).

*Teacher Scaffolding*
Teachers often provide adaptive, personalized scaffolds to students after noticing (van Es & Sherin, 2002) student difficulties (Scott, 1998). Studies have shown how teachers’ discursive scaffolds can promote student learning in relation to practices such as designing and carrying out investigations (Talanquer et al., 2013), engagement with evidence (Manz & Renga, 2017), and constructing scientific explanations (McNeill & Krajcik, 2008). Effective teacher scaffolding, however, involves identifying/noticing when students are struggling, determining student difficulties, and determining how to support students on their difficulties (Shulman, 1987; Shute, 2008).

It can be challenging for teachers to notice when students are struggling (Luna, 2018; Sherin et al., 2011; Talanquer et al., 2013; van Es & Sherin, 2002) because students do not always give clear behavioral signs when they are having difficulty with a subject (i.e., raising their hands; Sherin et al., 2011) and may not know when they need help (Aleven et al., 2016). Interventions exist to support the development of teacher noticing skills, such as having teachers review video recordings of their own teaching and reflect on indicators that students were struggling (Chen et al., 2020; Sherin et al., 2011; van Es & Sherin, 2002). These interventions, however, are implemented retroactively and do not provide direct insight into what students might be struggling with or how to support those difficulties in the moment.

Even after identifying students who are struggling, it can be challenging to determine the specific difficulties that students are experiencing with inquiry practices because the practices are difficult to assess, as noted previously (Luna, 2018; Pruitt, 2014; Talanquer et al., 2013). When teaching science inquiry in-person, teachers often rely on assessments centered around hands-on experimentation (e.g., Hofstein & Lunetta,
2004). During hands-on experiments, however, it can be challenging for teachers to keep track of each student’s progress (Deters, 2005) and corresponding summative assessments such as lab reports may not accurately reflect students’ practice competencies (Li et al., 2017). Teachers can use online assessments such as virtual labs within in-person and remote settings, but teachers still must make decisions regarding how to interpret the data from online environments in order to support their students (Crawford et al., 2008; Schifter et al., 2014).

Determining how to scaffold struggling students after identifying their specific difficulties requires Pedagogical Content Knowledge (PCK; Reimann & Markauskaite, 2018; Shulman, 1987). Teachers not only need to understand the competencies involved in a particular subject area, but also strategies to help others learn in that area (Shulman, 1987). In terms of science inquiry, teachers often require support to develop their Pedagogical Content Knowledge for Inquiry (PCKI; Bybee, 2014; Yang et al., 2018). Inquiry practices are more complex than straightforward, rote scientific content understandings (i.e., the focus of traditional science instruction; Pruitt, 2014) and teachers may be unfamiliar with how to support students on certain practice difficulties (Talanquer et al., 2013). Professional development programs can help teachers to improve their PCKI (Penuel et al., 2020; Schifter et al., 2014; Yang et al., 2018) and implement pedagogical practices that are effective in promoting student learning gains in science inquiry. For example, some programs have improved teacher scaffolding of inquiry through training teachers with pre-scripted instructional prompts (Morris & Chi, 2020; Oliveira, 2009). In addition to professional development opportunities, it is valuable to consider how technological tools can be implemented to help guide teacher scaffolding in
real-time in the classroom and further support teachers as well as students in science inquiry (Roschelle et al., 2017). These tools are especially important in contexts such as remote learning where teachers experience further challenges related to scaffolding students (Arnett, 2021).

**Remote Teacher Scaffolding.** Over 85% of students had access to a remote learning option in the United States in the fall of 2020 through winter of 2021 in response to the COVID-19 pandemic (IES, 2021). In particular, many educators had to shift their instruction to an online synchronous format (i.e., students met virtually with their teacher at a scheduled time to learn material; Arnett, 2021), which has added to existing challenges for STEM educators including assessing and monitoring students’ progress, as well as correspondingly scaffolding struggling students.

One of the most common challenges with remote synchronous instruction includes monitoring students’ progress because there are limited types of assessments that can be implemented in online settings (Dukes, 2020; Chatterjee, 2020; Kearns, 2012; Means et al., 2021). For instance, teachers cannot rely on traditional types of formative assessment to determine when students are struggling such as observations of classroom behaviors (Arnett, 2021; Kebritchi et al., 2017; Means et al., 2021) or hands-on laboratory activities (Chatterjee, 2020; Dukes, 2020; Kaden, 2020; Means et al., 2021). While there are numerous educational technologies developed for supporting assessment in STEM education, it can be overwhelming for educators to find tools that meet their instructional needs (Trust & Whalen, 2020) and, as noted previously, few online tools comprehensively assess students’ inquiry practice competencies. Additionally, remote teachers must navigate interruptions to online assessments (e.g., loss of internet; Means et
al., 2021) that can interfere with monitoring student progress. Thus, for a variety of reasons, it is more difficult for teachers to identify if/when students are struggling remotely and provide them with the help they need.

Furthermore, the likelihood of students seeking help from their teacher directly and receiving needed teacher support also differs in the context of remote synchronous learning relative to other formats (Archambault, 2010; Broadbent, 2017; Means et al., 2021). Specifically, students are less likely to ask their teacher for help when they are struggling in a synchronous remote context relative to students with opportunities for face-to-face interactions (i.e., in blended learning contexts; Broadbent, 2017). This could be in-part due to the fewer opportunities for meaningful interactions with instructors during online meetings (Archambault, 2010; Marshall et al., 2020) and because students instead might seek help from other sources such as the internet (Whipp & Chiarelli, 2004). Correspondingly, there is a greater likelihood of academic dishonesty when students are working remotely as students may resort to relying on and using information from the internet in their work (McAllister & Watkins, 2012). Additionally, students in general often may not realize that they are struggling with a subject and will continue to have difficulties until help is provided (Aleven et al., 2016). As a result, there is a greater need for tools that can be implemented in remote settings to help teachers identify when students are struggling in real-time and provide them with support when needed (Shute, 2008).

Educational technologies, such as teacher dashboards, can address these remote instructional challenges by helping STEM teachers identify and provide just-in-time, personalized scaffolding based on students’ performance within online learning
environments (Verbert et al., 2014). The personalized adaptive scaffolding provided by teachers in response to dashboard technologies can supplement automated adaptive scaffolding provided to students within online environments (e.g., Gobert et al., 2018; Holstein et al., 2019), allowing for further flexibility in how students are supported (Merrill et al., 1992; Molenaar & Knoop-van Campen, 2018). It is important to explore how features of dashboard technologies address central scaffolding challenges experienced by teachers across contexts (e.g., identifying students who are struggling, understanding student difficulties, and determining how to support students on their difficulties).

**Teacher Dashboards**

Teacher dashboards are monitoring tools that provide data to teachers on students’ progress and/or performance within online learning environments (Charleer et al., 2014; Verbert et al., 2014). Dashboards can be accessed through applications or web browsers via several different technologies (i.e., computers, tablets, smart phones, smart glasses etc.; Verbert et al., 2014; Holstein et al., 2019). The data presented within dashboard tools rely on an underlying learner model (Anderson et al., 1995) that is constructed based on automated assessment within the student environment. In particular, dashboards make this learner model visible to the teacher through different design features (i.e., an open learner model (OLM); Bull, 2020; Bull & Kay, 2016). The specific features available within dashboards vary depending on the context and student environment (Charleer et al., 2014; West, 2012) but can include graphics indicating progress (e.g., bar charts, pie charts, networks, progress bars, etc.; Martinez-Maldonado et al., 2013; Molenaar & Knoop-van Campen, 2017; Schwarz et al., 2018; Segal et al., 2017), live-
stream newsfeeds containing student contributions (Lajoie et al., 2020), augmented reality indicators of students’ learning state (i.e., using smart glasses; Holstein et al., 2019), and alerts on student progress (Knoop-van Campen & Molenaar, 2020; Tissenbaum & Slotta, 2019; VanLehn et al., 2019). The purpose and use of dashboards can also vary in terms of whether they are designed to support teacher adaptation at the whole class, small group, or individual student level (Charleer et al., 2014; Verbert et al., 2014).

Dashboard tools frequently support whole class orchestration (Dillenbourg, 2013) by providing information to teachers that can be used to inform decisions regarding the adaptation of overall lessons or activities (Martinez-Maldonado et al., 2015; Schwarz et al., 2018; Segal et al., 2017; Tissenbaum & Slotta, 2019; van Leeuwen & Rummel, 2018, 2020; VanLehn et al., 2019). For example, the SAGLET dashboard (Schwarz et al., 2018; Segal et al., 2017) provides summary data regarding class performance on assigned mathematics activities so that teachers can determine the next activities to assign or if additional supports are needed. Tools such as SAIL Smart Space (Tissenbaum & Slotta, 2019) make suggestions about how to divide the class into smaller groups based on students’ prior performance in the system.

Dashboards can also help with targeting instruction for small groups of students (Lajoie et al., 2020; Martinez-Maldonado et al., 2015; Schwarz et al., 2018; van Leeuwen & Rummel, 2018). For instance, the MTFeedback dashboard (Martinez-Maldonado et al., 2015) sends notifications to teachers with information on student groups’ progress as they work on mathematics problems using collaborative tabletops. The HOWARD dashboard (Lajoie et al., 2020) gives a visual representation of the proportion of contributions
students have made within a particular group while engaging in collaborative-reasoning problems in topics related to biology.

Teachers can also use STEM dashboard technologies to inform individualized support to students (Knoop-van Campen & Molenaar, 2020; Matuk et al., 2016; Roschelle et al., 2017). In the WISE dashboard (Matuk et al., 2016) teachers can see the steps of an inquiry lab investigation that students have completed as well as read the text responses that they submit within the lab. In SNAPPET (Knoop-van Campen & Molenaar, 2020), the dashboard displays the number of multiple-choice questions that students answered correctly (either on their first or second attempts). The Lumilo system (Holstein et al., 2018b, 2019) uses symbols to indicate individual students’ real-time learning states (i.e., confused, off-task, etc.) and provides a detailed breakdown of a student’s performance in the environment when selected by the teacher.

While dashboards include features that are designed to support instruction across a variety of levels (whole class, small group, individual), the competencies that are assessed and monitored vary across tools. In particular, there are a number of STEM dashboards that report on students’ content understandings or interactions in mathematics contexts: FACT (VanLehn et al., 2019), Lumilo (Holstein et al., 2019), MTDashboard (Martinez-Maldonado et al., 2015), SAGLET (Segal et al., 2017), Snappet (Molenaar & Knoop-van Campen, 2018), and general dashboards that correspond to math environments (Van Leeuwen & Rummel, 2018, 2020). There are also a number of systems that report on understandings or interactions in science contexts: CK Biology curriculum dashboard (Acosta & Slotta, 2018), HOWARD (Lajoie et al., 2020), SAIL Smart Space’s tablet tool (Tissenbaum & Slotta, 2019), and WISE dashboard (Matuk et
al., 2016). None of these systems, however, report on students’ science inquiry practice competencies, including Using Mathematics in a science inquiry context.

While the dashboards and corresponding student environments reviewed do not explicitly capture students’ inquiry practice performance, the features from STEM dashboards have the potential to address teacher challenges with assessing and supporting students on inquiry practices across contexts. Therefore, it is valuable to explore whether features of STEM dashboards are effective when implemented in actual classrooms (remote and in-person) in order to inform the design and development of dashboards that do address inquiry competencies. Specifically, evaluation studies allow for identifying if dashboard features function as intended and support teachers in actual classrooms in ways that promote student learning. The following section provides an overview of evaluation studies for STEM dashboards.

**Dashboard Evaluation Studies**

When developing technological innovations such as dashboard tools, it is essential to evaluate their effectiveness within authentic classroom contexts to ensure their ecological validity (Laurillard, 2008). Specifically, classroom implementation studies allow for understanding how design features can support common teacher challenges including identifying struggling students, determining student difficulties, and supporting student difficulties. Additionally, implementation studies allow for understanding the corresponding impact of these features on students’ learning outcomes.

Dashboards are often initially tested through methodologies such as interviewing teachers about design mock-ups (Matuk et al., 2016), asking teachers to respond to storyboards (Molenaar & Knoop-van Campen, 2018; van Leeuwen & Rummel, 2018), or
observing how teachers use the dashboard with simulated student data (Holstein et al., 2019; Lajoie et al., 2020; van Leeuwen & Rummel, 2020). These initial studies are valuable for making decisions about the features to include in final design iterations. For instance, a study by van Leeuwen and Rummel (2020) indicated that teachers were better able to identify struggling (simulated) students when the dashboard included a visual cue that alerted them to students’ difficulties with fractions. While this finding provided evidence of the potential for real-time alerting on guiding teacher support, further studies in classrooms with real students were needed to explore how this feature could direct teacher attention and the impact of this attention on student learning.

Studies investigating the effectiveness of dashboards in actual classrooms use data such as logs of teacher actions (Molenaar & Knoop-van Campen, 2017), videos of teacher use of the technology (Holstein et al., 2018b; Martinez-Maldonado et al., 2013; Tissenbaum & Slotta, 2019), and student performance data to evaluate the technologies (Holstein et al., 2018a; Martinez-Maldonado et al., 2015). Classroom implementation studies have shown that STEM dashboards that provide performance data to teachers in real-time (i.e., line charts reflecting accuracy of student responses to questions, maps showing accuracy of student work in relation to an ideal model, alerts showing students struggling with problems, etc.) were effective in directing teachers’ attention towards students who needed help most (Holstein et al., 2018b; Martinez-Maldonado et al., 2013; VanLehn et al., 2019). Dashboards that displayed information on the accuracy of student work were also found to help teachers determine areas of student difficulty (Molenaar & Knoop-van Campen, 2017). Additionally, the use of features such as real-time alerts or notifications containing performance data were found to guide teacher support that
promoted positive student learning outcomes on mathematics (Holstein et al., 2018a, 2019) and physics content (Tissenbaum & Slotta, 2019).

Few dashboard implementation studies, however, have closely examined other critical forms of process data associated with student learning in authentic settings such as classroom discourse (Gee, 2004; Lemke, 1990; Scott, 1998). While some recent studies have explored the types of discursive teacher feedback elicited by dashboards (e.g., Knoop-van Campen & Molenaar, 2020), these studies have yet to examine the relationship between teacher feedback and student performance. It is valuable to look at how teachers’ discursive supports in response to dashboards shape student learning in science (Chin, 2006, 2007; Manz & Renga, 2017; Scott, 1998). Discourse studies vary in how they categorize teacher discursive supports (i.e., the types of knowledge involved in inquiry elicited by questions; Roth, 1996), but tend to use frequencies to represent trends in teacher supports (Chen et al., 2020; Howe et al., 2019; Manz & Renga, 2017; Talanquer et al., 2013).

Other methodologies, such as Epistemic Network Analyses (ENA; Shaffer et al., 2016), have also been implemented to explore patterns in discourse in STEM contexts (Arastoopour et al., 2016; Herder et al., 2018; Siebert-Evenstone et al., 2018). ENA allows for examining connections between ideas in discourse, as well as making both quantitative and qualitative comparisons between patterns in discourse (Shaffer et al., 2016). For example, ENA could be used to map the types of information elicited from teachers by dashboard features. Comparisons could be made between patterns of teacher support associated with improvement or no improvement. This methodology builds on
the use of frequency counts alone by capturing the ways in which particular types of
information are presented in combination in response to dashboard use.

Furthermore, studies have examined the use of dashboards within in-person (e.g.,
Holstein et al., 2018a; Martinez-Maldonado et al., 2015; Tissenbaum & Slotta, 2019) and
asynchronous online settings (e.g., Lajoie et al., 2020; Verbert et al., 2014), but studies
have yet to examine the use of dashboards within synchronous remote settings. As noted
previously, remote instruction introduces additional challenges for teachers and different
dashboard features (or feature designs) may be needed to address these challenges.

It is important to consider how research can address gaps in the development of
dashboards to support science inquiry instruction and learning remotely. In particular, a
dashboard is needed that reports on students’ science inquiry practice competencies with
features such as alerts (to identify struggling students) and displays of performance data
(to determine student difficulties). This dashboard should then be evaluated within
remote contexts with attention to discourse data to understand its’ use in an authentic
classroom setting. The science Inquiry Intelligent Tutoring System (Inq-ITS) and
corresponding dashboard tool (Inq-Blotter) have recently been expanded to address a
range of inquiry practices at the high school level. These technologies have the potential
to address the challenges experienced by STEM teachers in remote settings, but further
work is needed to explore the implementation of these tools. The following sections
provide an overview of prior work on these technologies as well as the present design-
based research project that was conducted to explore their use in synchronous remote
settings.

Inq-ITS and Inq-Blotter
The Inq-ITS (Gobert et al., 2013; Gobert et al., 2014) environment was originally developed for middle school science using evidence centered design (ECD; Mislevy, 2012) to assess and support students on inquiry practices aligned to the NGSS (2013). As students progress through interactive inquiry stages in virtual lab investigations (i.e., asking questions/forming hypotheses, carrying out investigations using realistic simulations, analyzing and interpreting data collected using the simulations, and explaining findings in the claim, evidence, and reasoning format), all of their actions are logged and automatically scored using knowledge-engineered and educational data-mined algorithms (Gobert et al., 2013; Li et al., 2017; Moussavi et al., 2016; Sao Pedro et al., 2013). This automated scoring captures practice performance at the sub-practice level (i.e., each component that makes up the practice stage is scored) and is used to trigger fine-grained, real-time scaffolds to students from a pedagogical agent, Rex. Teachers have the option to have students complete investigations without Rex support or with Rex support. Prior studies have shown that students significantly improve without Rex on inquiry practices over time and across topics because of the structural scaffolds in Inq-ITS (i.e., carefully designed widgets that break down practices into sub-practices), but Rex can help students improve at a faster rate (Li et al., 2018).

Additionally, the automated scoring in Inq-ITS is used to trigger alerts to teachers on students’ specific practice difficulties in the dashboard, Inq-Blotter (Gobert et al., 2016, 2018). The Inq-Blotter alerts for the middle school practices of asking questions, carrying out investigations, and analyzing data were developed based on both observations of classroom implementations of Inq-ITS and interview data (Gobert & Sao Pedro, 2016). In particular, the researchers observed how teachers supported students
without a dashboard and explicitly asked teachers about the data that would be useful to help them address student difficulties. This information was then used to determine the content of alerts in Inq-Blotter. The alerts were reviewed and tested by teachers on paper over several iterations before being implemented into the online dashboard system (Gobert et al., 2018). A randomized controlled study comparing the effects of teacher support elicited by the Inq-Blotter dashboard with alerts versus without alerts (i.e., no performance data) found that students supported by a teacher using Inq-Blotter with alerts improved more so than students helped by a teacher without alerts (Sao Pedro et al., 2019). Studies then examined how teachers were using the alerting dashboard in their classrooms and how this promoted student learning.

In a recent study (Dickler et al., 2019b), researchers examined the relationship between the types of supports middle school teachers provided in response to alerts in Inq-Blotter and student performance on a subsequent activity following teacher assistance. Results showed that while the majority of students who were helped significantly improved on the inquiry practice on which they were helped, teachers tended to provide low-level supports that only oriented students to the practice with which they were struggling (versus high-level procedural or conceptual explanations of practices). Researchers also took a deeper look at the discursive patterns that occurred in response to alerts using ENA (ENA; Shaffer et al., 2016). By applying ENA to the discourse around the implementation of Inq-Blotter alerts, the researchers were able to uncover that the patterns of support that fostered student improvement significantly differed from patterns where students did not improve (i.e., students who did not improve
received primarily low-level orienting supports in combination with other low-level supports; Dickler et al., 2019a).

Another study showed that students maintained their improvement after being helped by a teacher, and that there were distinct differences between the support patterns elicited by Inq-Blotter that resulted in long term improvement or no improvement (Dickler et al., 2021a). This information was essential as researchers continued to iterate and improve on the alerts in Inq-Blotter for the middle school level to elicit the most effective forms of teacher support. In particular, the research team has begun to embed instructional supports within alerts to further guide teachers’ scaffolding at the middle school level (i.e., Teacher Inquiry Practice Supports, TIPS; Adair et al., 2020). These TIPS, however, have yet to be developed for alerts in the newly expanded Inq-Blotter dashboard for high school. Specifically, Inq-Blotter was expanded to accompany Inq-ITS high school labs that include additional inquiry practice stages for using mathematics in science.

The new Inq-ITS high school lab stages were specifically designed to support students on constructing graphs and applying equations in a science context (Sao Pedro, 2018; Sao Pedro & Betts, 2019). The stages were developed using ECD (Mislevy, 2012) over several iterations informed by NGSS (2013) performance expectations, teacher interviews, teacher surveys, student interviews, student surveys, and a review of the literature regarding student difficulties with using mathematics in science (Sao Pedro & Betts, 2019). Studies with the Inq-ITS Newton’s Law of Gravitation high school lab indicated that students struggled significantly with mathematical practices such as Applying Equations (Dickler et al., 2021b). Alerts were constructed for the Inq-Blotter
dashboard to tell teachers when students were having difficulty with a particular sub-practice of one of the new mathematics stages (using the same structure of the original alerts for middle school).

Studies, however, had yet to examine the use of the newly expanded Inq-Blotter dashboard for high school in actual classrooms (including remote synchronous classrooms). The present work explores the implementation of Inq-Blotter alerts to support using mathematics in a remote synchronous setting through a design-based research project.
Present Dissertation

Design-based research (DBR) is an approach to examining the relationship between: features of a tool developed based on theory, the use of the features (captured via process data), and corresponding outcomes to further inform theory (Brown, 1992; Cobb et al., 2003; Design-Based Research Collective, 2003). DBR is a frequently used approach in the field of the learning sciences because it addresses the core tenet of bridging the gap between research and practice (Fischer et al., 2018; Nathan & Alibali, 2010; Nathan & Sawyer, 2014). In particular, DBR studies allow for iteratively developing tools based on both theory and results of implementation studies in authentic settings. To help make the connections between theory, design, and outcomes explicit, Sandoval (2004, 2014) emphasized the value of conjecture mapping. In conjecture mapping, researchers outline their high-level conjectures (i.e., expectations for how practical challenges can be addressed based on theory), design conjectures (i.e., expectations for how embodied features designed based on theory will result in mediating processes that promote learning outcomes), and theoretical conjectures (i.e., expectations for how the mediating processes will promote learning outcomes; Sandoval, 2014). A conjecture map has been developed for each phase of the current dissertation to illustrate the connections between theory and design features in Inq-ITS and Inq-Blotter.

Phase 1: Inq-Blotter Alerts for Using Mathematics

In the first phase of the dissertation, the use of alerts for mathematical practices in Inq-Blotter were explored in remote synchronous contexts (see Figure 1). Specifically, Study 1 was conducted to examine the high-level conjecture that automated assessment within a student environment and real-time alerting in a corresponding teacher dashboard
could address challenges experienced by teachers when supporting students on mathematical practices in science inquiry remotely. The embodiments designed to address these challenges in Phase 1 were mathematical practice stages/tasks in Inq-ITS (i.e., Constructing Graphs and Applying Equations) and real-time alerts in Inq-Blotter with fine-grained information on student difficulties on the math practice stages at the sub-practice level.

**Phase 1 Conjecture Map**

<table>
<thead>
<tr>
<th>Embodiments</th>
<th>Mediating Processes</th>
<th>Theoretical Conjecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inquiry practice stages that capture Using Mathematics in science (i.e., constructing graphs and applying equations)</td>
<td>Students’ actions are logged and automatically scored on math practice stages in Inq-ITS</td>
<td>Valid assessment of students’ mathematical competencies (i.e., scores in Inq-ITS correlate with external assessments)</td>
</tr>
<tr>
<td>Real-time alerts that identify students who are struggling and the math practice on which students are struggling</td>
<td>Teacher identifies students who are struggling</td>
<td>Students improve on the math practice on which they were helped on their next opportunity</td>
</tr>
<tr>
<td>Details within alerts regarding students’ specific difficulties with Constructing Graphs and Applying Equations</td>
<td>Teacher provides discursive support to students on the specific math practice of difficulty</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1. Conjecture Map for Phase 1**

In terms of design conjectures, the new mathematical stages were developed to enable automatic logging and scoring of students’ practice competencies at the sub-practice level. The real-time alerts were developed to help teachers identify struggling students and correspondingly guide teachers in supporting student difficulties with
mathematical practices. Specifically, the diagnostic data on students’ specific difficulties at the sub-practice level were expected to elicit teacher support (as teachers could determine how to support the student based on their PCKI), and also ensure that the support provided by the teacher was targeted towards the difficulty the student was experiencing. Finally, the theoretical conjectures included that mathematical stages with automated scoring of students’ competencies would allow for fine-grained assessment of inquiry practice competencies remotely that correlates with other external assessments. Finally, real-time alerts containing detailed information about student difficulties at the sub-practice level would elicit teacher support that promotes student improvement on the mathematical inquiry practice on which they were helped.

**Phase 2: Inq-Blotter Alerts with TIPS for Using Mathematics**

In the second phase of the dissertation, embedded instructional supports (i.e., Teacher Inquiry Practice Supports (TIPS)) were added to the Inq-Blotter alerts for mathematical practices to further guide teacher scaffolding in remote synchronous contexts (see bolded text in Figure 2). The studies in Phase 2 (i.e., Study 2 and Study 3) were conducted to examine the high-level conjecture that adding instructional supports/prompts within dashboard alerts would address further challenges (identified in Phase 1) experienced by STEM educators teaching remotely. Therefore, the specific embodiment that was added in Phase 2 was the inclusion of TIPS instructional supports within alerts.

The corresponding new design conjectures included that these TIPS would promote higher level supports to students such as conceptual and procedural support, as well as provide enough context for teachers to support the PCKI needed to address
student difficulties. The theoretical conjectures included that teachers would identify and support struggling students based on alerts with TIPS, which would result in student improvement on the mathematical practices on which students were helped. Additionally, the high levels of teacher scaffolding elicited by the TIPS instructional supports would correspondingly result in greater student improvement (relative to when teachers helped students based on basic alerts without TIPS in Phase 1).

**Phase 2 Conjecture Map**

<table>
<thead>
<tr>
<th>High Level Conjecture</th>
<th>Theoretical Conjecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>A student environment and corresponding dashboard with real-time alerts containing data on student performance on math practices in an authentic inquiry environment as well as instructional supports will address challenges experienced by STEM educators when teaching remotely:</td>
<td>Students improve on the math practice on which they were helped on their next opportunity</td>
</tr>
<tr>
<td>• Assessing students on inquiry practices such as Using Mathematics in science</td>
<td>Students experience greater improvement on math practices when helped by a teacher using TIPS</td>
</tr>
<tr>
<td>• Identifying students struggling with math practices in science inquiry</td>
<td></td>
</tr>
<tr>
<td>• Determining student difficulties with math practices in science inquiry</td>
<td></td>
</tr>
<tr>
<td>• Supporting students on math practices in science inquiry at a high level</td>
<td></td>
</tr>
<tr>
<td>• Supporting teachers’ PCKI for using mathematics in science</td>
<td></td>
</tr>
</tbody>
</table>

**Embodiments**
- Real-time alerts that identify students who are struggling and the math practice on which students are struggling
- Details within alerts on students’ specific difficulties with Constructing Graphs and Applying Equations
- Instructional supports (i.e., TIPS) within alerts to guide teacher scaffolding of inquiry difficulties at multiple levels (e.g., orienting, conceptual, procedural, and instrumental) and support PCKI

**Mediating Processes**
- Teacher identifies students who are struggling
- Teacher provides discursive support to students on the specific math practice of difficulty
- Teacher provides higher levels of support to students with TIPS (e.g., conceptual and procedural support)

**Outcomes**
- Students improve on the math practice on which they were helped on their next opportunity
- Students experience greater improvement on math practices when helped by a teacher using TIPS

Figure 2. Conjecture Map for Phase 2

The results of the studies in each phase have important implications for designing dashboard alerts, supporting remote synchronous STEM instruction, and implementing
innovative methodologies to evaluate educational interventions. In particular, each study contributes understandings for how educational technologies can be designed to help realize the NGSS across educational settings.
Phase 1: Study 1

Abstract

Policy documents such as the Next Generation Science Standards (NGSS, 2013) have emphasized the need to support students on science inquiry practices. Teachers, however, face several challenges related to assessing students on inquiry competencies (Pruitt, 2014), especially in remote contexts (Chatterjee, 2020; Dukes, 2020). Innovative educational technologies, such as teacher dashboards and corresponding student online inquiry environments, present an opportunity to address these challenges. The present study examines how a teacher alerting dashboard, Inq-Blotter, can be used in combination with the student platform, Inq-ITS, to support teacher assessment and scaffolding of inquiry competencies in a remote setting. Specifically, this study was conducted to: 1) understand the relationship between teacher use of Inq-Blotter alerts and student learning outcomes on the practices involved in Using Mathematics in science, and 2) examine teachers’ experiences with remote dashboard use to inform the design of future dashboard alerts for supporting inquiry instruction remotely.

Introduction to Phase 1 Study 1

A central goal of Science, Technology, Engineering, and Mathematics (STEM) education is to prepare our students with the competencies needed to succeed in real-world STEM contexts (Noonan, 2017; OECD, 2016b). Policy efforts, such as the implementation of the NGSS (2013), have aimed to meet this goal by emphasizing that science education focuses on supporting students in developing critical skills in addition to content understandings. Specifically, the NGSS outlined eight inquiry practices that students should master by the end of high school including asking questions, planning
and carrying out investigations, and using mathematics and computational thinking (National Research Council, 2012). In order to realize the vision of the NGSS, teachers are expected to assess and support students on these inquiry practice competencies.

Science inquiry practice competencies, however, are not easily captured through traditional assessments designed for measuring science content understandings (Pruitt, 2014). For example, multiple-choice assessments can be insufficient in capturing variations in student inquiry practice performance at a fine-grained, sub-practice level (i.e., assessing components involved in each of the inquiry practices; Lee et al., 2011; Gobert et al., 2018). Additionally, open response items or lab reports that accompany hands-on experimentation may not always reflect students’ actual competencies (Li et al., 2017). Hands-on experiments themselves are difficult to use for assessment because teachers must monitor multiple groups of students simultaneously while also navigating classroom management and safety concerns (Deters, 2005). Therefore, it is important to consider the types of assessments that can be implemented to monitor students’ inquiry competencies across contexts to correspondingly provide support on inquiry difficulties.

Further challenges are introduced when assessing student competencies in remote contexts (as has been the case for many teachers due to the COVID-19 pandemic; Arnett, 2021) because educators have limited opportunities for monitoring student progress without traditional in-person assessments (e.g., hands-on experimentation is not a practical option in a remote context; Dukes, 2020; Chatterjee, 2020; Kearns, 2012). Additionally, educators have to navigate issues related to students cheating within online multiple-choice or open-response assessments (Whipp & Chiarelli, 2004), as well as issues related to fewer opportunities to monitor student progress through informal
observations or one-to-one interactions that are frequently used for formative assessment within in-person settings (Archambault, 2010; Arnett, 2021; Kebritchi et al., 2017; Means et al., 2021).

Fortunately, online science environments provide an opportunity to authentically capture student competencies (Quellmalz et al., 2012, 2013) and have the potential to support student learning remotely because they allow for formatively monitoring active student learning (i.e., within virtual labs; Chatterjee, 2020). In particular, several online science environments (e.g., WISE, Matuk et al., 2016; CK Biology curriculum, Acosta & Slotta, 2018; HOWARD, Lajoie et al., 2020; SAIL Smart Space; Tissenbaum & Slotta, 2019) have corresponding dashboard tools (Dillenbourg, 2013) that provide real-time data to teachers on students’ progress based on automated assessment in the online environment (Verbert et al., 2014). Teachers receive assessment data through the dashboards in the form of visualizations, alerts, and text descriptions (Charleer et al., 2014). While dashboards with these features have been shown to help guide teachers in assessing and supporting students (Lajoie et al., 2020; Roschelle et al., 2017; Tissenbaum & Slotta, 2019), the corresponding student environments are not assessing students’ competencies at science practices at a fine-grained level. Thus, students’ sub-practice performance on critical competencies, such as Using Mathematics (the focus of the present work), cannot be reported within the dashboard to guide teacher support on students’ specific inquiry practice difficulties.

It is important that teachers are able to comprehensively assess students’ inquiry practice competencies on practices such as Using Mathematics that are particularly difficult for students and often require teacher support (Lai et al., 2016; LópezLeiva et al.,
In terms of Using Mathematics, students are expected to be able to engage in practices such as constructing graphs of scientific phenomenon as well as applying equations to model the phenomenon in their graphs (HS-LS2-4/HS-PS1-7/HS-ESS1-4; NGSS, 2013, pp. 246-289). Students struggle with components of each of these practices including determining the scientific variables to place on axes of graphs (Lai et al., 2016; Nixon et al., 2016) and interpreting the type of mathematical relationship between variables in a graph (De Bock et al., 2017; Lai et al., 2016; Shah & Hoeffner, 2002). It is essential to capture and support these difficulties since mathematical practices are critical to success in science coursework (Basson, 2002; Hoban et al., 2013; Sadler & Tai, 2007; Sadler et al., 2014). Furthermore, mathematics is fundamental for acquiring a deep understanding of scientific phenomenon in real-world settings (e.g., in order to understand relationships between scientific variables) and related success in STEM careers (Gottfried & Bozick, 2016).

The Inquiry Intelligent Tutoring System, Inq-ITS, and corresponding teacher dashboard, Inq-Blotter, have recently been extended for high school to support the full range of practices including Using Mathematics (Sao Pedro, 2018; Sao Pedro & Betts, 2019). In the Inq-ITS system, students complete virtual lab investigations with inquiry practice stages and are automatically assessed on the inquiry practices at the sub-practice level (i.e., components of the inquiry practice stages are scored using patented algorithms; Gobert et al., 2014, 2016, 2018; Sao Pedro, 2018; Sao Pedro & Betts, 2019). This automated scoring has been essential for identifying student difficulties with the mathematical practice stages in the Inq-ITS Newton’s Law of Gravitation Lab including Applying Equations (Dickler et al., 2021b). Studies, however, have yet to validate the
interpretation of scores (Kane, 2013) within the mathematical practice stages in new labs such as the Inq-ITS Ramp with Graphing Lab, and examine the impact of the corresponding alerts informed by this assessment in the Inq-Blotter dashboard.

The automated scoring in Inq-ITS is used to trigger real-time alerts to teachers with actionable information on students’ specific practice difficulties in Inq-Blotter (Gobert et al., 2018). Recent work with Inq-Blotter has shown that the dashboard’s alerts are effective for helping teachers to identify and support students on inquiry practices at the middle school level (Dickler et al., 2021a). Specifically, triangulating classroom discourse with log data from both Inq-ITS and Inq-Blotter allowed for understanding how teachers used the dashboard to scaffold their students. While alerts in Inq-Blotter have been designed for student difficulties on the new mathematical practice stages in Inq-ITS, studies have yet to implement and test the effectiveness of these alerts for guiding teacher scaffolding at the high school level. Further, it is essential to examine how both the Inq-ITS and Inq-Blotter technologies function when implemented to support remote instruction, given the ubiquity of remote instruction during COVID-19 (IES, 2021). This work also serves to inform future iterations of the technologies for in-person and remote classroom use.

The present study examined the use of the Inq-ITS Ramp with Graphing Lab and corresponding Inq-Blotter dashboard during remote instruction. In particular, this study sought to validate the interpretations of the automated scoring in Inq-ITS for the new high school mathematical stages of Constructing Graphs and Applying Equations. Additionally, this work aimed to examine whether the alerting features in Inq-Blotter were effective for helping teachers to identify struggling students and provide support
that served to foster student learning of the mathematical practice competencies of interest (i.e., Constructing Graphs and Applying Equations). As such, the following four research questions were examined:

1) Is the interpretation of mathematical practice scores on the stages of Constructing Graphs and Applying Equations in Inq-ITS valid?

2) Are teacher discursive supports (elicited by Inq-Blotter alerts on using mathematics) associated with student improvement on the practice on which they were helped in their next opportunity?

3) What is the relationship between the types of discursive supports teachers provided to students and students’ prior knowledge of using mathematics in science (as measured by the pre-test)?

4) What unanticipated themes emerged in educators’ experiences with using the Inq-Blotter dashboard remotely to support students on mathematical practices?

Methods

Participants

The participants in the study included three teachers and their students from three high schools in the northeastern United States ($N = 119$ students total; see Table 1 for school demographics). All participating high schools were conducting classes completely remotely (synchronously) due to the COVID-19 pandemic and had a one-to-one computer policy (i.e., students could borrow a device from the school). The first teacher, Ms. A, taught two class sections at an urban district ($n = 29$ students; 73% of students were eligible for Free and Reduced Lunch) and conducted her classes virtually using Google Meets. Ms. B taught two class sections at a suburban district ($n = 42$ students; 32% of students were
eligible for Free and Reduced Lunch) and conducted her classes virtually using Zoom with breakout rooms. Mr. C taught three class sections at a suburban district ($n = 48$ students; 8% of students were eligible for Free and Reduced Lunch) and conducted his classes virtually using Google Meets with breakout rooms.

Table 1. Demographics of participating schools

<table>
<thead>
<tr>
<th>Teacher (School)</th>
<th>Ss</th>
<th>FRL</th>
<th>Asian</th>
<th>Black</th>
<th>Hisp.</th>
<th>Nat. Amer.</th>
<th>Pac. Isl.</th>
<th>White</th>
<th>Two or More Races</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ms. A (1)</td>
<td>29</td>
<td>73%</td>
<td>0.4%</td>
<td>45.8%</td>
<td>52.7%</td>
<td>0.1%</td>
<td>0%</td>
<td>0.9%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Ms. B (2)</td>
<td>42</td>
<td>32%</td>
<td>22.5%</td>
<td>11%</td>
<td>21.6%</td>
<td>0.2%</td>
<td>0.2%</td>
<td>39.6%</td>
<td>4.9%</td>
</tr>
<tr>
<td>Mr. C (3)</td>
<td>48</td>
<td>8%</td>
<td>8.1%</td>
<td>2.1%</td>
<td>10.8%</td>
<td>0%</td>
<td>0.2%</td>
<td>78.6%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

*Note.* Ss = number of students, FRL = Free and Reduced Lunch, Hisp = Hispanic, Nat. Amer. = Native American, Pac. Isl. = Pacific Islander

**Materials**

**Inq-ITS tutorial.** All students were assigned a brief Inq-ITS tutorial to complete for homework on their personal device (the tutorial could be accessed through any web browser). The topic of the tutorial was “Flower Growth” (i.e., exploring whether adding salt, sugar, or dye affected the growth of a flower) and required minimal scientific content knowledge so that students could focus on learning the components of the Inq-ITS environment. The tutorial walked students through four basic lab activity stages aligned to the NGSS practices (i.e., asking questions/forming a hypothesis to test, carrying out an investigation using a simulation, analyzing and interpreting data, and explaining findings; see Figure 3). The tutorial enabled students to become familiar with
the embedded tools (i.e., dropdown menus, interactive simulations) of the Inq-ITS lab environment as well as general structure of the lab.

Figure 3. Stages of the Inq-ITS Tutorial Lab

Pre-Test. During the virtual synchronous data collection sessions, students completed an 8-item multiple-choice pre-test embedded within the online Inq-ITS system (see the Appendix). The questions were explicitly compiled and adapted to assess
students’ competencies on constructing graphs and applying equations. Specifically, all questions were developed from prior assessments validated for capturing these competencies within questions that simulated an inquiry context (to the extent possible within a multiple-choice format). These assessments included the Graphing Inventory (Lai et al., 2016), Educational Testing Service (ETS) conversation-based assessments (Liu et al., 2016), ETS Study Companion (ETS, 2017), ETS HiSET Science Practice Test (ETS, 2013), and College Board SAT STEM subject and AP STEM preparation materials (College Board, 2019a, 2019b; see the Appendix). All questions were reviewed by K-16 mathematics and science educators to confirm the content validity (i.e., that the questions addressed the Construction of Graphs and Application of Equations in science, respectively).

Questions 1 - 4 were used to assess Constructing Graph competencies (e.g., selecting controlled trials to include in a graph, understanding the location of a point on the graph, determining the appropriate variables to put on the axes of a graph) and questions 5 - 8 were used to assess Applying Equation competencies (e.g., determining the mathematical relationship between data in a table, determining the mathematical relationship represented in a graph, interpreting the coefficient and constant in an equation, determining the equation that best represents data in a table; see the Appendix). The order of questions and the order of answers was randomized for all students.

**Inq-ITS Ramp with Graphing Lab.** Following the pre-test, students completed the Inq-ITS Ramp with Graphing Lab (i.e., Ramp Lab). This lab was explicitly designed to integrate the use of mathematics within a science investigation and included two new stages (i.e., Constructing Graphs and Applying Equations). The new stages of this lab
were initially developed and tested in prior studies during in-person learning for a different science topic (e.g., Newton’s Law of Gravitation; Dickler et al., 2021b; Sao Pedro & Betts, 2019). The Ramp Lab had three activities where students examined:

1) how the steepness of a ramp relates to the time it takes a sled to reach the end of the ramp (while the length of the ramp stays the same),

2) how the mass of a sled going down a ramp relates to the momentum of the sled when it reaches the end of the ramp,

3) how the roughness of the ramp relates to the acceleration of the sled when it reaches the end of the ramp.

Each of the three Ramp Lab activities has 6 stages aligned to the NGSS (2013) science inquiry practices that each involve several sub-practices (i.e., components involved in each inquiry practice stage; see Figure 4):

1) **Asking questions/Hypothesizing**: students formed a hypothesis about the mathematical relationship between an independent and dependent variable based on a given goal (e.g., if I change the roughness of the ramp, then I will be able to observe that the roughness of the ramp and the acceleration of the sled at the end of the ramp have a linear relationship).

2) **Carrying out an Investigation/Collecting Data**: students ran trials using a simulation to investigate the relationship between the variables that they outlined in their hypothesis (e.g., roughness of the ramp and acceleration at the end of the ramp). The data that they collected were automatically stored in a data table.

3) **Constructing Graphs**: students selected trials from the data they had collected in the Carrying out an Investigation stage to plot in a graph. Students selected the variable
to place on the x-axis of their graph and the y-axis of their graph (the data trials were automatically plotted on the graph).

4) **Applying Equations**: students selected the equation/functional relationship that best represented the data in their graph (e.g., linear, inverse, square/quadratic, inverse square/inverse quadratic, or horizontal equation). Students also determined the coefficient and constant for the equation (the line represented by the equation was automatically plotted on the graph) and checked the fit of the equation line to the graph of the data (the fit of the line to the graph was automatically calculated and stored in a table).

5) **Analyzing Data**: students interpreted the results of their graphs by making a claim about the relationship between the variables, identifying if it was the relationship that they had initially hypothesized, and selecting the graphs and corresponding equations that best demonstrated this relationship.

6) **Explaining Findings**: students wrote an explanation of their findings in the claim, evidence, and reasoning format.

The first four stages of each lab activity had automated scoring of students’ inquiry performance at the sub-practice level (automated scoring was still under development for Analyzing Data and Explaining Findings; see Measure section). The present study focused on students’ performance on the third (Constructing Graphs) and fourth (Applying Equations) stages, which were the newly designed mathematical stages constructed to capture student performance on Using Mathematics in science at the sub-practice level.
During the Asking Questions/Hypothesis stage, students had access to support from a pedagogical agent, Rex (triggered by the automated scoring), who would ensure that students created a question aligned to the goal of the investigation. In particular, Rex would not let students move on to the next stage until they selected the appropriate independent variable and dependent variable based on the activity goal. Rex would provide increasingly directive support to the student until eventually specifying to the student the specific variables to select (Rex was not available in any other stages of the lab). After completing the hypothesis stage, students could not make any changes to their hypothesis but could move freely between stages two through four to collect additional data, revise their graphs, or revise their equations. Students’ performance in the Ramp Lab was used to trigger real-time alerts in Inq-Blotter.

**Asking Questions**

**Carrying out an Investigation**
During data collection for this study, all teachers used the Inq-Blotter dashboard while students completed the Inq-ITS Ramp Lab. The
dashboard was accessed through a web-browser on the teachers’ devices and provided real-time alerts containing fine-grained, actionable information on student difficulties on inquiry practices in the Ramp Lab. All alerts in the dashboard are triggered based on the automated scoring within the Inq-ITS lab inquiry practice stages, which takes place at the sub-practice level (see Measures). At the time of the study, alerts were only available for student performance on the first four stages of each Ramp Lab investigation (i.e., Asking Questions, Carrying out Investigations, Constructing Graphs, and Applying Equations) because automated scoring was still being developed for the fifth (Analyzing Data) and sixth stages (Explaining Findings). Thus, actionable alerts at the sub-practice level were not yet possible for Analyzing Data and Explaining Findings. The present study focused on the use of alerts for the Constructing Graphs and Applying Equations stages.

Anytime students were assessed as performing below a proficiency threshold for a practice in Inq-ITS (i.e., scores of less than 75% on a practice), the teacher would receive an individual student alert on the Inq-Blotter dashboard (see Figure 5). The individual student alerts contained information regarding the inquiry practice stage that the student was struggling with, how long the student was on the stage, the specific sub-practice of the stage that was difficult for the student, other practices that the student was having difficulty with (and prior alerts that had appeared for the student), the activity the student was working on, and other contextual information. Teachers could also receive whole class alerts anytime more than 50% of participating students were assessed as performing below the pre-determined proficiency threshold for the practice (e.g., over half the students had a final score of less than 75% on the practice), as well as slow progress alerts anytime a student was spending more than 5 minutes without any action on an
inquiry practice stage. Teachers could adjust the percent of students needed to trigger a whole class alert and the time needed to trigger a slow progress alert.

![Figure 5. Inq-Blotter Teacher Dashboard](image)

The Inq-Blotter dashboard also included different options for how the alerts were organized on the screen (see left panel of Figure 5), including: alerts by time (i.e., the most recent alerts appeared at the top of the page), alerts by type (i.e., alerts were organized by inquiry practice stage), and students (i.e., all active students were listed alphabetically and could be selected to view their performance data).

**Procedure**

Prior to data collection, the principal investigator of the present dissertation (myself) met individually with each teacher to introduce the Inq-ITS tutorial and Inq-ITS Ramp with Graphing Lab, as well as provide an in-depth overview of the Inq-Blotter dashboard system. None of the participating teachers had used the Inq-ITS or Inq-Blotter technologies in their classrooms before the start of the study. During the initial meeting,
teachers had an opportunity to ask questions about the technologies and the formal data collection sessions were scheduled (during regularly occurring class periods for each participating class section). All participating teachers assigned their students the Inq-ITS Tutorial to complete for homework prior to data collection. Data collection took place for all teachers and students between December 2020 - January 2021.

The principal investigator attended all synchronous class sessions via the virtual meeting platform selected by the teacher (i.e., Google Meets or Zoom) and recorded audio data for consenting participants through the computer-based QuickTime application. The teachers introduced the principal investigator to the students and the principal investigator explained the goal of the Ramp with Graphing Lab to students. Students were told that the teacher would be able to monitor their progress in real-time as they worked on the Ramp Lab. Students were then instructed to complete the multiple-choice pre-test. Students could not return to the pre-test after they submitted their answers. After finishing the pre-test, students worked on the Ramp with Graphing Lab for the remainder of the class period (i.e., ranging from 45-55 minutes). Ms. B and Mr. C sent students into breakout rooms after the initial classroom announcements were made (students were told they could return to the main session if they had any questions), but Ms. A had all students remain in the main session of the virtual meeting platform. The organization of the virtual classroom (staying in one session or using breakout rooms) was determined by the teacher based on the format typically used in their remote classroom.

The teachers then used the Inq-Blotter dashboard to monitor their students’ progress as the students worked through the three Ramp Lab activities. Teachers
responded to dashboard alerts through one of three methods: initiating an interaction with a student verbally through the meeting platform (by talking to the student in the main session, joining the breakout room of the student, or moving the student from their breakout room into the main session), sending a chat message to the student via the virtual meeting platform, or making a whole class announcement via the virtual meeting platform. All interactions and announcements were recorded and then anonymized by the principal investigator.

At the end of the class session, students were brought back to the main session for Ms. B’s and Mr. C’s classes. All teachers made announcements to students explaining that the Ramp Lab was to be completed for homework within 48 hours. After all students left the virtual meeting, the principal investigator debriefed with the teacher regarding their experience using the Inq-Blotter dashboard.

Measures

Pre-Test. Students’ responses to the 8 multiple-choice pre-test questions were automatically scored and stored in Inq-ITS with anonymous user identification numbers. The scoring was based on whether students selected the correct answer to the multiple-choice question (1 point) or not (0 points). Students’ average scores on the first four questions related to constructing graphs (i.e., numbers 1 - 4 in the Appendix) and four questions related to applying equations (i.e., numbers 5 - 8 in the Appendix) were used for analyses (i.e., average scores ranged from 0 - 1 point). For instance, a student who answered 5 of the 8 questions correctly had a final pre-test score of 0.63.

Inq-ITS Ramp Lab Log Data. Student performance on each of the first four inquiry practice stages within the Inq-ITS Ramp with Graphing lab was automatically
scoring (Gobert et al., 2012, 2013, 2018; Sao Pedro, 2018; Sao Pedro & Betts, 2019) and
scores were logged with the students’ anonymous identification numbers. For the
purposes of the present study, only performance on each mathematics practice stage was
used for analyses (e.g., Constructing Graphs and Applying Equations; Sao Pedro & Betts,
2019). These math practice stages were scored within Inq-ITS using knowledge-
engineered algorithms that captured student performance on multiple components of each
stage (i.e., sub-practices or knowledge components; Corbett & Anderson, 1995).

Table 2. Mathematics practices and sub-practices automatically scored in Inq-ITS

<table>
<thead>
<tr>
<th>Practice</th>
<th>Sub-Practices (Scored as 0 or 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>- Selecting the appropriate independent variable for the x-axis and dependent variable for the y-axis</td>
</tr>
<tr>
<td></td>
<td>- Aligning the axes to the variables in goal</td>
</tr>
<tr>
<td></td>
<td>- Selecting controlled trials to plot (i.e., trials where only one independent variable is manipulated; Chen &amp; Klahr, 1999)</td>
</tr>
<tr>
<td></td>
<td>- Selecting sufficient trials to plot (i.e., at least five trials to see a trend in the data)</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>- Selecting the appropriate mathematical relationship between variables</td>
</tr>
<tr>
<td></td>
<td>- Adjusting the values in the equation to get a best fit of 70% or higher</td>
</tr>
<tr>
<td></td>
<td>- Creating a line with appropriate relationship and fit of at least 70%</td>
</tr>
</tbody>
</table>

The sub-practices of constructing graphs and applying equations (see Table 2)
were previously operationalized for the Inq-ITS Newton’s Law of Gravitation Lab based
on the NGSS (2013) performance expectations, interviews with teachers and students,
surveys with teachers and students, and the common difficulties with each practice
identified in the literature (Sao Pedro, 2018; Sao Pedro & Betts, 2019). The sub-practices
of the mathematical stages were scored in a binary manner (1 point if proficiency was demonstrated; 0 points if proficiency was not demonstrated) to allow for capturing nuanced differences in whether students had mastered the particular sub-components involved in the practice (Corbett & Anderson, 1995). The students’ scores on their first attempts on an inquiry practice stage (the average of the sub-practices) prior to receiving help and after receiving help from a teacher in their next activity were used for analyses. The interpretation of scores on each mathematical practice stage had yet to be validated within the Ramp Lab.

**Inq-Blotter Log Data.** Teachers’ actions within Inq-Blotter were automatically logged, timestamped, and stored with an anonymous user ID (Gobert et al., 2016, 2018; Gobert & Sao Pedro, 2017). The logs included data such as the alerts that appeared in the dashboard, the alerts accessed by the teacher, and the contents of the alerts (i.e., anonymized student ID number, inquiry practice difficulty, specific sub-practice difficulty, Inq-ITS lab, lab activity goal, etc.). For the purposes of the present study, only alerts related to mathematical practices (i.e., Constructing Graphs and Applying Equations) were examined.

**Audio-Recordings.** The audio-recordings captured by the principal investigator using QuickTime were timestamped, transcribed, and anonymized (each recording was assigned a Recording ID number). The recordings included interactions between teachers and students, whole class announcements made by the teacher, and debriefings between the principal investigator and teacher.

To identify the types of teacher supports elicited by alerts, teacher-student interactions that occurred in response to an individual student alert were identified and
transcriptions were segmented into speaker turns (i.e., all statements spoken by the
teacher were identified within each conversation and conversations were divided into
spoken turns). All teacher turns were coded for six types of supports based on a
previously developed coding scheme that accounted for levels of teacher scaffolding
(Dickler et al., 2021a) as well as an additional support type that was added based on the
remote context (7 support codes total), including:

- **orienting** (i.e., directing the student to focus on the inquiry practice or a component of
  the practice of difficulty),

- **conceptual** (i.e., explaining key terminology and components of the inquiry practice
to the student),

- **procedural** (i.e., providing guidance on the steps involved in successfully engaging in
  an inquiry practice),

- **instrumental** (i.e., telling the student the exact actions they need to take to complete
  the practice successfully),

- **content** (i.e., discussing information related to the scientific phenomenon being
  investigated),

- **evaluative** (i.e., providing feedback on the accuracy of student work),

- **technical** (i.e., helping with online challenges including navigating the virtual
  meeting platform or screen-sharing).

A teacher turn could be coded as representing between 0 - 7 types of teacher
support. The principal investigator trained two Ph.D. students on the coding scheme. The
teacher turns were split in half and each Ph.D. student coded half of all the teacher turns
for the types of teacher supports. The principal investigator coded all teacher turns.
Interrater agreement was calculated between the principal investigator and each Ph.D. student. Raters agreed on 81% and 91% of codes respectively, disagreements were discussed, and agreed upon codes were used for analyses.

Table 3. Challenges experienced by teachers when supporting math practices remotely

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Need for more contextual information            | Teacher requested a visual of the students’ work in order to provide support | T: [S] can you share your screen real quick?  
S: I was doing it on my phone.  
T: Nooo, uhhhh…ok.  |
| Discomfort with mathematical practices          | Teacher expressed uncertainty with graphing or applying equations         | T: If you know the math then you can probably calculate the slope and all that, but I don't remember any of that so I just choose numbers to put in there.  |
| Alternative conceptions of mathematics in science | Teacher provided inaccurate information related to graphing or applying equations | T: …On the X [axis], that's your dependent. |

Additionally, a thematic analysis (Braun & Clarke, 2006) of all transcribed audio data was used to identify any unanticipated challenges that emerged in teachers’ use of the dashboard. Codes were assigned to any instances where teachers noted or experienced difficulties with dashboard use and supporting students on mathematical practices. These codes were then consolidated into three major themes:

- *need for more contextual information* (i.e., teachers experienced difficulties supporting students remotely without being able to see student work products),
• *discomfort with mathematical practices* (i.e., teachers expressed that they did not perceive themselves to be experts in mathematics or feel confident in applying the mathematics),

• *alternative conceptions of mathematics in science* (i.e., teachers guidance included ideas that were not mathematically or scientifically accurate; see Table 3 for examples).

**Analyses**

**Research Question 1.** In order to examine the validity of the interpretation of student performance on mathematical practices in Inq-ITS, student pre-test data and scores on the Inq-ITS Ramp with Graphing Lab stages were triangulated using anonymous student IDs. Pearson correlations were used to examine the relationship between students’ scores on the pre-test and overall score on the Inq-ITS Ramp with Graphing Lab math stages (i.e., average of the scores on the Constructing Graphs and Applying Equations stages). Separate correlations were also computed for each math practice, respectively: the correlation between student performance on the four Constructing Graphs pre-test questions and Constructing Graphs Inq-ITS stage, and the correlation between student performance on the four Applying Equations pre-test questions and Applying Equations Inq-ITS stage. The strength of correlations was determined using Evans’ (1996) classification of: \( r < 0.19 \) is a very weak correlation, \( 0.20 < r < 0.39 \) is a weak correlation, \( 0.40 < r < 0.59 \) is a moderate correlation, \( 0.60 < r < 0.79 \) is a strong correlation, and \( r > 0.80 \) is a very strong correlation.

**Research Question 2.** Analyses were also conducted to examine whether students improved and if their improvement was significant after being helped by a
teacher in response to an alert on a mathematical practice. The student pre-test data, performance data from Inq-ITS, log file data from Inq-Blotter, and transcribed audio data were triangulated using anonymous IDs and timestamps. These triangulated data were used to identify the students who were helped on an inquiry practice by a teacher using Inq-Blotter, the practices on which students were helped, and students’ performance on practices from prior to after being helped. Only students who were supported on the inquiry practice stages of constructing graphs and applying equations, and continued on to complete a second activity after being helped were included in the analyses ($n = 18$ students).

Additionally, a matched cohort of students who were not helped by the teacher ($n = 18$ students) was selected to compare the amount of student improvement on mathematical practices when helped by the teacher in response to an Inq-Blotter alert versus when students did not receive teacher help. These students were identified based on the following criteria: the student was not helped by the teacher based on an alert in Inq-Blotter, the student had the same teacher (as the corresponding student who was helped based on an alert), the student received the same score (or slightly higher if there was no exact match; $n = 4$ matched students with higher performance who did not receive any help) on the mathematical practice of difficulty in the same activity as the matched student who was helped based on an alert, and the student completed at least one additional activity. While students who were in the matched cohort had alerts triggered based on their performance on mathematical practices, they were not helped by teachers for several reasons including that teachers did not have time to respond to all alerts.
during the class period or students moved on to the next inquiry activity (and their alerts were removed from the dashboard) before teachers had an opportunity to respond.

For the students who were helped by a teacher based on an Inq-Blotter alert, descriptive statistics were conducted to identify the percentage of students who improved their performance on the mathematical practice on which they were helped from before receiving help to after receiving help (i.e., in the next activity they completed). Paired-samples t-tests were used with an alpha of .05 to examine whether this improvement was significant. A Mixed Model Analysis of Variance (MM ANOVA) was used with an alpha of .05 to compare performance from the first to second opportunity on the practice of difficulty between students who were helped to the matched students who were not helped.

**Research Question 3.** To better understand how teachers distributed types of supports to students in response to alerts, the relative frequencies of types of supports provided were examined (i.e., proportion of total interactions with each type of support). Descriptive analyses (i.e., a median split identifying students with low versus high pre-test performance) were used to examine whether there was a relationship between students’ prior knowledge as determined by their pre-test performance and the types of supports provided by teachers.

**Research Question 4.** Finally, the challenges experienced by teachers based on the thematic analysis were explored. In particular, the frequency of each challenge and additional contextual information was examined. These analyses were essential for informing future iterations of the content of alerts, as well as other interface features.

**Results**
Research Question 1 Results

To answer RQ1 (Is the interpretation of mathematical practice scores on the stages of Constructing Graphs and Applying Equations in Inq-ITS valid?), Pearson correlations were used to assess the construct validity of the mathematical practice assessment in Inq-ITS. A significant strong positive correlation (Evans, 1996) was found when examining the total pre-test performance (i.e., overall average score across the eight pre-test questions out of 1 point; $M = .74$, $SD = .23$) relative to Inq-ITS performance averaged across both math stages (i.e., average of score on Constructing graphs stage and score on Applying Equations stage out of 1 point; $M = .65$, $SD = .28$), $r = .64$, $p < .001$. This result provides evidence that students competencies as measured within the mathematical stages in Inq-ITS aligned with the measurement of mathematical practices as defined by other external assessments. Correlations were also explored for each mathematical practice stage.

Results indicated that there was a significant positive moderate correlation between student performance on Constructing Graphs on the pre-test (i.e., average score across the four Constructing Graphs pre-test questions out of 1 point; $M = .75$, $SD = .27$) and on Constructing Graphs in the Inq-ITS lab stage (i.e., average of Constructing Graphs sub-practice components out of 1 point; $M = .71$, $SD = .27$), $r = .50$, $p < .001$. There was also a significant moderate positive correlation between student performance on Applying Equations on the pre-test (i.e., average score across the four Applying Equations pre-test questions out of 1 point; $M = .73$, $SD = .27$) and on Applying Equations in the Inq-ITS lab stage (i.e., average of Applying Equation sub-practice components out of 1 point; $M = .59$, $SD = .39$), $r = .43$, $p < .001$. 
These results indicate that the assessment within Inq-ITS captured student competencies on constructing graphs and applying equations in science as defined by other common multiple-choice STEM assessments. Given that the construct validity of the Inq-ITS assessment for mathematical practices was supported by this analysis, conclusions could be drawn regarding changes in students’ competencies on Constructing Graphs and Applying Equations in relation to teacher support in response to Inq-Blotter.

**Research Question 2 Results**

For RQ2 (Are teacher discursive supports associated with student improvement on the practice on which they were helped in their next opportunity?) descriptive statistics were first used to identify the proportion of students who improved on the mathematical practice on which they were helped. The findings indicated that 89% of students (16 out of 18 students) improved on the math practice of difficulty after receiving teacher help based on Inq-Blotter (see Table 4). The results of the paired-samples t-tests revealed that students’ improvement was significant from prior ($M = .35, SD = .16$) to after being helped ($M = .83, SD = .26$) on the math practice of difficulty, $t(17) = -5.89, p < .001, d = 2.32$. In particular, students significantly improved from prior to after being helped on each math practice (see Table 5). These findings demonstrate that teacher support informed by Inq-Blotter alerts was associated with student improvement on challenging mathematical practices.

It is important, however, to examine how students performed on mathematical practices across activities if they did not receive help from the teacher in response to an Inq-Blotter alert. Specifically, these analyses will help to account for whether students improved on mathematical practices regardless of whether they received help from their
teacher (e.g., due to structural scaffolds in the Inq-ITS lab activities; Li et al., 2018). The following analyses were conducted to compare performance between students who did receive help and students who did not receive help from the teacher in response to an alert.

Table 4. Proportion of students with improvement on mathematical practices

<table>
<thead>
<tr>
<th>Inquiry Practice Alert</th>
<th># of students who improved</th>
<th>% of students who improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>7/8</td>
<td>88%</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>9/10</td>
<td>90%</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>16/18</strong></td>
<td><strong>89%</strong></td>
</tr>
</tbody>
</table>

Table 5. Paired samples t-tests for math performance prior to and after teacher help

<table>
<thead>
<tr>
<th>Inquiry Practice</th>
<th># students helped</th>
<th>Pre Help $M$ $(SD)$</th>
<th>After Help $M$ $(SD)$</th>
<th>Paired-Samples t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>8 students</td>
<td>.47 (.16)</td>
<td>.75 (.19)</td>
<td>$t(7) = -4.97, p = .002, d = 1.59$</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>10 students</td>
<td>.23 (.16)</td>
<td>.90 (.32)</td>
<td>$t(9) = -5.51, p &lt; .001, d = 2.65$</td>
</tr>
</tbody>
</table>

A MM ANOVA was used to compare the performance of students who were helped on a mathematical practice (in response to an Inq-Blotter alert; $n = 18$ students) relative to students who were not helped on mathematical practices ($n = 18$ matched students). Results revealed a significant within-subjects main effect indicating that students across conditions significantly improved from their first opportunity ($M = .37$, $SD = .22$) to their second opportunity on the practice ($M = .82$, $SD = .28$), $F(1, 34) =$
58.13, \( p < .001 \), \( n^2 = .63 \). While the interaction between students’ condition (i.e., help versus no help) was not significant, \( F(1, 34) = .69, p = .41 \), a trend could be seen where students who were helped by the teacher did surpass the performance of students who were not helped (see Figure 6). This is especially important to note as students who were \textit{not} helped started off with a higher average performance (\( M = .41, SD = .25 \)) because there were no exact matches to four students who were helped (\( M = .34, SD = .20 \)).

![Figure 6. Student performance on math practices across opportunities](image)

These initial comparisons indicate the potential of Inq-Blotter alerts to guide teacher support associated with student improvement on challenging inquiry practices remotely, which is commensurate with results from prior studies during in-person
learning for inquiry practices at the middle school level (Dickler et al., 2021a). Further analyses are needed to examine how particular structural scaffolds within the Inq-ITS learning environment (Li et al., 2018) may be contributing to student improvement regardless of teacher support (i.e., widgets to support automated plotting of graph points on the Constructing Graphs stage may support student learning; Mokros & Tinker, 1987). Additionally, it will be valuable to explore changes in student performance in each condition over longer periods of time (i.e., across additional activities and topics).

Other factors also may have impacted how teachers supported students and the corresponding impact of the scaffolds on student learning outcomes including students’ prior knowledge (Shute, 2008). Teachers may have made decisions about which students to help and how to help the students based on contextual information related to students’ prior knowledge outside of the Inq-ITS environment. While teachers were not explicitly interviewed in the present work about their decision-making process for which students to help, the discourse exchanged in response to alerts can provide insights into the types of supports teachers decided to provide to their students. As a result, it is valuable to explore teachers’ discursive supports elicited by alerts overall and in relation to students’ prior knowledge (as assessed by the multiple-choice pre-test) in order to begin to understand how other factors may interact with the use and impact of Inq-Blotter alerts.

**Research Question 3 Results**

Relative frequencies of the types of supports provided by teachers in response to an alert for a mathematical practice was computed to answer RQ3 (What is the relationship between the types of discursive supports teachers provided to students and students’ prior knowledge of using mathematics in science?). First, the proportion of
interactions in which teachers provided a particular type of support was examined across all interactions (out of \( n = 18 \) interactions that occurred in response to an alert for using mathematics; see Table 6). Teachers were found to primarily provide evaluative support (89% of interactions) and technical support (83% of interactions). This finding indicates that teachers were helping students to navigate the virtual meeting platform in response to alerts and commenting on the accuracy of student work. For example, in response to an alert for which a student was struggling with labeling the axes of their graph, Ms. B walked the student through the steps involved in sharing their screen [Technical Support] so that they could look at if the student had placed correct variables on each graph axis [Orienting Support; Recording ID A93]. In this example, the teacher needed to first provide support to the student to get them to share their screen before being able to provide more focused support to the student on the inquiry practice of difficulty.

Teachers also provided higher-level inquiry supports such as conceptual, procedural, and orienting comments across over 56% of interactions, respectively. These supports were essential for students to understand the components of the practice, the steps involved in the practice, and what they would need to address to be successful in future opportunities. For instance, Mr. C responded to the same alert as above for graphing (i.e., where a student was struggling with labeling the axes of their graph) and directed student attention to the graph [Orienting Support], explained the difference between an independent variable and a dependent variable [Conceptual Support], and noted that the student should put the dependent variable on the y-axis [Procedural Support; Recording ID A109]. These supports build on orienting supports by helping students to understand the sub-practices involved in the mathematical practice of
difficulty and preparing them to successfully complete the practice on their next opportunity.

While these results illuminate the most frequently provided supports across students helped based on an Inq-Blotter alert, it is also valuable to investigate whether teachers distributed supports to students differently in relation to their awareness of students’ prior knowledge (outside of the Inq-ITS environment). To further explore this potential relationship, students’ prior knowledge of Using Mathematics in science was determined based on their performance on the multiple-choice pre-test and examined in the following analyses.

Table 6. Proportion of interactions with each support type

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>61%</td>
<td>56%</td>
<td>56%</td>
<td>39%</td>
<td>89%</td>
<td>39%</td>
<td>83%</td>
</tr>
</tbody>
</table>


**Frequencies of Support by Prior Knowledge.** A median split was used to identify students who were helped with high versus low prior knowledge on the practices of using mathematics in science as assessed by the multiple-choice pre-test. There were six students with overall pre-test scores at the median (with a score of 75%; i.e., middle performance), six students with pre-test scores below the median (i.e., low performance/low prior knowledge), and six students with pre-test scores above the median (i.e., high performance/high prior knowledge). Students at the median were not included in the comparisons (see italicized row in Table 7). When comparing the proportion of support
types provided for students with high versus low prior knowledge (see Table 7), there appeared to be no difference in the occurrence of orienting, conceptual, procedural, and evaluative supports in each group. There were, however, differences between supports provided to each group in relation to: instrumental, content, and technical supports, as detailed below.

Table 7. Proportion of interactions with each support type in relation to student prior knowledge

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>67%</td>
<td>67%</td>
<td>67%</td>
<td>33%</td>
<td>83%</td>
<td>33%</td>
<td>100%</td>
</tr>
<tr>
<td>Middle</td>
<td>50%</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
<td>100%</td>
<td>33%</td>
<td>100%</td>
</tr>
<tr>
<td>High</td>
<td>67%</td>
<td>67%</td>
<td>67%</td>
<td>50%</td>
<td>83%</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>


Students with higher prior knowledge were more likely to receive *instrumental* and *content* support relative to students with lower prior knowledge. This finding suggests that the supports provided to students with higher prior knowledge were focused more on helping students to move past the math practice stage of difficulty (i.e., telling them the exact steps to take to progress) or providing information related to the scientific phenomenon rather than giving information on the inquiry practice (i.e., explaining a science concept specific to the lab activity). For example, Ms. B received an alert that a student (pre-test score of .88 points out of 1 point) was struggling with Constructing Graphs because they had put incorrect variables on their axes. Ms. B asked the student
about the scientific phenomenon they were examining (i.e., how the height of the tower impacted the time it took the sled to reach the end of the ramp) and told them that the “the height of the tower should be on the x-axis” [Recording ID A92]. Ms. B essentially told the student exactly how to set up the axes of their graph for the specific scientific investigation in order to move forward to the next mathematical practice stage (i.e., Applying Equations). Without interview data, the exact reasoning behind these supports cannot be determined but perhaps Ms. B assumed this student with higher prior knowledge had the mathematical practice strategies to be successful in their next opportunity and made the decision that the student just needed an additional direct support in the moment (Shute, 2008).

Additionally, a chi-square analysis revealed a significant relationship between prior knowledge and relative frequency of technical support where students with lower prior knowledge received technical support in all interactions, \( x^2(2, N = 12) = 7.20, p = .027 \). This result indicates that when providing support to students with low prior knowledge in response to alerts, teachers provided substantial support related to navigating the online meeting platform (e.g., screen sharing). For instance, in one example Ms. A received an alert that a student (pre-test score of .63 points out of 1 point) was struggling with Constructing Graphs because the student was plotting uncontrolled trials. The first comment Ms. A made to the student was “[S] can you share your screen real quick?” [Recording ID A19]. Ms. A may have felt the need to see the student’s work and walk them through the practice of Constructing Graphs in greater detail because the student previously had difficulty with this practice, but further interviews are needed to fully understand how supports were distributed to specific students.
These exploratory findings suggest that in addition to the contextual information within dashboard alerts, teachers may have applied their own assessment and background knowledge of students’ understandings to inform the types of supports they provided to each student. Studies have shown the value in distributing feedback based on students’ prior knowledge and specific needs to best promote learning (e.g., Shute, 2008). A greater number of teachers and students, as well as interviews with teachers to inquire about their decision-making processes when providing support would be needed to further explore this relationship. The discursive data from the present study, however, is sufficient to begin exploring other factors related to teachers’ use of the dashboard in a remote setting including challenges experienced by teachers as explored in the final research question.

**Research Question 4 Results**

For the final research question, RQ4 (What unanticipated themes emerged in educators’ experiences with using the Inq-Blotter dashboard remotely to support students on mathematical practices?), the frequency of themes applied to the transcribed audio data reflecting the challenges experienced by educators were explored (see Table 8). For the first challenge regarding “need for more contextual information,” it was found that teachers needed additional details to provide support to students during a third of interactions (i.e., teacher-student conversations that occurred when the teacher responded to an alert). Teachers also commented (i.e., made statements outside of interactions) that they wished they had more information to be able to guide students without having to ask for students to share their screens. Screensharing was often time consuming and required substantial technical support as demonstrated by the frequency of technical support types
in the prior analyses. This challenge was specific to learning in a remote setting as teachers would be able to look over a student’s shoulder during in-person learning and access student materials directly.

Table 8. Frequency of teacher challenges with remote dashboard use

<table>
<thead>
<tr>
<th>Challenges</th>
<th># of Interactions</th>
<th># of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for more contextual information</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Discomfort with mathematical practices</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Alternative conceptions of math in science</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Another frequent challenge that emerged in almost a quarter of interactions was teachers’ “discomfort with mathematical practices” in science. In particular, teachers expressed that they did not remember the mathematics and made comments about their own challenges with mathematics. For instance, one teacher expressed that “the math part was kind of hard” for her because she “forgot all of the graphing math already” [Ms. B; Debriefing]. Using mathematics in science is a particularly difficult practice for teachers and students (Dickler et al., 2021b), so it is essential to consider how future alert designs might support teachers in this area.

Finally, a quarter of interactions reflected “alternative conceptions of mathematics in science” where teachers provided support that would not result in successful completion of the inquiry practice. Teachers also made comments demonstrating a misunderstanding of particular mathematical concepts (i.e., the role of the coefficient versus the constant in an equation). This theme indicated the need for additional guidance with accurate mathematical/scientific information on these challenging inquiry practices.
Discussion

Assessing and supporting the inquiry practices as outlined in policy documents such as the NGSS (2013) has introduced a number of challenges for teachers. There is a lack of tools that address the full range of practices (Pruitt, 2014), and essential practices such as Using Mathematics are often excluded from assessments. The need for materials to guide the assessment and scaffolding of these practices has been greatly amplified by remote learning due to the COVID-19 pandemic, which further limited opportunities to measure and support student competencies. The present study addressed these central challenges through a demonstration of technologies that guided teacher support on critical inquiry practices within a remote setting. The student environment, Inq-ITS, allowed for fine-grained assessment of students’ inquiry competencies, including valid assessment of mathematical practice competencies. The corresponding real-time alerting dashboard, Inq-Blotter, elicited teacher support associated with student improvement on inquiry practices.

The assessment of mathematical practices in science (i.e., Constructing Graphs and Applying Equations) in Inq-ITS was further validated based on the results of correlations between this assessment and an assessment constructed of multiple-choice items that capture mathematics in science competencies. This work builds on prior studies conducted to validate the assessment within the Newton’s Law of Gravitation Lab stages during in-person instruction (e.g., Sao Pedro & Betts, 2019). The findings of this study in combination with prior studies support the extrapolation that student performance on the stages of Constructing Graphs and Applying Equations in Inq-ITS reflects their competencies on these practices, and correspondingly changes in
performance on these stages reflects changes in their competencies. As a result, this work provides an authentic assessment operationalizing the core inquiry practice of Using Mathematics that can be applied across instructional contexts.

The present study also demonstrated how the use of the Inq-Blotter alerting dashboard was associated with student improvement on challenging mathematical practices. In particular, the majority of students who were helped by the teacher in response to a dashboard alert significantly improved on the mathematical practice on which they were helped in their next opportunity to use the same mathematical practice. This finding has significant implications for how teacher dashboards can support instruction of inquiry practices (when paired with student environments containing fine-grained, valid assessments of inquiry competencies). Additionally, this study provides evidence of how dashboards can be implemented remotely to enable teachers to continue to monitor and support students’ difficulties in real-time.

The present work also provided evidence of the types of teacher supports elicited by remote dashboard use and how teachers may decide to provide supports to students in relation to contextual factors beyond the dashboard alert. Teachers provided additional lower-level inquiry supports to students with higher prior knowledge, potentially because students could benefit from activity-specific supports. Teachers also provided more technical support to students with lower prior knowledge, reflecting the need to help these students navigate the technological environment to support inquiry challenges. Across each of these circumstances, teachers were able to initiate interactions and support students as a result of Inq-Blotter alerts. The dashboard tool provided an outlet for
teachers to support students in real-time remotely while also potentially building on their prior assessments of students’ understandings.

The final analyses in the present study revealed that there were some challenges experienced by teachers when using the dashboard to support students on using mathematics remotely. Teachers found that additional contextual information was needed to support students remotely because they could not access student work directly, teachers noted their own discomfort with challenging mathematical practices, and teachers expressed alternative conceptions of components involved in using mathematics in science. This work demonstrates the need for conducting implementation studies across contexts (i.e., synchronous remote instruction, in-person instruction) to fully understand how these technologies may function differently based on the circumstances and, correspondingly, inform future design iterations that will meet varying needs.

In terms of future work on Inq-Blotter, it will be essential to consider how alerts could be designed to address teacher challenges when using the dashboard remotely (and test the impact of the updated alerts; see Study 2 and Study 3). Instructional supports have been shown to be valuable in helping teachers provide inquiry guidance to students within professional development contexts (Morris & Chi, 2020; Oliveira, 2009). Future studies should investigate the impact of embedding instructional supports to guide teacher scaffolding on mathematical practices when limited contextual information is accessible to teachers in a remote setting. Studies could then explore whether there is a significant difference in student learning outcomes (see Study 2) and how teachers support students (see Study 3) when the embedded instructional supports are available to teachers within the alerts.
Additionally, it will be important to explore the use of these technologies with additional teachers and students to continue to capture any other differences in how the technologies are implemented over longer time periods. With additional students, further comparisons could be made regarding the relationship between teacher support and student improvement when helped by the teacher. Future studies might also explicitly ask teachers to reflect on their decisions to help particular students. It would also be beneficial to explore the effects of the technological interventions over longer periods of time and to see whether student learning transfers across different topics.

Overall, this work demonstrates how a real-time alerting dashboard and corresponding intelligent tutoring system for science inquiry support critical practices in remote settings. This work sets an example for both the processes involved in developing and validating the assessment of science inquiry competencies as well as testing the implementation of technologies in synchronous remote settings. The features of the tools involved in this work have the potential to alleviate many of the challenges experienced by educators in the current educational climate. It is important to remain flexible to educators’ needs and continuously iterate on technological designs to meet those needs.
Phase 2: Study 2

Abstract

Dashboard technologies have been developed to support teachers in monitoring their students’ progress within online environments for Science, Technology, Engineering, and Mathematics (STEM). It is important to consider how features of these tools can be developed to elicit guidance from teachers that supports student learning, even within remote settings. The dashboard, Inq-Blotter, was developed to support teacher guidance in science inquiry settings and has been associated with student improvement on inquiry practices in the intelligent tutoring system, Inq-ITS. Initial studies indicate, however, that teachers provided primarily lower level supports in response to dashboard alerts and that additional instructional supports were necessary to guide teachers, particularly on the practices involved in Using Mathematics in science. This study presents an overview of the development of Teacher Inquiry Practice Supports (TIPS) for mathematical practices that were embedded within the Inq-Blotter dashboard alerts and tested within remote classroom settings. Specifically, the present study quantitatively compared outcomes associated with the use of basic Inq-Blotter alerts (without TIPS) and Inq-Blotter alerts with TIPS. The results of this work have implications for how the design of alerts in teacher dashboards can impact student learning outcomes on complex inquiry practices.

Introduction to Phase 2 Study 2

STEM teacher dashboards are technologies that were designed to support orchestration of classroom activities (Dillenbourg, 2013). Dashboards can guide teachers in organizing whole class activities as well as in providing individualized support to
students based on their performance within online environments (Verbert et al., 2014). In terms of features that guide individualized support to students, dashboards for STEM often include visualizations of the number of activities students completed (Acosta & Slotta, 2018; Molenaar & Knoop-van Campen, 2018; van Leeuwen & Rummel, 2020), data on the accuracy of student work (Holstein et al., 2019; Martinez-Maldonado et al., 2015), and alerts on students’ progress (Schwarz et al., 2018; Tissenbaum & Slotta, 2019; VanLehn et al., 2019). These features enable an Open Learner Model (OLM; Bull, 2020; Bull & Kay, 2016) by making the results from assessment within the student environment visible to the teacher in a digestible format (Charleer et al., 2014). Teachers can then use the information from the features to make decisions regarding how to scaffold individual students in real-time (Knoop-van Campen & Molenaar, 2020). Evaluation studies, however, are necessary to understand the effectiveness of the use of these features in promoting student learning (Holstein et al., 2018a; Laurillard, 2008).

**Dashboard Evaluations**

Dashboard testing initially takes place in a laboratory setting as features undergo an iterative design process (e.g., Holstein et al., 2019). In particular, dashboards may be evaluated through asking teachers to review initial designs (Matuk et al., 2016; Molenaar & Knoop-van Campen, 2018; van Leeuwen & Rummel, 2018) as well as having teachers implement the tool using simulated data (Holstein et al., 2019; Lajoie et al., 2020; van Leeuwen & Rummel, 2020). These studies are valuable because they can help to elucidate effective features (e.g., real-time alerts to direct teacher attention; van Leeuwen & Rummel, 2020), but it is still necessary to examine the use of the tools in real
classrooms to fully understand the impact of feature use on student learning outcomes (since that is one of the primary goals of dashboard use).

It is also important to explore the use of dashboard systems within actual classroom settings to ensure the ecological validity of the technologies (Laurillard, 2008). Implementation studies often examine how the dashboard is used by the teacher within the classroom setting through examinations of log files and videos (Holstein et al., 2018b; Lajoie et al., 2020; Martinez-Maldonado et al., 2013; Molenaar & Knoop-van Campen, 2017; Tissenbaum & Slotta, 2019) as well as the corresponding impact on student learning outcomes (Holstein et al., 2018a; Martinez-Maldonado et al., 2015). Classroom settings, however, can vary extensively so it is necessary to explore the use of dashboards across different contexts (e.g., in-person classrooms versus remote classrooms), especially given the emergency shift to remote learning that occurred in response to the COVID-19 pandemic (Arnett, 2021; IES, 2021; UNESCO, 2020).

**Remote Dashboard Use**

While many studies have tested the use of dashboards within in-person classrooms (Holstein et al., 2018b, 2019a; Lajoie et al., 2020; Martinez-Maldonado et al., 2013, 2015; Molenaar & Knoop-van Campen, 2017; Tissenbaum & Slotta, 2019), few studies have explored the use of these tools within remote online settings (Lajoie et al., 2020; Verbert et al., 2014; see Study 1). Even amongst dashboards tested in remote settings, most studies were conducted asynchronously (e.g., Lajoie et al., 2020; Verbert et al., 2014) rather than synchronously (see Study 1).

Synchronous remote settings amplify and introduce new challenges for educators including finding ways to monitor and support student progress in real-time without
access to in-person materials (Archambault, 2010; Arnett, 2021; Broadbent, 2017; Kebritchi et al., 2017; Marshall et al., 2020; Means et al., 2021). It is important to investigate whether dashboard design changes may be needed to best support teachers and students in remote contexts, as was the case for the Inq-Blotter dashboard when initially tested in online synchronous high school classrooms (see Study 1).

**Prior work Inq-Blotter and Inq-ITS**

Inq-Blotter is a teacher dashboard that was initially developed to accompany a middle school science Inquiry Intelligent Tutoring System, Inq-ITS (Gobert et al., 2018) in order to support teachers’ pedagogical practices for the NGSS (2013). Inq-ITS was recently extended for high school and includes new virtual lab stages that assess students’ use of mathematics in science inquiry (i.e., Constructing Graphs and Applying Equations). The-Blotter dashboard was also extended for high school to include alerts that provided information to teachers when students were struggling with a particular component (or sub-practice) of the new mathematical stages (Sao Pedro & Betts, 2019). A recent study examining the use of Inq-Blotter with the alerts for mathematical stages during synchronous remote instruction in high school STEM classrooms revealed that students significantly improved on the mathematical inquiry practice on which they were helped (see Study 1). Additional qualitative analyses, however, indicated that teachers in remote contexts provided primarily lower-level inquiry supports to students (e.g., orienting supports to direct student attention, technical supports to guide use of online tools) and that remote teachers experienced challenges supporting difficult mathematical practices (e.g., lacking contextual information, discomfort with mathematical practices, and alternative conceptions of mathematics). As a result, it is essential to consider how
the design of Inq-Blotter alerts can be updated to best support teachers in providing high-
levels of support associated with student learning gains (Dickler et al., 2021) as well as to
give teachers enough contextual information to address challenges with supporting
complex inquiry practices remotely.

Researchers recently developed and integrated Teacher Inquiry Practice Supports
(TIPS) into Inq-Blotter alerts for middle school level inquiry practices including asking
questions/hypothesizing, carrying out investigations, and analyzing and interpreting data
(Adair et al., 2020). These TIPS were designed based on an assessment of teacher needs
and examples of higher-level teacher supports associated with student improvement at the
middle school level in prior studies (Dickler et al., 2019, 2021a). The instructional
supports within TIPS were intended to guide different types/levels of teacher scaffolding
to students based on the specific inquiry practice difficulty students were experiencing.
Teachers could choose to access the TIPS (embedded within the Inq-Blotter alert) and
either provide verbal feedback to students based on the content within the TIPS or read
the content of the TIPS directly to the student. Studies, however, had yet to investigate
the effects of these TIPS on teachers’ use of the dashboard and corresponding student
performance. Additionally, TIPS were not developed for high school level inquiry
practices involved in using mathematics in science.

The present study outlines the development of TIPS for the practices involved in
using mathematics (i.e., Constructing Graphs and Applying Equations) based on data
from a prior study (see Study 1). An experimental study was then conducted to examine
the effectiveness of TIPS for guiding teacher support and promoting student
improvement on inquiry practices.
Development of TIPS for Using Mathematics

TIPS are instructional supports that are embedded within Inq-Blotter alerts for student difficulties on the inquiry practice stages in Inq-ITS. The TIPS for middle school inquiry practices included four levels of increasingly targeted information that teachers could access (i.e., orienting supports, conceptual supports, procedural supports, and instrumental supports). These stages were previously determined based on a taxonomy for inquiry scaffolding needed to successfully carry out an inquiry practice (Gobert et al., 2013), and the same format was adopted for the TIPS for Constructing Graphs and Applying Equations. There are five alerts that teachers can receive in Inq-Blotter regarding student difficulties on the inquiry practice stage of Constructing Graphs and three alerts regarding student difficulties on the practice stage of Applying Equations (see Table 9). Therefore, TIPS were needed for each of these 8 alerts. The corresponding development of the content of the new TIPS took place over three stages.

Table 9. Inq-Blotter alerts for mathematical practice stages

<table>
<thead>
<tr>
<th>Math Practice</th>
<th>Brief Description of Possible Alerts Triggered by Student Difficulties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>- The variables on each axis do not align with the investigation</td>
</tr>
<tr>
<td></td>
<td>- The x-axis was labeled incorrectly</td>
</tr>
<tr>
<td></td>
<td>- Uncontrolled trials were plotted on the graph</td>
</tr>
<tr>
<td></td>
<td>- Graph axes were labeled incorrectly and plotted uncontrolled trials</td>
</tr>
<tr>
<td></td>
<td>- Insufficient data was plotted to see a trend in the graph</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>- Incorrect mathematical/functional relationship selected for the graph</td>
</tr>
<tr>
<td></td>
<td>- Line does not fit the data</td>
</tr>
<tr>
<td></td>
<td>- Incorrect mathematical relationship and line does not fit the data</td>
</tr>
</tbody>
</table>
In the first stage, the content of the TIPS was written for each alert by the principal investigator based on the supports provided by teachers on the mathematical practices in a prior study (see Study 1). In particular, the investigator identified all supports provided by teachers when students improved on the mathematical practice after receiving teacher help and organized supports according to the four levels of TIPS. The content of teacher support was then consolidated and reworded for clarity to be used in the TIPS text. The TIPS were drafted (4 levels of TIPS for each of the 8 alerts) and organized within a table (see examples in Table 10). A slide deck was also constructed with mock-ups of the alerts with TIPS and examples of student difficulties that would trigger the alert.

<table>
<thead>
<tr>
<th>Inq-Blotter Alert</th>
<th>Orienting TIPS</th>
<th>Conceptual TIPS</th>
<th>Procedural TIPS</th>
<th>Instrumental TIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>&lt;student name&gt;</em> is struggling to create a mathematical model because they do not understand the shape the data make.</td>
<td>Let’s look at your model. What is the trend in the data? What is the relationship between the variables in your graph?</td>
<td>Remember, the goal here is to choose a relationship (or mathematical equation) that represents the trend in the data. You can click on the &quot;Relationship&quot; dropdown menu to see what each mathematical relationship looks like.</td>
<td>Look at the points on your graph, think about the trend in the data, and choose the mathematical relationship. Also make sure your graph axes are labeled correctly and that you have at least 5 data points from controlled trials to better see the trend.</td>
<td>Try selecting different relationships until your line (or curve) follows the entire trend in your data points. You may need to go back to adjust the axes of your graph and add more data points to better see the trend.</td>
</tr>
<tr>
<td>&lt;student name&gt; is struggling to create a mathematical model. They have a model with the correct function, but none that have a fit of at least 70%.</td>
<td>Let's look at the models you created. How well does the line (or curve) in your model fit the data?</td>
<td>Remember, you want your line (or curve) to be as close to as many data points on your graph as possible to get the best fit. Changing the coefficient (a) will change the steepness or direction of the line (or stretch/shrink the curve). Changing the constant (b) will shift the line (or curve) up/down.</td>
<td>Think about how to change the coefficient to adjust the steepness or direction of the line (or stretch/shrink the curve) and the constant to shift the line (or curve) up/down.</td>
<td>Try out different values for both your coefficient and constant to get the line (or curve) as close to the data points as possible until you get a fit of at least 70%.</td>
</tr>
<tr>
<td>&lt;student name&gt; is struggling to create a mathematical model. They made models with the correct function and a fit of at least 70%, but no single model with both.</td>
<td>Let's look at a model you created that has a fit of over 70% to the data. Is this model representative of all of the data points in your graph? How can we improve this model?</td>
<td>Remember, it is possible for a representation to be very close to some data points, but the goal is to choose a relationship (or mathematical equation) that represents the entire trend in the data points. You can click on the &quot;Relationship&quot; dropdown menu to see what each mathematical relationship looks like.</td>
<td>Think about the model you created that had a high fit to the data (more than 70%). Is there another relationship you can select that might have a better fit to the data? Then think about how to change the coefficient to adjust the steepness or direction of the line (or stretch/shrink the curve) and the constant to shift the line (or curve) up/down.</td>
<td>Try selecting different relationships until your line (or curve) follows the entire trend in your data points. Try out different values for both your coefficient and constant to get the line (or curve) as close to the data points as possible until you get a fit of at least 70%.</td>
</tr>
</tbody>
</table>
In the second stage of TIPS development, the TIPS table and slide deck were reviewed by three high school and undergraduate STEM educators. All educators were sent the materials to review prior to a synchronous virtual meeting. The three educators and principal investigator then met virtually as a group for two hours to discuss the content, length, and presentation of the TIPS. The educators provided various suggestions including rewording instructional supports for mathematical clarity and brevity. The principal investigator made changes to the supports during the virtual meeting.

In the third and final stage of development, the TIPS were embedded into the Inq-Blotter system by two software engineers on the Inq-ITS team. After the principal investigator bug tested the updated alerts with TIPS for Using Mathematics as part of the present dissertation, this study (Study 2) was conducted to examine the use of the new instructional supports during synchronous remote instruction, as well as compare the use of Inq-Blotter alerts with TIPS to Inq-Blotter alerts without TIPS (i.e., basic alerts; Study 1). This work was essential for understanding how these instructional supports could guide teacher scaffolding remotely and if the supports could alleviate some of the challenges experienced by teachers when scaffolding mathematical practices in science remotely.

In the present study, the following research questions were examined regarding the remote use of Inq-Blotter alerts with TIPS for using mathematics (relative to basic Inq-Blotter alerts without TIPS):

1) Are teacher supports elicited by Inq-Blotter alerts with TIPS for using mathematics practices (i.e., Constructing Graphs and Applying Equations)
associated with student improvement on the practice on which they were helped in their next opportunity?

2) Are there differences in student performance on mathematical practices across Inq-ITS activities when teachers use Inq-Blotter alerts versus Inq-Blotter alerts with TIPS?

3) Are there any differences in how teachers use Inq-Blotter alerts versus Inq-Blotter alerts with TIPS?

Methods

Participants

The participants in the present study included 6 STEM teachers and their students from three high schools in the northeastern United States (N = 225 total students; see Table 11). One teacher from each high school used basic Inq-Blotter alerts (i.e., Inq-Blotter; n = 3 teachers and 119 students) and the other teacher from each high school used Inq-Blotter alerts with TIPS (i.e., Inq-Blotter with TIPS; n = 3 teachers and 106 students; see Table 12). All participation took place remotely between December 2020-February 2021.

Table 11. Demographics of participating schools

<table>
<thead>
<tr>
<th>HS</th>
<th>Ss</th>
<th>FRL</th>
<th>Asian</th>
<th>Black</th>
<th>Hisp.</th>
<th>Nat. Amer.</th>
<th>Pac. Isl.</th>
<th>White</th>
<th>Two or More Races</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36</td>
<td>73%</td>
<td>0.4%</td>
<td>45.8%</td>
<td>52.7%</td>
<td>0.1%</td>
<td>0%</td>
<td>0.9%</td>
<td>0.2%</td>
</tr>
<tr>
<td>2</td>
<td>115</td>
<td>32%</td>
<td>22.5%</td>
<td>11%</td>
<td>21.6%</td>
<td>0.2%</td>
<td>0.2%</td>
<td>39.6%</td>
<td>4.9%</td>
</tr>
<tr>
<td>3</td>
<td>74</td>
<td>8%</td>
<td>8.1%</td>
<td>2.1%</td>
<td>10.8%</td>
<td>0%</td>
<td>0.2%</td>
<td>78.6%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

Note. HS = High School, Ss = number of students, FRL = Free and Reduced Lunch, Hisp. = Hispanic, Nat. Amer. = Native American, Pac. Isl. = Pacific Islander
Table 12. Participant conditions and details

<table>
<thead>
<tr>
<th>High School</th>
<th>Teacher</th>
<th>Condition</th>
<th>STEM Subject</th>
<th># of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ms. A</td>
<td>Alerts</td>
<td>Engineering</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Ms. X</td>
<td>Alerts with TIPS</td>
<td>Physics</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Ms. B</td>
<td>Alerts</td>
<td>Chemistry</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Ms. Y</td>
<td>Alerts with TIPS</td>
<td>Biology</td>
<td>73</td>
</tr>
<tr>
<td>3</td>
<td>Mr. C</td>
<td>Alerts</td>
<td>Physics</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Ms. Z</td>
<td>Alerts with TIPS</td>
<td>Biology</td>
<td>26</td>
</tr>
</tbody>
</table>

Materials

Inq-ITS Tutorial Lab. All teachers assigned their students the Inq-ITS tutorial lab to complete for homework prior to the synchronous data collection session. The Inq-ITS tutorial lab was an investigation of whether adding red dye to a glass of water with a flower effected the redness of the flower petals. Students progressed through stages where they hypothesized about the impact of the red dye on the petal redness (i.e., Asking Questions/Hypothesis stage), used a simulation to test their hypothesis (i.e., Carrying out Investigations stage), made a claim based on the findings from their investigation with the simulation using widgets and selected the evidence that supported their claim (i.e., Interpreting Data stage), and communicated their findings in writing in the claim, evidence, and reasoning format (i.e., Explaining Findings stage; see Figure 7).

The four stages and corresponding widgets/tools that appeared in the tutorial were common across all Inq-ITS lab investigations. Therefore, the tutorial provided students with an opportunity to familiarize themselves with the Inq-ITS lab structure.
Inq-ITS Ramp with Graphing Lab. All participating students completed the Inq-ITS Ramp with Graphing Lab (i.e., Ramp Lab) during data collection. The goal of this lab was for students to practice using mathematics within a science context, a key NGSS (2013) practice that is often overlooked. Students investigated the mathematical relationships between different independent variables (i.e., height of a tower and corresponding steepness of a ramp, mass of sled, roughness of a ramp and corresponding friction) and dependent variables (i.e., the time it takes a sled to reach the end of a ramp, momentum of a sled at the end of the ramp, and acceleration of the sled at the end of the ramp). Students completed three investigation activities to explore these relationships and each investigation activity included six stages.

Four of the inquiry stages in the Ramp Lab were associated with the tutorial activity (i.e., Asking Questions/Hypothesizing, Carrying out Investigations, Interpreting...
Data, and Explaining Findings; within the Asking Questions/Hypothesis stage only, students could not move forward without developing a question that aligned with the investigation goal and received automated guidance on the variables to select. The other two stages that students completed occurred after the Carrying out Investigations stage:

- Constructing Graphs (i.e., students selected data that they collected to plot on a graph and determined the variables to use for each axis of the graph; see Figure 8).
- Applying Equations (i.e., students determined the mathematical relationship between the variables in their graph and adjusted the values within the mathematical equation to create a best fit line for their graph; see Figure 9).

Once students completed these stages, they moved on to further interpret their data and explain their findings in writing.

---

**Figure 8. Constructing Graphs stage of the Inq-ITS lab**
**Inq-Blotter.** Three participating teachers (one teacher from each high school) used Inq-Blotter while their students completed the Inq-ITS Ramp Lab. The Inq-Blotter dashboard provided real-time alerts to teachers on students’ performance on inquiry practice stages including Constructing Graphs and Applying Equations. The alerts for individual student difficulties on these stages were triggered by the automated scoring in Inq-ITS (see Measures section for more details). The individual student alerts appeared in the center panel of the dashboard with information on the specific inquiry practice stage that students were having difficulty with (see Figure 10). Teachers could select an alert to see additional information on the specific part of the practice the student was struggling with (i.e., the sub-practice of difficulty) as well as additional information in the right
panel including: the lab activity the student was working on, other alerts that had appeared for the student that day, and the student’s performance on other practices (see Figure 10).

![Figure 10. Inq-Blotter Teacher Dashboard (without TIPS)](image)

Teachers also received individual alerts for “slow progress” when a student was on a stage for longer than five minutes and alerts at the whole class level when more than 50% of students were struggling with a practice stage. Teachers could choose whether they preferred alerts to be organized by the most recent alerts appearing first or by inquiry practice stage.

**Inq-Blotter with TIPS.** The other three teachers used Inq-Blotter with TIPS while their students completed the Ramp Lab. Inq-Blotter alerts with TIPS included all of the same information as regular Inq-Blotter, plus embedded instructional supports when teachers selected an individual student alert for an inquiry practice difficulty.
Specifically, teachers would see an option in the right panel of the dashboard that said “Press for TIPS (Teacher Inquiry Practice Supports)” (see Figure 11). If teachers selected this option, then they were able to see a menu where they could navigate four levels of instructional support to teachers to inform scaffolding of student difficulties.

The first level of guidance available was the Orienting Support (see Figure 12), which was intended to direct the student attention to the component of the inquiry practice stage that required revisions. The next level of guidance was a Conceptual Support (see Figure 13), which included an explanation of the purpose of the inquiry practice stage and any relevant terminology. The Procedural support (see Figure 14) gave details on the general steps involved in completing the inquiry practice stage successfully. Instrumental support (see Figure 15) was the final level of support and involved instructions on the exact steps needed to successfully move forward in the specific lab activity. Teachers could click through these TIPS in any order (i.e., a teacher could read the Orienting Support and then click on the Procedural support).

Figure 11. Inq-Blotter with TIPS initial screen
Figure 12. Inq-Blotter TIPS Orienting Support

Figure 13. Inq-Blotter TIPS Conceptual Support
Figure 14. Inq-Blotter TIPS Procedural Support

Figure 15. Inq-Blotter TIPS Instrumental Support
**Procedure**

One teacher from each school was assigned to use either Inq-Blotter with basic alerts or Inq-Blotter with TIPS. The principal investigator met with each teacher individually to provide an overview of the Inq-ITS system and labs. The principal investigator then gave a walkthrough tutorial of the Inq-Blotter dashboard (with basic alerts or alerts with TIPS depending on the teacher’s condition). All teachers assigned the Inq-ITS tutorial lab for students to complete for homework on their own device (all participating students had access to either a personal or school provided device). After answering any questions, a date for remote synchronous data collection was scheduled during the teachers’ regular STEM class periods (Inq-Blotter was used between December 2020 - January 2021; Inq-Blotter with TIPS was used between January 2021 - February 2021).

Prior to data collection, all teachers assigned the Inq-ITS tutorial lab for students to complete for homework on their own device (all participating students had access to either a personal or school provided device). On the day of data collection, the principal investigator joined the virtual class meeting (via Zoom or Google Meets), introduced herself to the students, briefly explained the Ramp Lab, and audio-recorded interactions for consenting participants. Students then worked on the Ramp Lab for the remainder of the class period (45-55 minutes; students in the Inq-Blotter with basic alerts condition also completed a brief multiple-choice pre-test prior to starting the lab, but this pre-test was not examined in the present study). As students worked, the teachers monitored their performance using the Inq-Blotter dashboard (either with or without TIPS depending on their condition). Teachers provided support to students based on the dashboard alerts. At
the end of the class period, students were instructed to finish the rest of the Ramp Lab for homework if they had not completed all three investigation activities.

Teachers had continued access to the Inq-ITS and Inq-Blotter systems during their regular classes following the conclusion of the study.

**Measures**

**Inq-ITS Logs.** In the present study, students’ performance on the mathematical practice stages in Inq-ITS (i.e., Constructing Graphs and Applying Equations) were examined in relation to the use of Inq-Blotter with or without TIPS in their classroom. Students’ scores on the inquiry practice stages of Asking Questions/Hypothesizing (i.e., selecting an independent variable and dependent variable to investigate based on the activity goal) and Carrying out Investigations (i.e., running at least five controlled trials investigating the targeted scientific phenomenon in the investigation using the simulation; Gobert et al., 2013; Sao Pedro et al., 2013) were only included in analyses for baseline comparisons. Students’ performance on each mathematical practice was automatically scored and logged in Inq-ITS based on previously developed and validated knowledge-engineered algorithms (see Study 1; Sao Pedro & Betts, 2019). Each practice was scored based on student performance on the components of the inquiry practice stage.

The sub-practice components (each scored as 1 point if the student successfully completed the component or else 0 points if the student was unsuccessful) involved in Constructing Graphs included:

- selecting controlled trials to plot on the graph (i.e., where only one independent variable was changed),
- selecting at least five trials to plot on the graph to see the trend in the data,
selecting appropriate variables to label each axis based on the investigation,

• placing each variable on the correct axis (i.e., independent variable on the x-axis).

For the practice of Applying Equations, the sub-practice components (each scored as 1 point if the student successfully completed the component or else 0 points if the student was unsuccessful) included:

• identifying the correct mathematical relationship (i.e., linear, inverse, etc.) between variables in the graph,

• adjusting the values in the mathematical equation to get the best possible fit of the line to the data in the graph,

• creating a line with the correct mathematical relationship and a fit to the data of above 70%.

The average of each of these binarily scored sub-practice components was used to calculate the score for each math practice stage (i.e., ranging from 0 - 1 point). The inquiry practice score for the Constructing Graphs and Applying Equations stage for each separate lab investigation was used for analyses.

**Inq-Blotter Logs (with and without TIPS).** All teacher actions within Inq-Blotter (with or without TIPS) were automatically stored in log files and timestamped. The actions from Inq-Blotter that were captured included the alerts that appeared in the dashboard, the alerts selected by the teacher, and changes to any settings. The same information was captured for Inq-Blotter with TIPS, with the addition of whether teachers accessed the TIPS and the specific levels of guidance that were accessed (i.e., orienting TIP, procedural TIP, etc.).
Audio-Recordings. The principal investigator recorded all interactions for consenting participants during data collection. These data were transcribed, timestamped, assigned anonymous recording identifications numbers, and triangulated with the Inq-Blotter log data in order to identify students who were helped by a teacher in response to a dashboard alert (with or without TIPS). Additionally, the number of exchanges between teachers and students within each interaction were computed for analyses (i.e., a new teacher turn within an interaction indicated a new exchange).

Analyses

The Inq-ITS student performance data from the log files, teacher actions from the Inq-Blotter log files, and audio-recording data were triangulated using timestamps for the analyses.

Research Question 1. Analyses were first conducted to examine the relationship between student performance and teacher support in response to Inq-Blotter with TIPS (the relationship between student performance and teacher support without TIPS was examined in a prior study; see Study 1). The triangulated data were used to identify students who were helped by a teacher using Inq-Blotter with TIPS on a mathematical practice and whether students went on to complete another activity after being helped (to determine if students significantly improved on their next opportunity on the mathematical practice). Descriptive statistics were then used to identify whether the majority of these students \( n = 14 \) students total) improved on the math practice on which they were helped by a teacher using Inq-Blotter with TIPS and paired-samples t-tests were used to determine if this improvement was significant.
**Research Question 2.** The performance of students who were helped in the Inq-Blotter condition \((n = 18)\) was compared to the performance of students who were helped in the Inq-Blotter with TIPS condition (as identified via triangulated data; \(n = 14\)). Specifically, the amount of student improvement (the difference between students’ scores on the practice of difficulty from prior to after being helped) between conditions was examined using an independent samples t-test.

A Between-Subject MANOVA was used to identify whether there were any baseline differences between performance for all students in each condition (Inq-Blotter versus Inq-Blotter with TIPS) who completed the first activity \((n = 187)\). Next, the triangulated data were used to identify all students who completed all three Ramp Lab activities in each condition \((n = 133)\). A Mixed Model ANOVA was used to examine if there were any significant differences between student performance on mathematical practices in each group across activities.

**Research Question 3.** Descriptive statistics were used to explore differences in usage of basic Inq-Blotter alerts without TIPS compared to Inq-Blotter alerts with TIPS. Additionally, while interactions were qualitatively examined in a future study (see Study 3), an independent samples t-test was used to check for quantitative differences in interactions initiated by alerts between conditions. Specifically, the number of exchanges within each interaction (i.e., each time a new teacher turn occurred within an interaction) was computed and used for analyses.

**Results**

**Research Question 1 Results**
To answer RQ1 (Are teacher supports elicited by **Inq-Blotter alerts with TIPS** for using mathematics practices associated with student improvement on the practice on which they were helped in their next opportunity?), descriptive statistics were used to identify the proportion of students who improved on the mathematical practice on which they were helped by a teacher using Inq-Blotter with TIPS. Results indicated that 71% of students (10 out of 14 students helped by a teacher in response to an alert with TIPS) improved on the mathematical practice on which they were helped (see Table 13). The paired-samples t-test revealed that student improvement was significant overall from prior ($M = .33, SD = .26$) to after being helped ($M = .76, SD = .33$) across mathematical practices, $t(13) = -4.81, p < .001, d = 1.45$. Students also significantly improved on each practice (see Table 14). These findings demonstrate that teacher help elicited by Inq-Blotter alerts with TIPS was associated with student learning on the difficult mathematical practices of Constructing Graphs and Applying Equations. In the following section, student performance in the TIPS condition is compared to performance without TIPS in order to understand if the teacher support elicited by this design element further promoted student learning.

Table 13. Proportion of students with improvement on mathematical practices

<table>
<thead>
<tr>
<th>Inquiry Practice Alert</th>
<th># of students who improved/ # of students helped</th>
<th>% of students who improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>4/7</td>
<td>57%</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>6/7</td>
<td>86%</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>10/14</strong></td>
<td><strong>71%</strong></td>
</tr>
</tbody>
</table>
Table 14. Paired samples t-tests for math performance prior to after teacher help

<table>
<thead>
<tr>
<th>Inquiry Practice</th>
<th># students helped</th>
<th>Pre Help $M (SD)$</th>
<th>After Help $M (SD)$</th>
<th>Paired-Samples t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>7 students</td>
<td>.46 (.27)</td>
<td>.75 (.25)</td>
<td>$t(6) = -2.49, p = .047, d = 1.11$</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>7 students</td>
<td>.19 (.18)</td>
<td>.76 (.42)</td>
<td>$t(6) = -4.76, p = .003, d = 1.76$</td>
</tr>
</tbody>
</table>

Research Question 2 Results

Regarding RQ2 (Are there differences in student performance on mathematical practices across Inq-ITS activities when teachers used Inq-Blotter alerts versus Inq-Blotter alerts with TIPS?), analyses were conducted for students who were helped by a teacher using Inq-Blotter (with or without TIPS) as well as at the whole class level to examine the overall trend in student performance.

For students who were helped by a teacher using Inq-Blotter, an independent samples t-test was used to compare the amount of improvement students experienced from prior to after being helped in each condition ($n = 18$ students helped in the Inq-Blotter condition; $n = 14$ students helped in the Inq-Blotter with TIPS condition). The difference in the amount of improvement students experienced after being helped by a teacher using Inq-Blotter ($M = .49, SD = .36$) versus Inq-Blotter with TIPS ($M = .43, SD = .33$) was not significant, $t(30) = .54, p = .590$. These findings show that while students improved in both conditions (see Research Question 2 Results and Study 1), there was no difference in the amount of improvement after being helped in each condition. These analyses are useful for understanding the influence of support elicited by each version of the Inq-Blotter alerts for students who were helped directly, but it is also important to
examine the trajectory of student performance overall to identify any additional trends that emerge in relation to condition.

Prior to comparing student performance at the whole class level, it was necessary to identify whether there were any differences in baseline performance between students in the Inq-Blotter versus Inq-Blotter with TIPS conditions. Specifically, baseline differences between students’ performance on each inquiry practice stage (i.e., asking questions/hypothesizing, carrying out investigations, constructing graphs, and applying equations) in each condition were examined using a between-subjects MANOVA for students who completed the first ramp activity ($n = 187$ students). The overall MANOVA was not significant, $F(4, 182) = 2.21, p = .070, n^2 = .05$, which indicated that there were no significant differences overall between students’ inquiry performance in the Inq-Blotter ($M = .75, SD = .24$) versus the Inq-Blotter with TIPS condition ($M = .70, SD = .29$). Specifically, these results indicated that students across conditions started the Ramp Lab with similar competencies across practices (this is especially important to note given the different STEM subjects reflected in the current sample).

Follow-up analyses for the mathematical practices in particular revealed that there were no significant differences between student performance in the Inq-Blotter ($M = .72, SD = .27$) versus Inq-Blotter with TIPS condition ($M = .68, SD = .30$) for the Constructing Graphs inquiry practice stage, $F(1, 185) = 1.04, p = .310, n^2 = .01$. There was a difference, however, for Applying Equations where students in the Inq-Blotter condition ($M = .59, SD = .39$) had significantly higher performance to start relative to students in the Inq-Blotter with TIPS condition ($M = .44, SD = .41$), $F(1, 185) = 6.82, p = .010, n^2 = .04$. As a result, it is important to attend to how differences in performance
across activities on the inquiry practice stage of Applying Equations may relate to differences in prior knowledge (i.e., related to the science domain of physics versus chemistry versus biology versus engineering) in addition to experimental condition (i.e., teacher access to alerts with TIPS or without TIPS).

Following baseline comparisons, a MM MANOVA was used to examine student performance in each condition across the three ramp lab activities on the mathematical practice stages. Specifically, the repeated within-subjects factor was activity (i.e., the three Ramp Lab investigation activities) with condition as the between subjects factor and performance on Applying Equations and Constructing Graphs as the dependent variables for students who completed all three activities ($n = 133$ students). The results for the between-subjects effect, interaction effect, and within-subjects effect are explored below.

A significant between-subjects effect for condition was found where students in the Inq-Blotter condition ($M = .78$, $SD = .28$) had higher performance overall on mathematical practices relative to students in the Inq-Blotter with TIPS condition ($M = .67$, $SD = .33$), $F(2, 130) = 5.28$, $p = .006$, $n^2 = .08$. This finding indicates that students in the Inq-Blotter condition may have maintained higher performance on math practices when data were collapsed across activities, but does not indicate that these students improved more than students in the Inq-Blotter with TIPS condition from activity 1 to activity 3 (which is important to note given the higher performance of the alert condition students for applying equations). In particular, there was no interaction found between condition and performance (see Table 15), $F(4, 128) = 1.76$, $p = .140$, $n^2 = .05$. Therefore, Inq-Blotter alerts and Inq-Blotter alerts with TIPS were *similarly* effective in guiding teacher support that promoted student learning on mathematics practices overall across
activities. While there were no significant differences in student performance across activities by condition, the within-subjects effect was explored below to identify the overall trends in student performance on mathematics practices across activities.

Table 15. Average scores on math practices across activities in each condition

<table>
<thead>
<tr>
<th>Practice</th>
<th>Activity</th>
<th>Alerts $M (SD)$</th>
<th>Alerts with Tips $M (SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>1</td>
<td>0.79 (.19)</td>
<td>0.66 (.29)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.84 (.21)</td>
<td>0.79 (.23)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.87 (.19)</td>
<td>0.80 (.23)</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>1</td>
<td>0.62 (.37)</td>
<td>0.45 (.41)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.88 (.26)</td>
<td>0.81 (.35)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.69 (.43)</td>
<td>0.49 (.44)</td>
</tr>
<tr>
<td>Overall Math</td>
<td>1</td>
<td>0.71 (.28)</td>
<td>0.56 (.35)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.86 (.24)</td>
<td>0.80 (.29)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.78 (.31)</td>
<td>0.65 (.34)</td>
</tr>
</tbody>
</table>

There was a significant within-subjects effect indicating that students overall had differences in performance across each Ramp Lab activity, $F(4, 128) = 29.69, p < .001$, $n^2 = .48$. Specifically, follow-up comparisons using t-tests revealed that students significantly improved from the first to the second and from the first to the third activities, but had a significant decline in performance from the second to the third activity (see Table 16 and Figure 16). It is promising to see that students improved overall from the first to the third activity, but the decline in performance from the second to the third activity can be better understood when looking at the practices of Constructing Graphs and Applying Equations independently. Specifically, follow-up analyses for each practice indicated a different trend in performance on Constructing Graphs relative to Applying Equations as explained below.
Table 16. Comparisons of student performance across activities on math practices

<table>
<thead>
<tr>
<th>Practice</th>
<th>Comparison</th>
<th>Results of t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructing Graphs</td>
<td>Activity 1 ($M = .74$, $SD = .24$) &amp; Activity 2 ($M = .82$, $SD = .21$)</td>
<td>$t(132) = -4.13$, $p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>Activity 2 ($M = .82$, $SD = .21$) &amp; Activity 3 ($M = .85$, $SD = .21$)</td>
<td>$t(132) = -1.71$, $p = .090$</td>
</tr>
<tr>
<td></td>
<td>Activity 1 ($M = .74$, $SD = .24$) &amp; Activity 3 ($M = .85$, $SD = .21$)</td>
<td>$t(132) = -5.35$, $p &lt; .001$</td>
</tr>
<tr>
<td>Applying Equations</td>
<td>Activity 1 ($M = .56$, $SD = .39$) &amp; Activity 2 ($M = .86$, $SD = .29$)</td>
<td>$t(132) = -8.31$, $p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>Activity 2 ($M = .86$, $SD = .29$) &amp; Activity 3 ($M = .62$, $SD = .45$)</td>
<td>$t(132) = 5.81$, $p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>Activity 1 ($M = .56$, $SD = .39$) &amp; Activity 3 ($M = .62$, $SD = .45$)</td>
<td>$t(132) = -1.27$, $p = .205$</td>
</tr>
<tr>
<td>Overall Math</td>
<td>Activity 1 ($M = .65$, $SD = .27$) &amp; Activity 2 ($M = .84$, $SD = .22$)</td>
<td>$t(132) = -8.66$, $p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>Activity 2 ($M = .84$, $SD = .22$) &amp; Activity 3 ($M = .73$, $SD = .27$)</td>
<td>$t(132) = 4.67$, $p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>Activity 1 ($M = .65$, $SD = .27$) &amp; Activity 3 ($M = .73$, $SD = .27$)</td>
<td>$t(132) = -3.13$, $p = .002$</td>
</tr>
</tbody>
</table>

Figure 16. Student performance on mathematical practices across activities
A significant within-subjects effect was found for the practice of Constructing Graphs indicating differences in student performance across activities, $F(2, 262) = 20.83$, $p < .001$, $n^2 = .14$. Follow-up comparisons revealed that students significantly improved on constructing graphs across conditions from the first to the second, the second to the third, and the first to the third activities (see Table 16 and Figure 17). Thus, students in classrooms where teachers used Inq-Blotter or Inq-Blotter with TIPS primarily learned how to construct graphs successfully by their third activity. The decline in student performance seen in the overall math scores from the second to the third activity would therefore be explained by students’ performance on Applying Equations.

![Constructing Graphs Performance Across Activities](image)

**Figure 17.** Student performance on constructing graphs across activities

In terms of the practice of Applying Equations, a significant within-subjects effect was also found indicating a difference in student performance across activities, $F(2, 262) = 30.69$, $p < .001$, $n^2 = .19$. The results of follow-up comparisons indicated that students significantly improved from the first to the second activity on Applying Equations, but
had a significant decline in performance from the second to the third activity (see Table 16 and Figure 18). Part of the decline in student performance from the second to the third activity might be explained by the increased difficulty of the mathematical relationship. Specifically, students were applying a positive linear function in the second activity versus a negative linear function in the third activity. Negative linear relationships have been shown to be particularly difficult for students in terms of understanding how the coefficient can change the direction of the line (in addition to steepness) and how to correspondingly adjust the constant to reflect the y-intercept (e.g., De Bock et al., 2017). It is possible that the support teachers provided on Applying Equations in the first and second activities did not include the mathematical content understandings necessary to help students successfully apply equations in the third activity (i.e., that slope can also change the direction of a line).

Figure 18. Student performance on applying equations across activities
Future iterations of TIPS should emphasize the role of the coefficient (and related changes that are necessary to the constant) in changing the direction of the line, even when the relationship is positive in order for students to be able to apply their understandings across activities. In the interim, it is possible to better understand student performance in each condition through a quantitative exploration of teachers’ use of the dashboard alerts and corresponding interactions that occurred.

Research Question 3 Results

For RQ3 (Are there any differences in how teachers use Inq-Blotter alerts versus Inq-Blotter alerts with TIPS?), the usage data for teachers who implemented Inq-Blotter versus Inq-Blotter with TIPS in their classrooms were explored. Over 400 alerts were generated across teachers across conditions (there were some data not stored for one teacher, Ms. A, in the basic Inq-Blotter alerts condition due to a system logging error, so the number of alerts generated/accessed was calculated based on audio-recordings). A greater proportion of alerts across practices were opened in the alerts condition (45% of alerts opened) relative to the alerts with TIPS condition (16% of alerts opened), but a majority of students who were helped on mathematical practices and went on to complete a second activity improved after being helped in each condition (i.e., over 70% of students improved across conditions; see Table 13). Some reasons that teachers may not have opened alerts included that students moved on to the next inquiry activity before the teacher had an opportunity to access the alert and also that teachers did not have enough time in the class period to open all alerts. It is important to use caution when interpreting these usage data given the differences in number of participants and other contextual variables across conditions (i.e., length of class periods), but these data do still provide
insight into how teachers are able to drill down to provide individualized support to
students on their specific difficulties. For example, an exploration of the number of
exchanges between teachers and students that occurred within interactions in response to
alerts allows for beginning to identify potential differences in how teachers responded to
alerts to support students in each condition.

Table 17. Inq-Blotter usage data (with and without TIPS)

<table>
<thead>
<tr>
<th>Condition</th>
<th># of students</th>
<th># of alerts generated</th>
<th># of alerts opened</th>
<th># of students helped on math (completed second activity)</th>
<th># of students improved on math</th>
<th>Average # of teacher turns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alerts</td>
<td>119</td>
<td>&gt; 413</td>
<td>&gt; 184</td>
<td>18</td>
<td>16</td>
<td>7.67</td>
</tr>
<tr>
<td>Alerts with TIPS</td>
<td>106</td>
<td>535</td>
<td>86</td>
<td>14</td>
<td>10</td>
<td>1.14</td>
</tr>
</tbody>
</table>

In terms of the interactions between teachers and students, an independent
samples t-test revealed that there was a significant difference in the number of exchanges
when teachers responded to Inq-Blotter alerts ($M = 7.67$, $SD = 6.33$) versus Inq-Blotter
alerts with TIPS ($M = 1.14$, $SD = .36$), $t(30) = 3.84$, $p = .001$, $d = 1.46$. Specifically, there
were fewer back-and-forth turns between teachers and students within interactions in
response to Inq-Blotter alerts with TIPS. Therefore, teachers in the Inq-Blotter with TIPS
condition often provided support to students without requiring significant student
interaction.

One potential explanation for the difference in interactions between conditions
could be that teachers with TIPS did not need to ask students for additional contextual
information or to share their screens to be able to provide support (see Table 18). For
example, in the first interaction in Table 18 (Inq-Blotter Alerts Condition), five exchanges occurred in relation to the student sharing their screen before the teacher provided any specific inquiry support and the teacher was still unsure about the students’ specific difficulty without being able to see their screen.

Table 18. Examples of interactions elicited by Inq-Blotter alerts with and without TIPS

<table>
<thead>
<tr>
<th>Condition</th>
<th># of exchanges</th>
<th>Example interaction in response to Applying Equations Alert [Recording ID]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inq-Blotter Alerts</td>
<td>8</td>
<td>T: Okay, how are you guys doing over here? I saw you guys might have some trouble with fitting the data…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: Ummm I think I’m okay, but I keep looking at some things that I saw that ummm…I thought it was gonna be linear but I guess like…I, I’m not completely sure but I think it’s supposed to be inverse…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: Why don’t you share your screen and we’ll take a look at the data together?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: Sure…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: …You wanna share your screen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: Yeah sorry it’s just taking a long time…I’m trying to share it but it’s just…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: Well if you click the “present now” and just share your entire screen then it will just mirror what you have in front of you, so you can just click on the tab…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: Maybe if we go back on the actual meet?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: Maybe? For now another group was asking for some help so as long as you guys have a general idea of where you’re going, then I’ll leave you here for right now and I’ll pop back in in a couple of minutes or let me know umm if you have any other issues.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: Alright, thank you</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: Uh huh, and so you might have to do, might have to change the coefficient so that it will tilt it a little bit better and so on, so…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: Ummm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: You’re still working on the height one?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S: I think this is the height one?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T: Yes. Okay, alright, continue working, okay? [A111]</td>
</tr>
</tbody>
</table>

| Inq-Blotter Alerts with TIPS | 1              | T: [S] take a look at your graph. How well does the line in your model fit your data? You wanna get your curve as close to as many data points as possible. So if you change the coefficient, aka “a” that will change the steepness or direction of the line. And if you change the constant “b” it will shift the line or your curve up or down…So try and put different numbers in for your coefficient and for your constant to get that best fit line as close as possible. If that makes sense? [ B60] |
This varied from the second interaction in Table 18 (Inq-Blotter Alerts with TIPS Condition) where the teacher was able to provide specific supports to the student regarding their difficulty with creating an equation that fit the data and no screensharing was necessary. The additional information within the TIPS that were accessed by teachers (see Table 19) may have provided enough detail for teachers to evaluate students and also determine the type of help needed by students. For instance, teachers were able to read directly from the TIPS or base their scaffolding directly from the instructional supports within the TIPS. Student performance, however, did not significantly benefit from the teacher support elicited by TIPS relative to basic alerts without TIPS. Further analysis of the discourse within interactions is needed to fully understand the relationship between the information within alerts, the supports provided by teachers, and students’ corresponding performance (see Study 3).

Table 19. Inq-Blotter with TIPS usage data

<table>
<thead>
<tr>
<th># of times TIPS were accessed</th>
<th># of Orienting TIPS</th>
<th># of Conceptual TIPS</th>
<th># of Procedural TIPS</th>
<th># of Instrumental TIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>36</td>
<td>18</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

**Discussion**

As previously stated, it is essential to attend to the ecological validity (Laurillard, 2008) of innovative educational technologies to ensure their efficacy for teachers and inform future design iterations to meet teachers’ needs. While several new technologies have been developed to support STEM educators (e.g., Holstein et al., 2019; Martinez-Maldonado et al., 2015; Schwarz et al., 2018), these tools do not report on key science inquiry practices and have yet to be evaluated during synchronous remote instruction.
The present study demonstrates an approach to developing and testing technologies that support critical STEM practices within remote classroom contexts. Specifically, alerts within the science inquiry dashboard Inq-Blotter were expanded to include TIPS that guided teacher support to students on difficult mathematical practices in the intelligent tutoring system, Inq-ITS. This work outlined the TIPS development process, demonstrated how TIPS informed teacher support associated with students’ performance on mathematical practices, and compared differences in student performance and teacher actions when using Inq-Blotter alerts with TIPS versus without TIPS.

TIPS were developed based on previously successful teacher scaffolds and were reviewed by educators with the goal of providing the instructional support (contextual and content) needed for teachers to guide students on mathematical practices remotely. The results of this study showed how the majority of students helped by a teacher using TIPS significantly improved on the math practice on which they were helped. This finding has important implications for the potential of alerting dashboards to not only direct teacher attention towards struggling students, but also elicit interactions that result in learning. There were interesting findings, however, related to the impact of support elicited by alerts with TIPS relative to basic Inq-Blotter alerts.

Further comparisons of the use of Inq-Blotter alerts with versus without TIPS revealed no significant differences in the effectiveness of each version of the technology in guiding teacher support associated with improvement on mathematical practices. While students in the alerts with TIPS condition had slightly lower performance in general, there was no relationship between the amount of improvement students experienced and their condition. Students across conditions were found to demonstrate
significant learning gains on the practice of Constructing Graphs. For the practice of Applying Equations, however, students were found to struggle with the third activity involving a negative linear relationship across conditions (versus the positive linear relationship in the second activity). This result implied the need for further revisions of the content within the dashboard to better initiate support on varying types of mathematical relationships. Given that there were minimal differences in student performance across conditions, it was important to examine if there were differences in teachers’ implementations of Inq-Blotter alerts with versus without TIPS.

While the usage data extracted from log files provide insight into the alerts generated, accessed, and acted upon across conditions, caution should be taken in drawing conclusions from these data alone given the differences in number of participants and other factors that may have impacted these logs (i.e., length of class periods). There were significant differences, however, found in the number of exchanges that occurred between teachers and students in response to alerts with TIPS versus without TIPS. In particular, there were substantially more back-and-forth exchanges between teachers and students when using basic Inq-Blotter alerts (without TIPS). This finding may have been due to a number of factors including that teachers with only basic alerts needed to request more information from students or needed students to share their screen (and therefore walk them through that process) before scaffolding them on the inquiry practice of difficulty. Further analyses are needed to explore the discursive supports elicited by teachers in greater detail and to unpack the connection to student performance (see Study 3).
In addition to examining the types of discursive supports provided by teachers in response to alerts (with and without TIPS), future studies should examine the use of the technologies with additional teachers and students. These analyses would help to account for possible differences in prior knowledge of both teachers and students that may have interacted with their performance on inquiry practices and the effectiveness of supports (i.e., in the present study, students in the Inq-Blotter condition had higher performance to start on the mathematical practice of applying equations). In particular, studies should investigate potential differences that may emerge in relation to the content of the participating STEM classes (i.e., physics and engineering students in Phase 1 may have had greater prior knowledge related to Applying Equations). Additional data would also enable controlled comparisons of performance when students received help versus did not receive help from a teacher in response to an alert with TIPS.

Studies might also more closely examine the factors that influence teachers’ decisions to respond to particular alerts. For instance, teachers could be interviewed regarding whether their choices to respond to alerts were related to the recency of alerts, type of alerts, background on the students’ competencies outside of Inq-ITS, or other factors. In the TIPS condition, it would be helpful to know more information about the decision-making process behind when teachers decided to access TIPS in general as well as each of the specific levels of TIPS. Further, it will be important to collect data on the use of TIPS within in-person contexts to fully assess the ecological validity of the tool. There may be different needs as well as different ways of using the technology within an in-person classroom setting (i.e., the teacher has access to other classroom cues that are used to inform guidance).
In conclusion, the results of this study show the potential of an alerting dashboard to support students on critical STEM practices. Importantly, this study demonstrates the need to evaluate the efficacy of new technological design elements to identify if/when they add value. Researchers can then determine potential reasons why the tools did not work as expected in order to inform future design iterations. Overall, the present study is an important step towards better understanding the connection between dashboard use and student performance on science inquiry practices.
Phase 2: Study 3

Abstract

Classroom discourse can have a strong impact on student understandings and learning outcomes in STEM contexts (Lemke, 1990), so it is important to provide educators with the instructional supports needed to promote productive discourse (Michaels & O’Connor, 2017). Technologies such as dashboards can present teachers with data to inform classroom discourse and support student learning (Gee, 2004; Lemke, 1990; Scott, 1998). It is important to better understand how teacher dashboards are used in real time and how the different types of discursive supports used by teachers impact student learning of complex STEM practices (NGSS, 2013), including Using Mathematics in science (the focus of the present work). The teacher dashboard, Inq-Blotter, is one of the only dashboards known to date that provides real-time data to teachers on students’ inquiry practice competencies as they complete virtual lab investigations in the Inquiry Intelligent Tutoring System, Inq-ITS.

The present study investigated the discourse elicited by Inq-Blotter in relation to newly embedded instructional supports for teachers within the dashboard alerts. Specifically, Epistemic Network Analyses (ENA) were used to make comparisons between: 1) patterns in teacher support elicited by basic Inq-Blotter alerts versus Inq-Blotter alerts with instructional supports and 2) patterns of support associated with student improvement or no improvement. Additionally, the present study examined the effectiveness of different representations of discursive supports for predicting student learning outcomes (i.e., frequencies versus weights of connections from ENA). The results of this study have implications for both how actionable alerts with fine-grained
data on students’ competencies within dashboards can shape classroom discourse, and the effectiveness of representations of discourse for capturing patterns in teacher support.

**Introduction to Phase 2 Study 3**

A number of core ideas in STEM classrooms are shared between teachers and students through classroom discourse (i.e., exchange of oral language; Gee, 2004; Lemke, 1990; Scott, 1998). Studies have shown that the wording teachers use for particular questions or feedback can impact how students learn core STEM competencies such as the science inquiry practices outlined in the Next Generation Science Standards (Manz & Renga, 2017; McNeill & Krajcik, 2008; Talanquer et al., 2013). Providing higher level discursive support to students on complex STEM competencies requires both identifying student difficulties and determining how to support those difficulties (Luna, 2018; van Es & Sherin, 2002; Shulman, 1987; Talanquer et al., 2013), which can be extremely challenging for teachers. These challenges are further magnified by remote learning contexts where STEM teachers have limited assessment opportunities (Arnett, 2021; Chatterjee, 2020; Dukes, 2020; Means et al., 2021). Even when assessments are available, difficult STEM practices such as Using Mathematics (Lai et al., 2016; LópezLeiva et al., 2016; McDermott et al., 1987; Potgieter et al., 2008; see Study 1) can be challenging for teachers to support (Pruitt, 2014). It is important to evaluate teacher discourse in STEM to identify areas that require further instructional support.

Researchers have created frameworks (e.g., Classroom Discourse Analysis Tool, Lee & Irving, 2018; Electronic Quality of Inquiry Protocol, Smart & Marshall, 2012) and typologies of types of teacher supports (e.g., teacher talk moves including question types, Manz & Renga, 2017; Michaels & O’Connor, 2017) in order to examine the relationship
between teacher discourse and student STEM learning outcomes. Most discourse studies implement analytic approaches such as frequency calculations to analyze trends in discourse (Howe et al., 2019; Lee & Irving, 2018; Manz & Renga, 2017; Talanquer et al., 2013), but finer-grained approaches such as Epistemic Network Analysis (ENA; Shaffer et al., 2016) have more recently been developed that can capture rich, dynamic patterns in discourse in addition to the frequencies of types of discourse. Analytic approaches such as ENA are important for characterizing and evaluating teacher discourse to determine if instructional supports are needed to better guide teacher scaffolding (Scott, 1990). Specifically, these analyses can be used to inform the design of advanced educational technologies including teacher dashboards to help guide teacher support of student difficulties in STEM across settings (Roschelle et al., 2017).

STEM dashboards can help teachers to monitor student progress in online environments and provide discursive supports based on their formative assessment of student difficulties (Holstein et al., 2018a; Knoop-van Campen & Molenaar, 2020; Martinez-Maldonado et al., 2013; Tissenbaum & Slotta, 2019; Verbert et al., 2014). Studies have demonstrated how dashboards can elicit higher levels of teacher support in STEM contexts (e.g., feedback on processes in addition to just evaluative feedback; Knoop-van Campen & Molenaar, 2020), but few studies have explicitly examined the connections between teacher discourse elicited by dashboards and student learning outcomes (Dickler et al., 2021a; see Study 1).

Recent studies on the Inq-Blotter teacher dashboard at the middle school level have demonstrated how teacher support elicited by real-time alerts with data on students’ inquiry competencies was associated with middle school student improvement on inquiry
practices in the intelligent tutoring system, Inq-ITS (Dickler et al., 2021a). Specifically, students’ improvement on inquiry practices was related to teachers providing higher level conceptual supports (i.e., explanations of the inquiry practice) and procedural supports (i.e., overview of general sub-practices involved in an inquiry practice) on the inquiry practice of difficulty in response to a dashboard alert. Studies with Inq-Blotter at the high school level have also shown that real-time alerts can elicit teacher support associated with student improvement on mathematical practices (i.e., Constructing Graphs and Applying Equations; see Study 1 and Study 2). In particular, two versions of Inq-Blotter alerts were tested with either basic information on the inquiry practice difficulty experienced by a student or with additional instructional supports to guide teacher scaffolding (i.e., Teacher Inquiry Practice Supports, TIPS). While the majority of students improved on the mathematical inquiry practice on which they were helped regardless of teacher condition (i.e., basic alerts or alerts with TIPS), there were no significant differences in the amount of improvement in each condition on students’ next opportunity to engage in the practice on which they were helped (in the short-term). Further research is necessary to determine whether the basic alerts versus alerts with TIPS elicited different patterns in teachers’ discursive supports. In particular, it would be expected that the alerts with TIPS would help to elicit higher-level supports from educators (relative to teachers using basic alerts without instructional supports).

To investigate this hypothesis, the present study examined the teacher discourse elicited by the Inq-Blotter dashboard alerts with and without TIPS in remote high school STEM classrooms as students completed investigations in Inq-ITS. Comparisons were made to understand differences in patterns of support elicited by each version of the
dashboard alerts as well as differences in patterns of support associated with student learning outcomes on mathematical practices (i.e., whether students improved or did not improve on the mathematical inquiry practice on which they were helped). Additionally, both frequencies of discursive supports and weights of connections between combinations of supports (from ENA) were applied to predict the amount of student after being helped by a teacher using Inq-Blotter. Specifically, the following research questions were examined:

1) Are there significant differences in the patterns of teacher support (examined using ENA) elicited by Inq-Blotter alerts versus Inq-Blotter alerts with TIPS for mathematical practices?

2) Are there significant differences in the patterns of teacher support (examined using ENA) when students improved versus did not improve on mathematical practices?

3) What representation of discursive supports (i.e., frequencies versus weights of connections from ENA) best predict student improvement on mathematical practices?

Methods

Participants

The participants in the present study included six teachers (two teachers each from three participating high schools) and their high school students ($N = 212$ students total; see Table 20). One teacher from each school was assigned to use basic Inq-Blotter alerts with their students ($n = 119$ students total) and the other teacher from each school was assigned to use Inq-Blotter alerts with TIPS with their students ($n = 106$ students)
total). All participation occurred remotely between December of 2020 and February of 2021 during regularly scheduled synchronous class periods.

Table 20. Demographics for participating high schools

<table>
<thead>
<tr>
<th>High School</th>
<th>Alert Teacher (# of students)</th>
<th>Alerts+TIPS Teacher (# of students)</th>
<th>Total # of Students</th>
<th>% of students eligible for FRL</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School 1</td>
<td>Ms. A - Engineering (n = 29 students)</td>
<td>Ms. X - Physics (n = 7 students)</td>
<td>36</td>
<td>73%</td>
</tr>
<tr>
<td>High School 2</td>
<td>Ms. B - Chemistry (n = 42 students)</td>
<td>Ms. Y - Biology (n = 73 students)</td>
<td>115</td>
<td>32%</td>
</tr>
<tr>
<td>High School 3</td>
<td>Mr. C - Physics (n = 48 students)</td>
<td>Ms. Z - Biology (n = 26 students)</td>
<td>74</td>
<td>8%</td>
</tr>
</tbody>
</table>

**Materials**

**Inq-ITS Tutorial Lab.** Students completed a brief tutorial lab in the Inq-ITS system prior to the start of data collection. The tutorial lab involved an investigation (i.e., exploring if adding red dye to a flower in a vase impacted the redness of the flower petals) to introduce students to the widgets (i.e., tools) and structure of the labs within Inq-ITS.

**Inq-ITS Ramp with Graphing Lab.** Students completed a Ramp with Graphing Lab (i.e., Ramp Lab) during data collection. The Ramp Lab was explicitly designed for high school students to practice applying mathematics within a science context. The three lab investigation activities involved exploring the mathematical relationship between: the height of a tower a ramp was leaning on (i.e., steepness of the ramp) and the time it took a sled to reach the end of the ramp, the mass of a sled and the momentum of the sled going down a ramp, and the roughness of the ramp (i.e., friction between the sled and the ramp) and the acceleration at the end of the ramp.
Each lab investigation activity included 6 inquiry practice stages aligned to the NGSS (2013) practices. In the first two stages, students asked questions/hypothesized about the mathematical relationship between variables and then used a simulation to carry out an investigation to test their hypothesis. Students were not able to move to the simulation until they created a question/hypothesis that aligned to the goal of the investigation (and received automated guidance on the variables to select if they continued to struggle). Students began applying mathematics in the third stage of the lab, Constructing Graphs (see left of Figure 19), where they selected data that they collected to plot on a graph and selected the variables to label each axis of the graph. Students then moved on to the fourth stage, Applying Equations (see right of Figure 19), where they
selected the mathematical relationship between the variables in their graph and adjusted values in the equation to create a best fit line. In the final two stages, students summarized their findings using widgets and text boxes.

**Inq-Blotter.** One teacher from each high school \((n = 3\) teachers total) used regular Inq-Blotter (without TIPS) during data collection. The Inq-Blotter dashboard supported real-time monitoring of students’ performance within the first four stages of Inq-ITS (i.e., Asking Questions/Hypothesizing, Carrying out Investigations, Constructing graphs, and Applying Equations). Teachers received real-time alerts regarding individual students who were having difficulty with an inquiry practice and the specific part of the stage that was challenging for them (e.g., a student was struggling with Constructing Graphs because they were not selecting the appropriate variable for their x-axis; see Figure 20).

![Figure 20. Inq-Blotter Dashboard with a Constructing Graphs Alert](image)

Teachers could also receive alerts related to whole class difficulties on an inquiry practice stage (e.g., more than 50% of students were struggling with Constructing Graphs), and individual students who were making slow progress on an inquiry practice
stage (i.e., a student was on the Constructing Graphs stage for longer than 5 minutes).

Teachers could organize the presentation of the alerts by recency or by type of inquiry practice alert. Teachers could also view a list of all active students and select a student name to see their progress.

**Inq-Blotter with TIPS.** The other teacher from each high school (n = 3 teachers total) used Inq-Blotter with TIPS during data collection. All of the alerts in Inq-Blotter were accessible to teachers in this condition, but teachers also had access to instructional supports embedded within alerts by clicking “Press for TIPS (Teacher Inquiry Practice Supports)” (see Figure 21). After clicking the link, teachers could access four levels of support in the right panel of the alert (see Figure 22), including:

- orienting (i.e., directing the students’ attention towards the components of the inquiry practice stage that required revision),
- conceptual (i.e., explaining the purpose and underlying components of the inquiry practice),
- procedural (i.e., outlining the general steps involved in completing the practice successfully),
- instrumental (i.e., telling the student the explicit steps needed to complete the practice successfully).

Teachers were able to navigate between the support types by clicking on the blue arrows and correspondingly could skip between the different levels of support. Teachers could also choose to minimize the TIPS at any time.
Figure 21. Inq-Blotter Dashboard with TIPS for Constructing Graphs Alert

Figure 22. Four levels of Inq-Blotter TIPS for a Constructing Graphs Alerts
Procedure

Prior to the start of data collection, the principal investigator virtually met with each teacher individually to provide a tutorial of the Inq-ITS system and teacher resources, as well as coordinate the plan for the data collection session. The teachers were assigned to either the Inq-Blotter or Inq-Blotter with TIPS condition, and received a detailed overview of the respective technology. The teachers specified their preferred virtual classroom meeting platform and date for data collection for each of their participating class periods. The teachers then assigned the Inq-ITS tutorial lab to their students to complete on their own for homework prior to the scheduled data collection date (between December of 2020 and February of 2021).

On the day of data collection, the principal investigator joined the virtual meeting platform designated by the teacher (e.g., Zoom or Google Meets), was introduced to the students, and briefly introduced the Inq-ITS Ramp Lab. The investigator then began audio-recording all interactions for consenting teachers and students. The teachers assigned students the Ramp Lab to complete during the class period (45-55 minutes; students in the Inq-Blotter without TIPS condition also completed an 8-item multiple choice pre-test prior to beginning the Ramp Lab that was not examined in the present study). While students completed the Ramp Lab, the teachers monitored students’ performance using either Inq-Blotter or Inq-Blotter with TIPS depending on their respective condition. At the end of the class period, the teachers instructed students to finish any remaining Ramp Lab activities for homework. All teachers and students had
continued access to the technologies for their regular classroom use following completion of the study.

**Measures**

**Mathematical Practice Performance in Inq-ITS.** For the purposes of the present study, students’ scores on only the mathematical practice stages in Inq-ITS (i.e., Constructing Graphs and Applying Equations) were examined. The mathematical stages were automatically scored using knowledge-engineered algorithms (Sao Pedro & Betts, 2019; see Study 1). Specifically, students were scored based on whether they successfully completed components (i.e., sub-practices) of each stage successfully (1 point) or not (0 points). The components involved in Constructing Graphs included correctly labeling the axes of the graph (i.e., the independent variable on the x-axis and dependent variable on the y-axis), selecting variables related to the investigation for the graph axes (i.e., based on the goal of the activity), selecting data from controlled trials to plot in the graph, and selecting sufficient data to plot in the graph. The components involved in Applying Equations included selecting the correct mathematical relationship between the variables in the graph, updating the values in the equation to construct a line that has at least a 70% fit to the data, and constructing at least one line with both the appropriate mathematical relationship and fit to the data of at least 70%. The students’ scores on each mathematics stage in each activity were calculated by taking the average of the component scores (i.e., final scores between 0 - 1 points on each stage).

**Teacher Actions in Inq-Blotter.** All teacher actions within the Inq-Blotter dashboard (e.g., selecting an alert) and data from the dashboard (e.g., alerts that appeared,
content of alerts) were automatically stored. Therefore, timestamped data from the log files were extracted and used for analyses.

**Teacher Actions in Inq-Blotter with TIPS.** The same data from Inq-Blotter (without TIPS) were logged for Inq-Blotter with TIPS, with the addition of teachers’ actions in regards to accessing the TIPS (i.e., selecting the TIPS, selecting an orienting support, selecting a conceptual support, etc.). The timestamped data related to teachers use of alerts with TIPS were extracted for use in analyses.

**Audio-Recordings.** The principal investigator audio-recorded classroom interactions for consenting participants through a laptop using QuickTime. Audio-recordings were transcribed, anonymized, and timestamped.

**Analyses**

The student performance data from Inq-ITS, Inq-Blotter log data, Inq-Blotter with TIPS log data, and audio recordings were triangulated using timestamps. The triangulated data were used to identify all interactions that occurred in response to an Inq-Blotter alert (with or without TIPS) for the practice stages of Constructing Graphs and Applying Equations. These data were also used to determine when students improved after receiving teacher help (for students who completed an activity after being helped) and the amount of improvement (i.e., the difference in performance on the mathematical inquiry practice from prior to being helped versus after being helped). There were interactions in which the teacher conversed with multiple students, but only data for the student who had an alert activated most recently was included in the present analyses. There were also some students who were supported across multiple interactions on different practices (e.g., a student received help from their teacher on Applying Equations based on an alert,
and the teacher later returned to the same alert because the student had not progressed; a student received support from their teacher on Applying Equations, and later received support from their teacher on Constructing Graphs) and each of these interactions were included in the analyses.

The audio-recordings for interactions were divided into turns by speaker and only the teacher turns were used for analyses to examine patterns within the teacher support \( (N = 277 \text{ teacher turns}) \). The teacher turns from all interactions were coded using the scheme applied in prior studies (Dickler et al., 2021a; see Study 1) to capture the different levels of support aligned to the levels of support in TIPS (i.e., orienting, conceptual, procedural, instrumental), evaluative comments, and content comments (see Table 21). Additionally, a new code was added to capture the presence of additional technical supports given by teachers that emerged in the discourse when teachers were using the dashboard remotely (i.e., screensharing support; see Table 21).

Two doctoral student researchers were given an overview of the coding scheme by the principal investigator and trained with a set of data from a prior study. The three researchers then coded 31 turns together from the current data set and disagreements were discussed. The two doctoral student researchers split the remaining turns, each coding 123 turns. The principal investigator coded all remaining teacher turns \( (n = 246 \text{ turns}) \). Agreements between the principal investigator and each researcher were computed (81% agreement between the first student the principal investigator, and 91% agreement between the second student and the principal investigator). Disagreements were discussed amongst all three researchers and agreed upon codes were used for analyses.
Table 21. Teacher discursive support codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Applying Equations Examples</th>
<th>[Recording ID]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orienting Scaffold</td>
<td>Reminds the student/directs the students’ attention towards a practice</td>
<td>Looking at the data on your graph. What is the trend in the data? [B135]</td>
<td></td>
</tr>
<tr>
<td>Conceptual Scaffold</td>
<td>Explains components involved in an inquiry practice to a student</td>
<td>So the goal is to choose a relationship or a mathematical equation that represents the trend in the data. So you can click on the relationship dropdown menu to see what each mathematical relationship looks like. [B104]</td>
<td></td>
</tr>
<tr>
<td>Procedural Scaffold</td>
<td>Tells the students the steps involved in an inquiry practice</td>
<td>So you should always have at least 5 data points…and the more the better so you can actually see what the graph's gonna look like [A71]</td>
<td></td>
</tr>
<tr>
<td>Instrumental Scaffold</td>
<td>Tells the student exactly what to do/click to move forward on an inquiry practice</td>
<td>So just keep sliding around [the coefficient] until you find the right fit [A17]</td>
<td></td>
</tr>
<tr>
<td>Content Comment</td>
<td>Tells or asks the student about some scientific content</td>
<td>…if I change the height, the time’s gonna change right? [A94]</td>
<td></td>
</tr>
<tr>
<td>Evaluative Comment</td>
<td>Makes a comment regarding whether something is right, wrong, correct, or has an issue</td>
<td>It looks like you are having trouble with your line. [B43]</td>
<td></td>
</tr>
<tr>
<td>Technical Comment</td>
<td>There are comments about the use of external technology (zoom, google, internet, sharing screen, etc.)</td>
<td>Why don't you share your screen and we'll take a look at the data together [A111]</td>
<td></td>
</tr>
</tbody>
</table>

**Research Question 1.** Epistemic Network Analyses (ENA; Shaffer et al., 2016) were used to compare patterns of teacher support across Inq-Blotter alert conditions. The coded teacher turns from each interaction that occurred in response to a mathematical practice alert in Inq-Blotter (with or without TIPS) were used to build the Epistemic
Networks. The networks were constructed using the ENA Web Tool (version 1.7.0; Marquart et al., 2018) to first compare teacher support patterns in response to basic alerts (without TIPS) versus alerts with TIPS. All seven teacher discursive support codes were used as the nodes in the networks and the unit of analysis was each interaction elicited by an alert in both conditions (without TIPS versus with TIPS). The stanza size was set to be the whole conversation (i.e., each interaction was treated as a conversation). Contributions from all participating teachers were represented in the networks, however, the majority of the contributions for the alerts with TIPS condition (90% of interactions) were from Ms. Z. The corresponding networks for alerts and alerts with TIPS were quantitatively compared using t-tests (i.e., examining differences in the mean centroids of the interactions in each network) and qualitatively compared using the subtracted network (i.e., the difference in weighted connections between the two networks).

**Research Question 2.** Epistemic networks were also constructed to compare the difference in support patterns associated with student improvement versus no improvement on the mathematical practices on which students were helped (by a teacher using basic alerts or alerts with TIPS). The seven discourse codes were again used as the nodes, the unit of analysis was each interaction associated with student improvement or no improvement, and the stanza size was the whole conversation. Networks were quantitatively compared using t-tests and qualitatively compared based on the subtracted network.

**Research Question 3.** Finally, the effectiveness of representations (i.e., frequencies of discursive supports versus weights of connections between supports from the ENA) for capturing the relationship between discourse and student learning outcomes
were compared. The average frequency of each support type across interactions was computed for when students improved or did not improve. The interpretation of support types associated with improvement based on frequencies was compared to the interpretation based on the epistemic networks.

In order to explore if frequencies would be more effective than ENA at predicting student improvement, a linear regression model was constructed with the frequency for each discursive support code as the predictor variables and the amount of student improvement as the dependent variable (i.e., the difference between student performance on the practice from prior to after being helped).

Another linear regression model was built with the weight of connections between discursive support codes in the ENA as the predictor variables (i.e., the weight of the connection in the epistemic network for each interaction between conceptual support and procedural support, between conceptual support and orienting support, etc.) and the amount of student improvement as the dependent variable. Given that there were 21 combinations of connections between support codes, only the higher-level support code connections (described next) were included as predictor variables in the final model to prevent over fitting ($n = 11$ connections). Specifically, connections between procedural supports with other support types (procedural-conceptual, procedural-orienting, procedural-instrumental, procedural-evaluative, procedural-content, procedural-technical) and conceptual supports with other support types (conceptual-orienting, conceptual-instrumental, conceptual-evaluative, conceptual-content, conceptual-technical) were included based on their importance as identified in the subtracted network analysis and
overall frequency (see Table 23). Only independent observations were included in the models.

Results

Research Question 1 Results

For RQ1 (Are there significant differences in the patterns of teacher support elicited by Inq-Blotter alerts versus Inq-Blotter alerts with TIPS for mathematical practices?) an epistemic network analysis (ENA) was conducted to compare the discursive supports provided by teachers in the two conditions (Inq-Blotter alerts versus Inq-Blotter alerts with TIPS). A significant difference was found between the types of supports provided by teachers with Inq-Blotter alerts with TIPS (n = 29 interactions; M = -.27, SD = .17; see right of Figure 23) relative to teachers with basic Inq-Blotter alerts (n = 16 interactions; M = .50, SD = .28; see left of Figure 23), t(44) = 10.05, p < .001, d = 3.60.

Figure 23. Epistemic network of support patterns for teachers with alerts (left) versus alerts with TIPS (right)
Figure 24. Subtracted network comparing teacher support in response to alerts (orange) versus alerts with TIPS (blue)

Qualitative analyses examining the subtracted network (i.e., reflecting the difference between the weights of connections in each network) further highlighted the differences in the support patterns (see Figure 24). Specifically, teachers using basic Inq-Blotter alerts provided more technical feedback (e.g., guidance on how to navigate the virtual meeting platform) in combination with mostly orienting (i.e., directing student attention) and evaluative supports (i.e., commenting on the accuracy of student work) relative to teachers using Inq-Blotter with TIPS. For example, Ms. A (a teacher at High School 1) received an alert that her student was having difficulty with Constructing Graphs and provided the following support to the student:

Ms. A: Uhhh could you **share your screen real quick**? [Technical Support]... **I want to see your graph** [Orienting Support; Recording ID A17]

In this interaction, the teacher requests to see the student’s screen to be able to look at the graph with the student, and correspondingly anchor support to the visualization. This
teacher later noted the need to have students share their screens “because in the dashboard it just tells me like ‘they're struggling with like data points’ or whatever but I don't know specifically like which part.” Therefore, the teacher required additional contextual support to be able to guide the student beyond just an orienting support.

On the other hand, teachers using Inq-Blotter alerts with TIPS provided relatively more orienting supports in combination with higher-level conceptual supports (i.e., explaining key components of the practice) and procedural supports (i.e., providing general guidance on the steps involved in engaging in the inquiry practice). This finding provides initial evidence of how TIPS were effective in eliciting higher-level supports (for which they were designed) that have been associated with student improvement in prior studies (Dickler et al., 2021a). Additionally, teachers in the TIPS condition did not provide any technical supports, which may indicate that the TIPS provided enough contextual information for educators to provide other inquiry supports without requiring students to share their screen. For instance, the following support was provided by Ms. X (the matched teacher to Ms. A from High School 1) after receiving an alert with TIPS regarding a student difficulty with Constructing Graphs:

Ms. X: In regards to your graph, for your axes, **which variable should you put along the x axis?** [Orienting Support] Or as the **independent variable**?…And the **dependent variable always goes along the y axis** [Procedural Support; Recording ID B1]

This interaction provides an example of how the teacher was able to direct the student’s attention to the specific difficulty with constructing graphs (i.e., selecting variables to place on each axis) as well as provide an overview of the steps that were necessary for...
the student to move forward with creating an appropriate graph. These results show the potential for TIPS to guide higher levels of support, and additional examples further demonstrate how teachers often used direct language from the higher-level TIPS when supporting their students (see Table 22). It is also valuable to understand the patterns of support that were associated with student improvement to understand the role of support combinations in promoting student learning, as is explored in the following section.

Table 22. Example of teacher supports in response to TIPS for Applying Equations

<table>
<thead>
<tr>
<th>TIPS Level</th>
<th>Teacher Support [Ms. Z; Recording ID B104]</th>
<th>TIPS Instructional Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orienting</td>
<td>T: Alright…make sure, take a look at your model, what's the trend in the data?...What is the relationship between the variables in your graph?... Let's look at your model. What is the trend in the data? What is the relationship between the variables in your graph?</td>
<td></td>
</tr>
<tr>
<td>Conceptual</td>
<td>T: So the goal is to choose a relationship or a mathematical equation that represents the trend in the data. So you can click on the relationship dropdown menu to see what each mathematical relationship looks like. Remember, the goal here is to choose a relationship (or mathematical equation) that represents the trend in the data. You can click on the &quot;Relationship&quot; dropdown menu to see what each mathematical relationship looks like.</td>
<td></td>
</tr>
<tr>
<td>Procedural</td>
<td>T: So look at the points on your graph, think about the trend in the data, is it curved? Is it a straight line? And then choose the mathematical relationship that best represents that trend. You should have at least 5 data points on your graph in order to view the trend. So make sure you only changed the one variable when you plot those data points. Look at the points on your graph, think about the trend in the data, and choose the mathematical relationship. Also make sure your graph axes are labeled correctly and that you have at least 5 data points from controlled trials to better see the trend.</td>
<td></td>
</tr>
</tbody>
</table>
**Research Question 2 Results**

Regarding RQ2 (Are there significant differences in the patterns of teacher support when students improved versus did not improve on mathematical practices?), an ENA was conducted to compare teacher support patterns associated with student improvement versus no improvement on mathematical practices (i.e., Constructing Graphs and Applying Equations). Only a subset of the interactions ($n = 33$ of 45 interactions) from the prior analyses were included as not all students completed an additional activity after being helped by the teacher and, as a result, improvement could not be computed for those students. A significant difference was found between support networks associated with improvement ($n = 25$ interactions; $M = -.16$, $SD = .45$; see left of Figure 25) versus no improvement ($n = 8$ interactions; $M = .50$, $SD = .26$; see right of Figure 25), $t(32) = 5.07$, $p < .001$, $d = 1.57$.

![Figure 25. Epistemic network of support patterns when students improved (left) versus did not improve (right)](image-url)
These results indicate that the types of supports teachers provided in response to alerts (with and without TIPS) were related to student learning outcomes on mathematical practices. Even teachers without TIPS provided combinations of support types associated with student improvement such as orienting with procedural support (see left of Figure 23 and Figure 25), but not to the extent of teachers in the alerts with TIPS condition (see Figure 24). The combinations of support associated with improvement are further illuminated within the subtracted network.

An examination of the subtracted network (see Figure 26) indicated strong connections between evaluative and other inquiry supports when students improved, versus stronger connections between instrumental and other inquiry supports when students did not improve. In particular, interactions where students improved involved more heavily weighted connections between evaluative support and higher-level procedural and conceptual supports. There was also a relatively strong connection between evaluative support and lower-level supports such as technical, content, and instrumental support when students improved. This finding demonstrates the importance
of using a dynamic network such as ENA to fully capture how the combination of higher-
level and lower-level supports could promote student improvement. For example, Ms. Z
(from High School 3) received an alert with TIPS that a student had placed a dependent
variable on the x-axis and gave the following supports to the student:

Ms. Z: When you're making your graph, think about **what should go on the**

**x axis** [Orienting Support]. So the **x axis should be your independent

**variable**. That's **what you're testing** [Conceptual Support], should **go on

your x axis** [Procedural Support]. …Make sure you set up your graph
correctly [Evaluative Support; Recording ID B42]

The student improved from prior to receiving this help (score of .25 out of 1 on
Constructing Graphs) to after receiving help (score of .75 out of 1 on Constructing
Graphs) and, in particular, successfully labeled their axes on their next activity. It is
likely that after focusing their attention on their axes, the student benefited from receiving
an explanation and instructions that they were able to apply to make correct graphs
moving forward.

For students who did not improve, the strongest relative connection was between
two lower-level supports: content and instrumental support. Content support is specific to
the scientific phenomenon being examined within an activity and instrumental support
often involves a bottom-out hint of the exact actions students must take to move forward
in the activity. As a result, these types of supports are unlikely to apply in future activities
that involve varying scientific content and may require differing procedures, thus,
transfer was not borne out.
Additionally, while instrumental support was sometimes provided in combination with orienting and procedural supports, students may have focused in on the exact instructions within the instrumental support that would not necessarily apply to future (i.e., if a student is told that they need to think about the trend in their data and that the trend is linear, then the student may select the relationship is linear without understanding why and may not understand how to interpret future mathematical relationships). For instance, the same teacher, Ms. Z, was using alerts with TIPS and also provided the following support for Constructing Graphs that did not result in improvement:

Ms. Z: [Student], x-axis should be the mass of the sled. Y axis should be the momentum. [Instrumental; Content; Recording ID B188]

The student who was helped had a decline in performance from prior (score of .75 out of 1 point on Constructing Graphs) to after being helped (score of 0 out of 1 point on Constructing Graphs). The next activity the student completed (i.e., examining how the roughness of the ramp impacted the acceleration of the sled at the end of the ramp) involved different variables and therefore the support provided previously was not applicable to the new investigation.

This epistemic network analysis enabled an understanding of the dynamic interaction between different support types associated with student improvement, but it is important to explore whether other discourse representations (e.g., frequencies) might better capture the relationship between support types and student performance.

Research Question 3 Results
In terms of RQ3 (What representation of discursive supports (i.e., frequencies versus weights of connections from ENA) best predict student improvement on mathematical practices?), it was first necessary to compute the average frequencies of the support types across interactions associated with student improvement or no improvement on mathematical practices (see Table 23). This analysis revealed that the most frequently provided supports when students improved (and most frequently provided supports overall across interactions) included conceptual and procedural supports. The most frequently occurring supports when students did not improve were procedural and instrumental supports. These findings align with the results from the ENA, but do not fully capture the connections between support types. As a result, linear regression models were built to predict the amount of student improvement using: 1) frequency of supports across interactions versus 2) the weights of connections between support types from the epistemic network analysis. Only a subset of the interactions ($n = 25$ interactions; see Table 24) were included in the regression analyses to account for the assumption of independence of observations.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement ($n = 25$)</td>
<td>0.96</td>
<td>1.24</td>
<td>1.16</td>
<td>0.80</td>
<td>0.96</td>
<td>0.60</td>
<td>0.75</td>
</tr>
<tr>
<td>No Improvement ($n = 8$)</td>
<td>0.53</td>
<td>0.50</td>
<td>0.88</td>
<td>0.88</td>
<td>0.00</td>
<td>0.38</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>0.79</strong></td>
<td><strong>0.87</strong></td>
<td><strong>1.02</strong></td>
<td><strong>0.84</strong></td>
<td><strong>0.48</strong></td>
<td><strong>0.49</strong></td>
<td><strong>0.66</strong></td>
</tr>
</tbody>
</table>

Table 24. Average frequency of supports when students improved versus did not improve (for subset of the data)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement (n = 19)</td>
<td>0.95</td>
<td>1.42</td>
<td>1.26</td>
<td>0.90</td>
<td>1.11</td>
<td>0.68</td>
<td>1.00</td>
</tr>
<tr>
<td>No Improvement (n = 6)</td>
<td>0.67</td>
<td>0.50</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.33</td>
<td>0.74</td>
</tr>
<tr>
<td>Overall</td>
<td>0.81</td>
<td>0.96</td>
<td>1.13</td>
<td>0.95</td>
<td>0.55</td>
<td>0.51</td>
<td>0.87</td>
</tr>
</tbody>
</table>


The linear regression model using the frequency of each support within interactions as the predictor variables was not significant, $F(7, 17) = 0.85, p = .564, R^2 = .26$. Accordingly, none of the variables were found to be significant predictors of student improvement (see Table 25). This indicates that the independent frequency of each support within interactions could not account for the amount of student improvement after being helped by a teacher using Inq-Blotter (with and without TIPS).

Table 25. Linear regression model coefficients (frequency to predict improvement)

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>B</th>
<th>t(13)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orienting Frequency</td>
<td>.09</td>
<td>.90</td>
<td>.382</td>
</tr>
<tr>
<td>Conceptual Frequency</td>
<td>.02</td>
<td>.23</td>
<td>.822</td>
</tr>
<tr>
<td>Procedural Frequency</td>
<td>-.15</td>
<td>-.85</td>
<td>.406</td>
</tr>
<tr>
<td>Instrumental Frequency</td>
<td>.00</td>
<td>.03</td>
<td>.977</td>
</tr>
<tr>
<td>Evaluative Frequency</td>
<td>.15</td>
<td>1.47</td>
<td>.160</td>
</tr>
<tr>
<td>Content Frequency</td>
<td>-.05</td>
<td>-.26</td>
<td>.797</td>
</tr>
<tr>
<td>Technical Frequency</td>
<td>-.05</td>
<td>-.80</td>
<td>.435</td>
</tr>
</tbody>
</table>
The linear regression model with the weight of connections based on the ENA as the predictor variables was found to be significant, $F(11, 13) = 2.88, p = .037, R^2 = .71$. This result shows that the weights of connections (specifically related to conceptual and procedural supports) were able to predict the amount of student improvement and explain 71% of the variance in the amount of improvement. In particular, the weight of the connection between conceptual and instrumental support was found to significantly, positively predict student performance (see Table 26). Therefore, increased higher-level conceptual support paired with lower-level instrumental support was likely to be associated with learning gains. Additionally, the weighted connection between procedural and evaluative support was found to significantly, positively predict amount of improvement (see Table 26). This result supports that increased higher-level procedural guidance with an evaluation of the accuracy of student work may benefit student learning, as indicated within the ENA.

Overall, these regression analyses provide initial evidence that the weighted connections from the ENA may better capture (relative to frequency counts) the relationship between dynamic teacher support and student performance on the practices on which they were helped. ENA allows for identifying how particular combinations of supports influence learning outcomes versus focusing on each support type independently. Further analyses are needed with additional participants to explore all possible combinations between types of support, but these initial results demonstrate the value of attending to combinations of discursive supports to predict learning outcomes.
Table 26. Linear regression model coefficients (ENA connection weight to predict improvement)

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>b</th>
<th>t(13)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual-Orienting Weight</td>
<td>-0.43</td>
<td>-0.87</td>
<td>.401</td>
</tr>
<tr>
<td>Conceptual-Procedural Weight</td>
<td>-0.34</td>
<td>-1.31</td>
<td>.214</td>
</tr>
<tr>
<td>Conceptual-Instrumental Weight</td>
<td><strong>2.37</strong></td>
<td><strong>3.29</strong></td>
<td><strong>.006</strong></td>
</tr>
<tr>
<td>Conceptual-Evaluative Weight</td>
<td>-0.29</td>
<td>-1.09</td>
<td>.295</td>
</tr>
<tr>
<td>Conceptual-Content Weight</td>
<td>-2.05</td>
<td>-2.07</td>
<td>.059</td>
</tr>
<tr>
<td>Conceptual-Technical Weight</td>
<td>1.38</td>
<td>1.17</td>
<td>.264</td>
</tr>
<tr>
<td>Procedural-Orienting Weight</td>
<td>-0.06</td>
<td>-0.21</td>
<td>.835</td>
</tr>
<tr>
<td>Procedural-Instrumental Weight</td>
<td>-0.96</td>
<td>-1.98</td>
<td>.069</td>
</tr>
<tr>
<td>Procedural-Evaluative Weight</td>
<td><strong>1.42</strong></td>
<td><strong>2.28</strong></td>
<td><strong>.040</strong></td>
</tr>
<tr>
<td>Procedural-Content Weight</td>
<td>-0.18</td>
<td>-0.21</td>
<td>.839</td>
</tr>
<tr>
<td>Procedural-Technical Weight</td>
<td>-1.14</td>
<td>-1.55</td>
<td>.146</td>
</tr>
</tbody>
</table>

Discussion

Teachers’ feedback can help promote student learning on complex inquiry practices (Manz & Renga, 2017; McNeill & Krajcik, 2008; Talanquer et al., 2013) and technologies such as dashboards can help to direct feedback in meaningful ways (e.g., Knoop-van Campen & Molenaar, 2020). Dashboard tools have been shown to help teachers identify struggling students and provide corresponding discursive support (Holstein et al., 2018a; Knoop-van Campen & Molenaar, 2020; Martinez-Maldonado et al., 2013; Tissenbaum & Slotta, 2019), but it is important to examine the types of support elicited by different types of dashboard alerts (i.e., with or without instructional supports) and the relationship to student learning outcomes. The present study provided initial evidence of how detailed TIPS embedded within Inq-Blotter alerts could foster higher level support from educators remotely as students completed virtual inquiry investigations in Inq-ITS.
A major contribution of the present study is the evidence of how the TIPS within Inq-Blotter alerts were successful in guiding teachers to provide higher-level supports to students on mathematical practices. Specifically, the results of the ENA comparing teacher support with Inq-Blotter alerts versus Inq-Blotter alerts with TIPS indicated that teachers using basic Inq-Blotter alerts (without TIPS) provided primarily technical support in combination with orienting students toward the practice. Teachers using Inq-Blotter alerts with TIPS, however, provided higher-level conceptual and procedural supports in combination with orienting students’ attention. These findings indicate how the TIPS instructional supports could influence the ways in which teachers support their students. Future work is therefore needed to explore why this higher-level support elicited by alerts with TIPS did not promote greater learning gains relative to basic alerts without TIPS (see Study 2).

An ENA comparing the patterns of support when students improved or did not improve also revealed a significant difference in the types of supports associated with student learning. In particular, teachers provided primarily lower-level, context specific supports in situations where students did not improve on mathematical practices on their next opportunity. Students who did improve their performance received a number of supports in combination with evaluative feedback from the teacher. Future work should attend to why lower-level contextual supports were still provided to some students and the decision-making processes behind support choices. For instance, results from exploratory analyses in earlier studies (see Study 1) provide evidence that other contextual information related to students’ prior knowledge may inform the types of supports that teachers provide to students (i.e., teachers tended to provide more
instrumental and content supports to students with higher prior knowledge on mathematical practices).

Finally, the present study compared the effectiveness of discourse representations (i.e., weighted connections from the ENA relative to frequency counts) in predicting the amount of student improvement on inquiry practices and found that the ENA-based model could significantly predict student performance. This finding demonstrates how frequency counts alone may not sufficiently capture the dynamics of discursive supports that promote student learning (i.e., how particular supports used in combination are successful at guiding students). ENA provides an opportunity to more fully capture how teachers distribute supports and the corresponding relationship to student learning and performance.

In terms of future studies, it will be important to recruit a greater number of participating teachers and students to continue to compare support patterns across conditions. The present study was conducted with schools in the northeastern United States with one-to-one technology access, so additional work is needed to understand how these tools might be used in schools with differing access to technology. Additionally, it is critical to work with participating schools reflecting diverse demographics in order to ensure that the development of tools meet the needs of teachers across contexts. While the schools in the present study represented diverse demographics, the number of participating students from each school varied in each phase. Therefore, caution should be taken when generalizing the results of the present study across teachers and students (especially given the variations in how remote learning has been implemented across schools in response to the COVID-19 pandemic; Arnett, 2021), but
the results are still important for understanding how the design of dashboard features can impact technology use.

Future studies will also need to attend to how Inq-Blotter alerts with and without TIPS are implemented within in-person classroom contexts. In-person classroom contexts enable teachers to have direct access to student materials (i.e., teachers can look over a student’s shoulder to see the student’s work), so the information provided within TIPS may be used differently. It will also be valuable to attend further to how teachers determine the types of supports that are provided to each student and which student alerts to address. This can be done through interviewing teachers about their decision-making processes during or after data collection, as well as examining the relationship between student prior knowledge on an external measure and supports received (see Study 1).

The in-depth analysis of teacher supports elicited by Inq-Blotter alerts without and with TIPS builds on prior work by demonstrating ways in which design elements in the dashboard impact teacher discursive guidance on inquiry practice competencies. This work also provides evidence of the value of applying analytics to discourse, such as ENA, to better understand how discursive patterns relate to performance outcomes. These findings have important implications for the development of educational technology both in terms of demonstrating the potential of a dashboard to guide teacher support remotely, but also the methodologies that can be implemented to examine the use of innovative STEM technologies.
Conclusion

Overall, the findings from the present dissertation demonstrate the potential for innovative technologies to help prepare students in the United States with the competencies needed for success in STEM. In particular, difficulties with using mathematics in science can be a barrier to STEM attainment (Gottfried & Bozick, 2016; NGSS, 2013) so it is essential that these difficulties are identified and supported. Teachers, however, require supports to help with assessing and scaffolding students on these difficulties, particularly in the context of remote learning. The present design-based research dissertation provides evidence of how innovative technologies with the capacity for real-time alerting can meet these challenges in STEM education and informs theory to guide the future development of educational tools.

Phase 1 Conjectures

Phase 1 (Study 1) of the present dissertation examined the implementation of the real-time alerting dashboard, Inq-Blotter, in remote high school STEM classrooms while students completed investigations in Inq-ITS. The results of Study 1 provided evidence for the high-level conjecture that real-time alerting dashboards could address teacher challenges with supporting students in STEM remotely including assessing student performance, identifying struggling students, determining students’ difficulties, and supporting students on Using Mathematics in science inquiry (see Figure 27).

The embodiment of additional stages in Inq-ITS with tools for mathematical practices was successful in logging and automatically scoring student performance on the practices of Constructing Graphs and Applying Equations at the fine-grained, sub-practice level. The Theoretical Conjecture was also supported that the interpretation of
scores on these stages was a valid assessment of student competencies for using mathematics as defined by other assessments (College Board, 2019a, 2019b; ETS, 2013, 2017; Lai et al., 2016; Liu et al, 2016). This embodiment was critical both for supporting teachers’ remote assessment of student competencies overall as well as triggering real-time alerts in the Inq-Blotter dashboard.

Additional embodiments tested as part of Phase 1 included real-time alerts with fine-grained, actionable information on which students needed help and their specific difficulties on mathematical practices (i.e., Constructing Graphs and Applying Equations). Results supported the design conjectures that real-time alerts help teachers to identify struggling students and that the details within alerts help teachers to determine students’ specific difficulties in order to then provide discursive support. The support provided by teachers, however, was relatively low-level (i.e., discourse consisted of primarily orienting and technical supports) and teachers experienced some challenges related to supporting students remotely (e.g., lacking contextual information). Even with these challenges, the support provided by teachers was associated with student improvement on mathematical practices for the majority of students, which aligns with the theoretical conjecture.

These findings provide support for how dashboards with real-time alerts could support remote instruction of critical practices in STEM, but indicated that there was still a need to further iterate on alerts to address new challenges that were not anticipated within the remote context. In particular, it was important to consider whether additional information could be integrated within alerts to help activate teachers’ PCKI (Yang et al., 2018) specifically related to supporting Using Mathematics and correspondingly promote
higher-level support to students on inquiry practices. The development and integration of the alert design iteration with instructional supports was tested as part of Phase 2 of the dissertation.

### Phase 1 Conjecture Map Results

#### High Level Conjecture
A student environment and corresponding dashboard with real-time alerts containing data on student performance on math practices in an authentic inquiry environment addressed challenges experienced by STEM educators when teaching remotely:

- Assessing students on inquiry practices such as Using Mathematics in science
- Identifying students struggling with math practices in science inquiry
- Determining student difficulties with math practices in science inquiry
- Supporting students on math practices in science inquiry

#### Design Conjecture

**Embodiments**

- Inquiry practice stages that capture Using Mathematics in science (i.e., constructing graphs and applying equations)
- Real-time alerts that identify students who are struggling and the math practice on which students are struggling
- Details within alerts regarding students’ specific difficulties with Constructing Graphs and Applying Equations

**Mediating Processes**

- Students’ actions were logged and automatically scored on math practice stages in Inq-ITS
- Teacher identified students who are struggling
- Teacher provided discursive support to students on the specific math practice of difficulty (but support was relatively low-level and teachers experienced some challenges remotely)

#### Theoretical Conjecture

**Outcomes**

- Valid assessment of students’ mathematical competencies (i.e., scores in Inq-ITS correlated with pre-test scores)
- Students improved on the math practice on which they were helped on their next opportunity

---

Figure 27. Phase 1 Conjecture Map Results

### Phase 2 Conjectures

In Phase 2 (Study 2 and Study 3) of the dissertation, Teacher Inquiry Practice Supports (TIPS) for mathematical practices were developed and embedded within Inq-Blotter to further guide teacher support on mathematical practices. Study 2 and Study 3 contributed evidence towards the high-level conjecture that additional information within alerts (e.g., TIPS) could address some of the challenges that were found for teachers in Study 1. Study 2 examined the embodiment of alerts with TIPS (relative to basic alerts...
without TIPS) and the design conjecture was supported that these instructional supports helped teachers to identify and support struggling students. Study 3 provided further evidence that the embodiment of TIPS elicited discourse with higher levels of support on student difficulties as demonstrated using ENA. In terms of the corresponding theoretical conjectures, Study 2 provided evidence that both Inq-Blotter alerts and Inq-Blotter alerts with TIPS were found to foster support associated with student improvement on practices, but there were no significant differences in the amount of improvement experienced by students in each condition. Therefore, the results of Phase 2 did not support the theoretical conjecture that alerts with TIPS (that elicited higher-level teacher support) would also promote greater student learning outcomes.

### Phase 2 Conjecture Map Results

**High Level Conjecture**

A student environment and corresponding dashboard with real-time alerts containing data on student performance on math practices in an authentic inquiry environment as well as instructional supports addressed challenges experienced by STEM educators when teaching remotely:

- Assessing students on inquiry practices such as Using Mathematics in science
- Identifying students struggling with math practices in science inquiry
- Determining student difficulties with math practices in science inquiry
- Supporting students on math practices in science inquiry at a high level
- Supporting teachers’ PCKI for using mathematics in science

**Design Conjecture**

- **Embodiments**
  - Real-time alerts that identify students who are struggling and the math practice on which students are struggling
  - Details within alerts on students’ specific difficulties with Constructing Graphs and Applying Equations
  - Instructional supports (i.e., TIPS) within alerts to guide teacher scaffolding of inquiry difficulties at multiple levels (e.g., orienting, conceptual, procedural, and instrumental) and support PCKI

- **Mediating Processes**
  - Teacher identified students who were struggling
  - Teacher provided discursive support to students on the specific math practice of difficulty
  - Teacher provided higher levels of discursive support to students with TIPS (e.g., conceptual and procedural support)

**Theoretical Conjecture**

- **Outcomes**
  - Students improved on the math practice on which they were helped on their next opportunity
  - There were no significant differences in student improvement when helped by a teacher using alerts with TIPS versus without TIPS

---

*Figure 28. Phase 2 Conjecture Map Results*
Further work is needed to better understand why the higher-level support from teachers did not significantly impact student performance on the mathematical inquiry practices of difficulty. It is possible that certain contextual variables (e.g., students receiving support remotely through their computer versus face-to-face during in-person learning) may have impacted the efficacy of the teachers’ feedback because students may not have fully attended to teacher support (i.e., staying engaged in an online setting requires self-regulation on the part of the student, which can be difficult to maintain; Archambault, 2010; Whipp & Chiarelli, 2004). It is important to review potential limitations and determine areas of future research to better situate the implications of the present work.

**Limitations**

There are limitations to the present dissertation including first the sample size. With only six teachers participating in total, caution must be taken when generalizing the results from the present studies. Additionally, all participants were recruited from three school districts in the northeastern United States. While there was diversity within the socio-demographics of each participating district, the sample was not representative of the overall population. The in-depth examination of the data, however, does still provide insight into dashboard use remotely and potential needs experienced by teachers.

Additionally, the length of data collection was restricted to one class period for each participating class section and students only completed one lab set. As a result, it is unclear from the present data whether student learning gains might transfer over time and across lab topics. Prior studies at the middle school level have shown transfer for other
inquiry practices with scaffolded support from Rex (Li et al., 2019), but additional work is needed to understand how student performance transfers for mathematical practices on the high school level activities (especially given the interaction between the difficulty of mathematical relationships and performance on applying equations that was found in Study 2) and in relation to teacher scaffolding.

The present work focused primarily on the relationship between teacher support and students’ corresponding performance in relation to individual alerts. There were some situations where multiple students were supported within the same interaction (based on each of their individual alerts) or where one student was assisted multiple times. Additional analyses would be needed to understand the impact of these variations in delivery of support. Studies could also examine the impacts of whole-class level alerts in greater detail as well as other alert types such as “slow progress” alerts. These studies could then inform the development of TIPS for these types of alerts.

Given the remote nature of the study, there are certain limitations to the data including factors that may have had a negative impact on student performance. For example, it was unclear if students were attending to the support provided by the teachers (i.e., students did not always have their cameras on or reply to teacher feedback) and technical difficulties sometimes interfered with teacher support (i.e., the student could not determine how to share their screen or the internet was slow). Other remote factors may have unintentionally benefited student performance including that students could overhear the help provided by the teacher to other students when in the main session of the virtual meeting and students may have received help from outside sources while working remotely (i.e., a parent or through a source on the internet). These limitations
could apply across remote contexts, but are important to note when thinking about applying the findings of this dissertation to in-person settings.

**Future Directions**

In addition to future directions noted previously including collecting data with larger samples representing diverse demographics, extending the length of data collection, exploring the effects of other alert types, and conducting studies in in-person settings, future studies should iterate on the design of TIPS for mathematical practices for which students did not experience improvement due to the complex interaction with mathematical content (e.g., understandings related to negative linear relationships). In particular, students across conditions experienced a decline in performance for the practice of Applying Equations indicating that additional support was needed specific to the application of a negative linear relationship. TIPS can be updated to encourage teachers to explain the role of equation components across different types of relationships, even if only one type of relationship is being examined at a point in time (i.e., when supporting a student on a positive linear relationship, it would be valuable to provide a more detailed explanation of how the coefficient can also impact the direction of the line in addition to steepness). While TIPS do not currently address content understandings, a new level of support specific to content could be added or external supplementary documents/resources could be constructed to better support teachers.

**Overall Conclusions**

In conclusion, the present work illustrates the potential for dashboard technologies to guide teachers in supporting struggling students on critical practices in remote STEM settings. Triangulating log file data with nuanced discourse data resulted in
rich, deep, and rigorous evidence indicating how these technologies can shape remote STEM instruction, while also informing the needs for future design iterations. Overall, this work addressed a crucial issue in STEM education in terms of analyzing how innovative, data-informed technologies can be used to support students on core competencies in STEM in remote classroom settings.
Appendix

Pre-Test for Mathematical Practices

**Question 1**
You want to test whether the speed of a ball changes when it is dropped from different heights. Which of the following trials are appropriate to test this phenomenon:

a) Trials 1, 2, and 3  
b) Trials 2, 3, and 4  
c) Trials 4 and 5  
d) Trials 1 and 5

<table>
<thead>
<tr>
<th>Trial Number</th>
<th>Mass of Ball (kg)</th>
<th>Height of Drop (m)</th>
<th>Speed of Ball (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>10.10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>30</td>
<td>24.21</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>50</td>
<td>31.26</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>70</td>
<td>40.41</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>70</td>
<td>45.46</td>
</tr>
</tbody>
</table>

**Question 2**
Use the graph to the right for the following question. When there are three bars of gold, what is the force of gravity on the gold?

a) 0.5 Newtons  
b) 2 Newtons  
c) 4.5 Newtons  
d) 7 Newtons

**Question 3**
If scientists want to create a graph to examine whether an increase in body temperature causes an increase in heart rate, then how should they label the x-axis and y-axis of their graph?

a) x-axis = Body Temperature; y-axis = Exercise  
b) x-axis = Exercise; y-axis = Body Temperature  
c) x-axis = Heart Rate; y-axis = Body Temperature  
d) x-axis = Body Temperature; y-axis = Heart Rate
**Question 4**

Scientists want to examine how the depth of a diver in water affects the pressure experienced by the diver. Based on this investigation, which of the following graphs has the axes labeled correctly?

a) ![Graph a)](image)
b) ![Graph b)](image)
c) ![Graph c)](image)
d) ![Graph d)](image)

**Question 5**

Which graph below best represents the data in the table?

<table>
<thead>
<tr>
<th>Trial Number</th>
<th>Height of Drop (m)</th>
<th>Speed of Ball (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>14.01</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>24.21</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>31.26</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>40.41</td>
</tr>
</tbody>
</table>

a) ![Graph a)](image)
b) ![Graph b)](image)
c) ![Graph c)](image)
d) ![Graph d)](image)

**Question 6**

Which graph below best represents an inverse relationship between time and temperature?

a) ![Graph a)](image)
b) ![Graph b)](image)
c) ![Graph c)](image)
d) ![Graph d)](image)
**QUESTION 7**

The equation to the right represents the relationship between the percent of fat (F) and carbohydrates (C) in cereal. Which of the following statements about the equation are true?

- a) Statement 1 only
- b) Statements 1 and 2
- c) Statements 2 and 3
- d) Statements 1 and 3

\[ C = -1.02F + 93.63 \]

1. The constant in the equation is negative
2. The equation represents a linear relationship
3. The slope of the equation is -1.02

**QUESTION 8**

Which equation below is the best model of the data in the table?

- a) \( y = 15x + 1 \)
- b) \( y = \frac{15}{x} \)
- c) \( y = 15x^2 \)
- d) \( y = 15 + x \)

<table>
<thead>
<tr>
<th>X (Time)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y (Speed of Train)</td>
<td>15</td>
<td>7.5</td>
<td>5</td>
<td>3.75</td>
</tr>
</tbody>
</table>
References


Jong (Eds), *Cyber-physical laboratories in engineering and science education* (pp. 191-217). Switzerland: Springer.


