Exploring the Impact of Small-group Synchronous Discourse Sessions in Online Math Learning

Jinnie Choi and Alyssa Walters
Pearson

Abstract

Students in a fully online learning environment have limited opportunities to talk about math problem solving. While discourse is an effective pedagogical model, less is known about how it translates to online math learning. We analyzed online platform data from two fully online virtual elementary schools in the United States to address the following research questions: (a) Is participation in math discourse about reasoning and problem solving related to students' math confidence, self-efficacy toward math, and math mindset in an online learning environment? (b) Is participation in math discourse related to math performance? The study design involved school-level voluntary student participation and a retrospective analysis of online platform data. The results showed that repeated participation in discourse sessions in a course was not directly related to changes in math confidence, self-efficacy toward math, and math mindset. However, student participation in higher numbers of sessions was associated with higher final math course score and math state assessment performance, after controlling for prior performance and confidence, self-efficacy, and mindset scores. This study has important implications for the future design and implementation of online math instruction as well as the training and development of fully online math teachers.

Keywords: math discourse, mathematics instruction, online learning

Exploring the Impact of Small-group Synchronous Discourse Sessions in Online Math Learning

K-12 virtual schools in the United States serve a highly mobile student population (Gatti, 2018), and mobility (i.e., changing schools for reasons other than grade promotion) has a consistent and severe negative impact on math performance (Rumberger, 2015). Indeed, studies have shown low average state assessment scores in math (Woodworth, Raymond, Chirbas, Gonzalez, Negassi, Snow, & Van Donge, 2015; Ahn, 2016). However, after adjusting for district-mean student mobility and other demographic factors, a recent study found no statistical difference in the percentage of students scoring proficient in math between student cohorts in fully online virtual schools and cohorts in brick-and-mortar schools that were matched on prior achievement (Gatti, 2018). This work highlights the importance of finding effective methods to remediate the negative

effects of high mobility on math performance. However, research shows a lack of rigorous studies regarding the practices of successful school-level strategies to improve math outcomes of virtual school students (Choi, Belenky, DiCerbo, Lai, & Wardlow, 2016). How can we remediate the negative effects of high mobility by special interventions designed to help support math learning?

Fully online learning environments provide an experience of learning math that differs substantially from traditional brick-and-mortar classrooms. One important distinction is that these students have limited access to opportunities to talk about math problem solving, given that virtual instruction incorporates a large amount of asynchronous and self-paced lessons and activities. The impact of reduced talking in an online environment may also include decreased social and constructive math practices, decreased peer-to-peer teamwork and collaborative problem solving, and decreased exposure to positive and supportive feedback from others.

Arguably, promoting a deep understanding of complex mathematical concepts and robust problem-solving skills is not a simple task to achieve in any learning environment. Among the interdisciplinary efforts to address this problem, a body of research (Yackel & Cobb, 1996; Resnick, Bill, & Lesgold, 1992; Lehrer & Schauble, 2005; Lampert, Rittenhouse, & Crumbaugh, 1996; Boaler, 1999) emphasizes the role of constructivist and sociocultural principles that build on the importance of social practices and, in particular, the careful orchestration of talk and tasks in academic learning (Michaels, O'Connor, & Resnick, 2008). The core idea includes the discourse experience for students to explicate their reasoning, provide warranted evidence for their claims, and make their voices heard and recognized as offering reasonable contributions to further the understanding and development of math concepts and problem-solving skills. While it is an effective pedagogical model, less is known about how the discourse experience within traditional learning environments translates to online math learning. In particular, research is needed to understand how we can best support the experience of learning math in fully online learning environments by incorporating time to talk about math.

Research Questions

In this study, we explored an intervention for fully online virtual school students where students have opportunities to engage in meaningful discussions about math problem solving through small-group synchronous discourse sessions. Research suggests that carefully designed math discourse interventions can improve students' math learning outcomes such as problem-solving skills. Also, research supports the idea that self-efficacy, self-confidence, and a growth-fixed mindset are related to math learning. With those findings in mind, we designed and implemented small-group synchronous discourse sessions and examined how participation relates to online math learning by analyzing online platform assessment data from two virtual elementary schools. This study addressed the following research questions. In an online learning environment, (a) is participation in math discourse about reasoning and problem-solving related to students' math confidence, self-efficacy toward math, and math mindset in online learning environment? (b) Is participation in math discourse about reasoning and problem solving related to math performance?

Theoretical Framework

Discourse and Math Performance. In the classroom, students are not only expected to master a body of authoritative conceptual knowledge but also to demonstrate reasoning and the use of procedural tools. Evidence from mathematics and science education literature suggests that sense-making and scaffolded discussion in classrooms can promote this deeper understanding of

complex concepts and robust reasoning. Specifically, discourse increases student conceptual understanding, develops precision with academic vocabulary, and increases student ability to justify and reason in math problem solving. (Michaels et al., 2008; Yackel & Cobb, 1996; Resnick et al., 1992; Lehrer & Schauble, 2005; Lampert et al., 1996).

Several researchers described strategies for designing and implementing math discourse in classroom teaching. For instance, Mercer and Sams (2006) evaluated whether a program that provides elementary school students with guidance and practice in the use of language for reasoning would enable them to use language more effectively as a tool for collaborative math problem solving. They found evidence that improving the quality of students' use of language for collaborative reasoning improves their individual learning and understanding of mathematics. In emphasizing the teacher's role in modeling and guiding students' use of language for reasoning, they recommended that teacher-student discourse should be designed and implemented with a careful combination of peer group interaction and expert guidance and should involve an explicit and practical introduction to the use of language for collaborative reasoning for better learning outcomes.

In her influential article on mathematics education, Lampert (1990) described an approach for teachers and students to talk about mathematical problems and solutions. The focus of the approach is to bring the practice of math teaching and learning in school closer to the real disciplinary knowledge of mathematics. Specifically, the teacher describes the problem and then initiates and supports social interactions appropriate to making mathematical arguments in response to students' conjectures. Students are responsible for figuring out how to solve the problem as well as finding the solution. The content of the lesson is arguing for the legitimacy of the solution strategies, rather than the finding of answers. "Students' strategies yield answers to teachers' questions, but the solution is more than the answer, just as the problem is more than the question" (Lampert, 1990, p. 40).

In the current study, we designed the math discourse intervention for fully online virtual school students with these principles in mind. The roles of session facilitators and students were carefully defined to support the practice of mathematical problem solving by training facilitators on a framework for facilitating meaningful math discourse described by the National Council of Teachers of Mathematics (2014). Specifically, facilitators were asked to engage students in purposeful sharing of mathematical ideas using varied representations, select and sequence student responses to create group analysis and discussion, encourage students to explain and defend to one another their solution strategies, and make explicit connections within the discussion to ensure progress toward goals. Students, in turn, were expected both to present ideas to others and to hear, seek understanding about, and critique the ideas of others. Finally, students were expected to compare and contrast the various approaches to the problems being solved.

Confidence, Self-efficacy, and Growth/Fixed Mindset. Among the constructs that we examined as outcomes and covariates in this study were self-efficacy, self-confidence, and growth-fixed mindset, given the large body of research that supports their relation with math learning. Self-efficacy is defined as "the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations" (Bandura, 1995). Research has shown that self-efficacy beliefs affect a variety of behaviors known to predict learning and academic achievement (Schunk & Pajares, 2002), and specifically, self-efficacy has been found to relate to higher performance in online math courses (Kim, Park, & Cozart, 2014). Similarly, self-confidence is a belief about the extent to which one is good at something (Zimmerman, 2000). Low self-

confidence in math has been linked to math anxiety, a sense of not belonging, and discomfort with math learning (Ashcraft, 2002). In this study, we hypothesized that the math discourse intervention would have a positive impact on students' self-confidence and self-efficacy in math by explicitly encouraging the effort and process of problem solving rather than asking students to simply provide correct answers.

A significant body of research has also identified benefits associated with adopting a "growth mindset," or belief that intelligence/math ability can change (Dweck, 2006; Dweck, 2017). For example, students with a growth mindset are more likely to adopt more learning-oriented goals, to persist longer (Diener & Dweck, 1980), to use better learning strategies, and ultimately, to achieve better grades (Yeager & Dweck, 2012). To change mindset from fixed to growth, Dweck (2006, 2017) suggests a few types of interventions including (a) discussing the growth mindset directly in lessons and instructions , (b) presenting analogies between math and other disciplines such as sports and music (e.g., "You have to practice to get better," (c) providing growth-oriented feedback: rather than focusing on correct and incorrect student responses and (d) having students reflect on their own progress and growth. In this study, we asked facilitators of the math discourse sessions to integrate growth mindset interventions in the discussions and hypothesized that students' growth mindset would increase alongside performance. Thus, facilitators of the discourse sessions received mindset training prior to the implementation of discourse sessions so that they could promote students' growth mindset.

Gaps in the Literature. K-12 students in fully online virtual schools have two distinct characteristics that have a negative impact on math performance: high mobility and reduced opportunities to talk about math problem solving. There is limited research on whether providing discourse experiences for students has a positive impact on online math learning or whether, and how, participating in discourse has a positive impact on other constructs that are known to be related to math learning. Also, online instruction for upper elementary age students is rarely addressed in the research literature. In this study, we implemented math discourse in two virtual elementary schools to explore the impact of discourse on online math learning.

Methods

Participants

During the 2016-17 academic year, the study was conducted at two fully online virtual schools, one that serves students in a Midwestern state and the other that serves students in a Western state in the United States. These schools, similar in terms of overall performance level and years in operation, voluntarily participated in the pilot program of the math discourse intervention. Across the schools, a total of 898 students in grades 3, 4, and 5 from 10 classrooms participated in the pilot. The discourse sessions were offered in two phases: the first during the Fall of 2016 (semester A) and the second during the Spring of 2017 (semester B). In total, the number of students who participated in at least one discourse session in semester A included 789 students: 228 (3rd grade); 271 (4th grade); and 290 (5th grade). In semester B, the participants included 719 students: 199 (3rd grade); 254 (4th grade); and 266 (5th grade). Among all the students, 20% participated in at least one session in semester A only (53, 3rd grade; 55, 4th grade; and 71, 5th grade), 67.9% participated in at least one session in semester A and at least one session in semester B (175, 3rd grade; 216, 4th grade; and 219, 5th grade), and 12.1% participated in at least one session in semester B only (24, 3rd grade; 38, 4th grade; 47, 5th grade).

Students varied in terms of how many discourse sessions they participated in per semester and whether they participated in a single or multiple semesters. Based on the mode of the distribution, we created two groups for later analyses: a high-participation group that attended six or more sessions in a semester and a low-participation group that attended five or fewer sessions. Six (sessions) was chosen, because it was the global mean of the number of participated sessions across all courses, grade levels, and semesters. In semester A, 155 students (68% of participating 3rd graders), 214 students (79% of participating 4th graders) and 205 students (71% of participating 5th graders) attended six or more sessions. In semester B, 149 students (75% of participating 3rd graders), 185 students (73% of participating 4th graders), and 178 students (66.9% of participating 5th graders) attended six or more sessions. We separated the semesters in the analyses, because the courses in different semesters were graded independently from each other.

Study Design

The study design involved school-level voluntary participation and a retrospective analysis of online platform data. Student participation in the discourse sessions was voluntary but strongly recommended at the classroom and school levels. Teachers used a sample welcome letter and webmail communication to encourage students to participate. Also, students received temporary zeros in their grade book for non-participation, which had no consequence for the final grade. Participation was tracked in terms of three variables: 1) number of sessions participated in per semester, 2) high (i.e., attended six or more sessions in a semester) versus low participation, and 3) semester participation pattern (A only, B only, or A and B). The details about these variables are described in a later section.

Implementation of Discourse Sessions

The discourse sessions consisted of synchronous, small-group online sessions with verbal and visual communication. The student-facilitator ratio ranged from 1:1 to 10:1. These synchronous sessions were embedded in the math courses that are normally asynchronous with flexible schedules.

The sessions occurred approximately once every seven lessons during a 16- to 17-week semester. The queue was open during normal school hours on weekdays. Each session lasted 30 minutes as a small-group discussion format. Students accessed the sessions through a link in their course homepage for each designated lesson. The link brought them to a queue where the math specialist (a facilitator) transferred them to a synchronous live lesson room. The facilitators were not known to the students prior to the experience. Each week, students encountered new math problems to discuss and the facilitator moderated the difficulty of the problems as needed.

In total, students were given opportunities to participate in 9 to 11 discourse sessions per semester (the number varied, depending on the grade level and the course). At the end of each discourse session, participating students were asked to complete reflection assessments that contained items designed to measure their math confidence, self-efficacy towards math, and growth/fixed mindset. These assessments were automatically scored and did not count towards the final course grade.

Roles of Session Facilitator and Participants

Each of the sessions was facilitated by one of eight math subject experts who held a degree in mathematics. Prior to implementation, facilitators received formal training about their roles,

which included (a) presenting the math problem, (b) guiding students to focus on the process rather than arriving at the solution and a variety of ways to approach the problem, , (c) encouraging students to talk to one another about their thought processes, and (d) giving feedback to students that promotes a growth mindset. Facilitators encouraged participants to engage in specific actions including communicating interactively with each other about mathematical reasoning and problem-solving using screen sharing, explaining and justifying, listening carefully, seeking understanding, asking questions that clarify, and comparing different approaches to the same problem.

Instruments and Data Sources

In this study, we collected data to explore how participation in math discourse relates to changes in mindset, math self-views and math performance, when other variables such as prior year math performance scores and different participation patterns are considered. Data were collected from participation logs on the platform, automated scores from reflection assessments, and end-of-course scores and state assessment results that were matched by non-identifiable web user IDs.

Confidence, self-efficacy, and growth / fixed mindset. We developed the reflection assessments following the discourse sessions to measure confidence, self-efficacy, and growth/fixed mindset (see Figure 1 for the actual questions). Each question included four items that were scored yes or no. We used a collection of items from a previous study targeted to measure the same constructs (Choi, Walters, & Hoge, 2017). The reflection assessments were administered every three sessions immediately after the discourse to capture potential changes over time. Considering the select-all options as "items," all four confidence and self-efficacy items were scored 1 or 0 and the composite score at a given session was calculated by summing up all four scores, resulting in the range of 0 to 4. Two growth mindset items were scored 1 or 0, and two fixed mindset items were scored 0 or -1. The composite score at a given session was also calculated by summing up all four scores resulting in the score range of -2 (fixed) to 2 (growth). Across all participants at all measurement occasions, the confidence and self-efficacy score was on average 2.64 (SD = 1.30) for semester A and 2.68 (SD = 1.33) for semester B. The mindset score was on average 0.95 (SD = 0.844) for semester A and 0.84 (SD = 0.893) for semester B. Cronbach's alpha was 0.45 (95% CI = [0.39, 0.51]) for the confidence, self-efficacy items and 0.40 (95% CI = [0.23, 0.57]) for the mindset items (when -1 for the fixed items was considered as a correct answer).

A. Confidence and self-efficacy item
Select all that you believe to be true.
1. I like math.
2. I feel like I belong in a math classroom.
3. I am good at math.
4. I am sure that I can solve the problems that my math teacher gives me.

B. Growth Mindset question
Select all that you believe to be true.
1. I think anyone can be good at math.
2. I can always get better at math if I try
3. I think some people are just better at math than others.
4. I can't do a lot to change how good I am at math.

Figure 1. Reflection questions after each discourse session.

Math performance measures. We collected evidence of math performance from multiple sources. Final course grade for the 2016-2017 school year represented one of the main outcome variables for this study. These end-of-course grades were measured on the scale of 0 to 100 percent. Across all participants, the final course score was on average 85.3 (SD = 14.4) for semester A and 79.8 (SD = 19.0) for semester B.

In addition, we collected the state assessment results at the end of the 2016-2017 school year using the following four categories: "Advanced," "Proficient," "Basic Proficiency," and "Below Basic Proficiency." The variable was dummy coded for a logistic regression to have 1 if "Proficient" or "Advanced" and 0 if "Basic Proficiency" or "Below Basic Proficiency." Across all participants, the average was 0.318 (SD = 0.466) for semester A and 0.295 (SD = 0.456) for semester B.

Participation variables. Individual student participation in each discourse session was recorded on the online platform as "yes" (did participate) or "no" (did not participate). We counted the number of sessions a student participated in a semester (*numPart*). The mean number of participated sessions in a semester was 5.12 (SD = 3.91), with a minimum of 0 and a maximum of 11 sessions. We then used the information to distinguish two groups of students: a high-participation group that participated in more than six sessions in a semester and the other group that did not (*highPart*). The cutoff of six sessions was determined based on the global median of the distribution of the number of participated sessions across grade levels and courses. Across the two semesters, 53.4% of students participated in more than six sessions. We also created another variable that indicated the semester(s) in which the student participated in at least one discourse session (*semPart*). In total, the number of students who participated in at least one discourse session was 789 students for semester A and 719 students for semester B. Among them, 179 students participated in semester A only (20%), 610 students participated in semester A and B (67.9%), and 109 students participated in semester B only (12.1%).

Other background variables. To explain additional variance in math performance, we utilized the repeated measures of math confidence, self-efficacy towards math, and math mindset to calculate average scores for each student in each semester (*avgCS*, *avgMD*). The mean of avgCS was 2.614 (SD = 1.16) on the 0 to 4 scale, and the mean of avgMD was 0.893 (SD = 0.734) on the -2 to 2 scale. When available, we collected the final course scores for the students who also attended the school the previous school year (2015-2016) to further explain the variation in the main performance outcome (*score1516*). The mean was 82.45 (SD = 13.12). Also, we collected basic background information such as school locations (75.7% are from location 1, and 24.3% are from location 2), which semester the data are from (about 1:1), and grade level (described in the participants section).

For an additional validation analysis, we utilized a metric that was calculated multiple times during each semester to track overall student engagement in activities associated with the math course. The calculation included multiple indicators such as course retention, attendance, instructional hours logged, escalation metrics, and course participation. For this study, we used the 3rd week engagement metric for each semester to explain the level of course engagement recorded prior to participation in math discourse. The combination of variables varied between the measurement occasions. Thus, while the score values ranged from 0 to 1, with 1 being high engagement, the 3rd week engagement metric had values ranging from 0 to 0.33. To increase interpretability, the scale was multiplied by 30.3, resulting in a 0 to 9.99 scale (*engagement3rdWeek*). The mean was 8.818 (SD = 2.316).

The above set of variables were included in generalized linear models to address the research questions. The independent variables consisted of the following:

- *numPart*: number of discourse sessions that the student participated in per semester
- *highPart*: 1 if participated in six or more sessions in a semester (six is the global median); 0 if participated in five or fewer sessions in a semester
- *semPart*: semester participation pattern. "A" if a student participated in at least one discourse session in semester A only; "A, B" if a student participated in at least one discourse session in semester A and at least one discourse session in semester B; "B" if a student participated in at least one discourse session in semester B only
- *avgCS*: a student's confidence and self-efficacy score (0 to 4 scale), averaged across repeated measurements within a semester
- *avgMD*: a student's mindset score (-2 to 2 scale), averaged across repeated measurements within a semester
- *score1516*: prior-year final course score ranging from 0 to 100 from the semester A or B math course taken in the 2015-2016 school year that is a grade-level below the course taken in the corresponding semester in the 2016-2017 school year
- *semester*: "A" if the participation and performance record is from semester A (Fall 2016); "B" if from semester B (Spring 2017)
- *location*: participant's school (one of the two participating schools)
- *grade*: grade level in the 2016-2017 school year—either 3, 4, or 5
- *engagement3rdWeek*: a composite score ranging from 0 to 10 with 0 being low engagement and 10 being high engagement

The dependent variables were the following:

- *score1617*: score ranging from 0 to 100 from the semester A or B math course taken in the 2016-2017 school year
- *state1617*: 1 if student achieved either Advanced or Proficient levels on the state assessment in 2016-2017 school year, 0 otherwise (Basic Proficiency or Below Basic Proficiency)

Analysis

We conducted a series of retrospective data analyses to address the research questions. For all analyses, the unit of analysis was a student's record within a course during a semester. First, to answer the first research question, we compared how students who participated in the discourse sessions changed between the sessions in terms of self-efficacy, confidence, and growth/fixed mindset measures. The statistical comparisons of these measures between sessions after controlling for grade levels were conducted by fitting a multiple regression model. The sample for the analysis was limited to students who answered every time the measures were administered.

To answer the second research question, we used generalized linear models to evaluate the effects of discourse participation on the final course performance and state assessment results after controlling for the participation variables, prior-year final math course performance, semester, location, and grade level. We evaluated the model fit using McFadden's pseudo R^2, which is defined as $R^2_{McF} = 1 - \ln(L_M) / \ln(L_0)$ where $\ln(L_M)$ is the log likelihood value for the fitted model and $\ln(L_0)$ is the log likelihood for the null model with only an intercept as a predictor. The measure ranges from 0 to near 1, with zero indicating no predictive power of the model.

We did not include in the models some popular background variables such as gender or English language learning status, because we did not find sufficient variation in the outcome variables linked to those variables.

Results

Trend in Math Confidence, Self-efficacy Toward Math, and Mindset Confidence and Self-efficacy

Overall, most students who participated in discourse sessions liked math and thought they were good at math. Among the four options, "I like math" was the most popular answer chosen by the students in 20 out of 21 measurement occasions across all grade levels and sessions during the whole school year. The overall trend of the raw responses from each item was consistent over time: it did not necessarily increase or decrease.

The result was consistent when we looked at the sum scores within students measured over time. Figure 2 shows that for all grade levels in semesters A, mean confidence and self-efficacy scores increased only slightly, not noticeably, with consecutive sessions. The trend for semester B did not necessarily change over time.
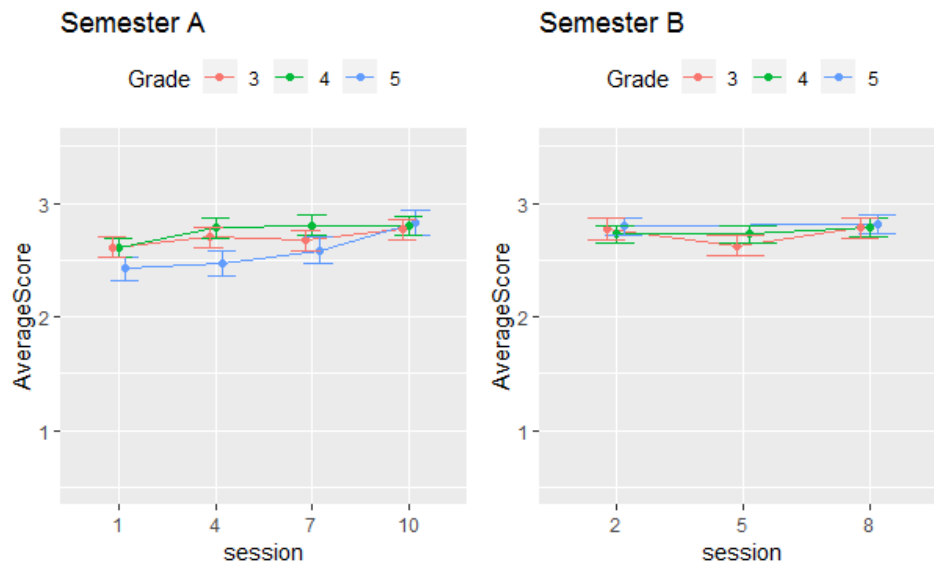

Figure 2. Trend in confidence and self-efficacy on the scale of 0 to 4.

We fitted a multiple regression model to confirm the descriptive and visual patterns found in the data. Consistent with those results, in semester A, the sessions and grade level explained only 0.4% of the variance in the confidence and self-efficacy scores ($F(5, 2238) = 2.831$, p-value $= 0.015$, effect size (Cohen's f2) $= R^2 / (1 - R^2) = 0.004$). In semester B, the model explained close to zero percent of the variance in the confidence and self-efficacy scores ($F(4, 1957) = 0.741$, p-value $= 0.564$). Although the model fit was insignificant, semester A data showed some evidence of increasing confidence and self-efficacy. After controlling for grade level, students in the later

sessions (sessions 7 and 10) had slightly higher confidence and self-efficacy scores than in session 1. The estimated mean difference of session 10 compared to session 1 was 0.223 (t = 2.884 and p-value = 0.004). In semester B, after controlling for grade level, students in the later sessions did not have confidence and self-efficacy scores that differed significantly from their scores in session 2.

Growth Mindset. Overall, most students who participated in discourse sessions across all sessions and all grade levels agreed mostly on the growth mindset items. The raw responses indicated that among the four options, most students chose either "I can always get better at math if I try" or "I think anyone can be good at math" each time the measure was administered across all grades and sessions during the entire school year. Overall trends in raw responses were mostly consistent over time.

Using the scale scores, we also observed a pattern that neither increased nor decreased. Figure 3 shows that in both semesters A and B, average mindset scores stayed in the range of growth mindset (i.e. above zero) across consecutive sessions.
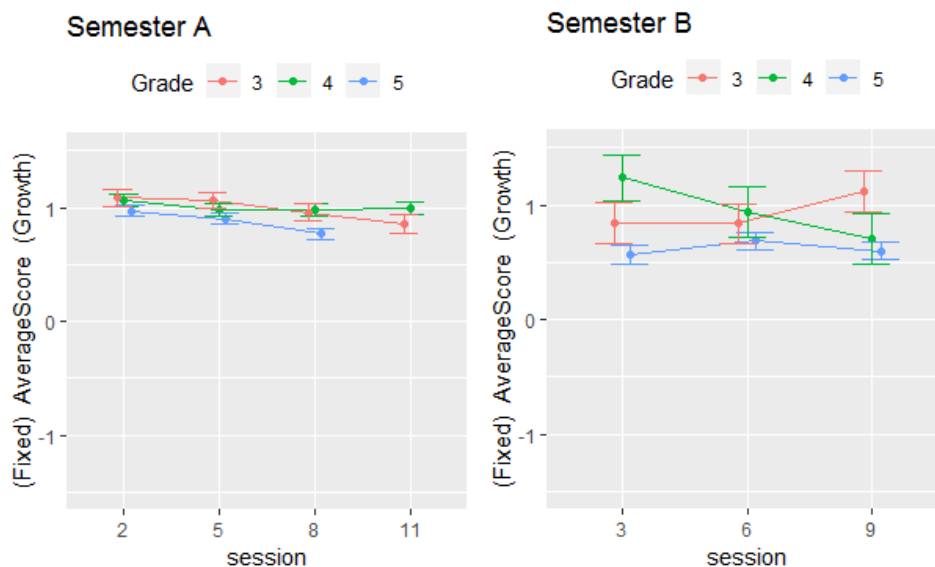

Figure 3. Trend in mindset on the scale of -2 to 2.

A multiple regression model also confirmed the descriptive and visual results. Consistent with those results, in semester A, the sessions and grade level explained only 0.7% of the variance in the mindset scores (F(5, 2296) = 4.573, p-value = 0.000, effect size (Cohen's f2) = 0.007). In semester B, the model explained only 1.5% percent of the variance in the mindset scores (F(4, 487) = 0.741, p-value = 0.021, effect size (Cohen's f2) = 0.015).

Together, we found that confidence, self-efficacy, and mindset scores did not change significantly over time across all grade levels. The differences between sessions were either not statistically significant or practically small. This result supports the idea that participation in multiple math discourse sessions does not necessarily increase confidence, self-efficacy towards math, or math mindset. Based on this finding, we decided not to consider the change or gain in these variables as covariates in the later analyses. However, because there was between-student

variance in these variables, we included average confidence, self-efficacy and mindset scores for each participant as covariates in the model.

Effects of Participating in Discourse Sessions on Math Performance

Final course scores. In order to examine how participation in discourse sessions related to math performance, we fitted multiple regression models and compared the estimated means of final course scores between groups with different amounts and patterns of participation. We fitted two models: the first model included all independent variables except for the 3rd week engagement metric which had a significant number of missing data. After fitting the first model, we included the 3rd week engagement metric in the second model, which resulted in reducing significantly the variance of other covariates. Thus, in the second model, we removed the background variables of which the effects were not significant in Model F1 ("final" 1) and also were not related to participation patterns.

Model F1 explained 46.5% of the variance in the final course scores. The results showed that after controlling for all other variables, students who participated in a higher number of sessions in a semester had significantly higher final course scores than the group that did not. For an added number of participated sessions, the final course score was estimated to increase by 1.135 (t = 4.173, p-value = 0.000). Students who participated in at least one discourse session in semester B only were estimated to have higher final scores than students who participated in both semesters A and B (estimate = 13.341, t = 1.655, p-value = 0.098). A one point increase in average confidence and self-efficacy scores was associated with an average 0.815 increase in the final course scores (t = 2.263, p-value = 0.024), and a one point increase in average mindset scores was associated with on average 1.038 increase in the final course scores (t = 1.894, p-value = 0.059). The prior year final course score also had a significant and positive effect (estimate = 0.643, t = 18.659, p-value = 0.000). The effects that were not statistically significant included difference in final course scores between the two semesters, difference between high participation (six or more sessions in a semester) vs. low participation groups, difference between semester A only vs. semesters A and B, and differences between two locations and grade levels.

In Model F2, we did not find sufficient evidence that the 3rd week engagement level has a significant effect on the final course score, after controlling for discourse participation and all other variables. Also, the effect of the confidence and self-efficacy score on the final course score was not statistically significant. Similar to Model F1 results, after controlling for all other variables, the more students participated in the discourse sessions, their final course grade was higher (estimate = 1.387, t = 2.894, p-value = 0.004). Participation in semester B only was associated with higher final course score than participation in both semesters A and B (estimate = 12.529, t = 1.668, p-value = 0.097). Higher average mindset score was associated with higher course scores (t = 2.611, p-value = 0.010).

Table 1.
Regression Results for 2016-2017 Final Course Scores

	Model F1			Model F2		
	Estimate	Std. Error		Estimate	Std. Error	
(Intercept)	17.088	(3.081)	***	16.581	(6.443)	*
numPart	1.135	(0.272)	***	1.387	(0.479)	**
highPart	-0.218	(1.970)		2.149	(3.388)	
semPartA	-2.187	(2.156)		-0.686	(4.082)	
semPartB	13.341	(8.059)	.	12.529	(7.513)	.
avgCS	0.815	(0.360)	*	0.331	(0.682)	
avgMD	1.038	(0.548)	.	2.929	(1.122)	**
score1516	0.643	(0.034)	***	0.551	(0.069)	***
semesterB	0.361	(0.888)				
locationB	-0.726	(0.914)				
grade4	0.225	(0.979)				
grade5	-0.550	(0.964)				
engagement3rdWeek				0.418	(0.346)	
AIC, BIC	6560.11, 6621.83			1618.99, 1652.65		
Adjusted R^2	0.465			0.452		
Effect size (Cohen's f2)	0.869			0.825		
F-statistic (DF)	68.36 (11, 840) ***			22.94 (8, 205) ***		
N	852			214		

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Although the two models are not nested, we evaluated general model fit of the two models. Model F2 was as good as Model F1, despite a drastic decrease in sample size. Learning that the effect of the engagement metric was not significant, with the results from Model F1, we calculated the fitted final score values given the model parameter estimates. The results shown in Figure 4 indicate that the more sessions students participated in, the higher their predicted final scores were, after holding all other variables consistent.
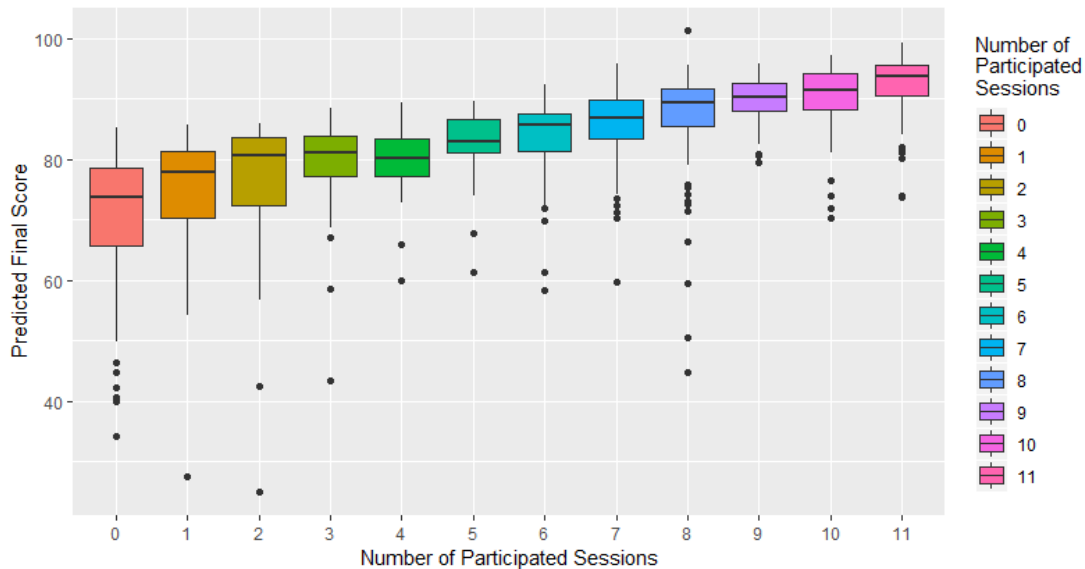
Figure 4. Predicted final score by number of participated discourse sessions.

State Assessment Results

To examine how participation in discourse sessions relates to math state test performance, we fitted two logistic regression models. The first model included all independent variables except for the 3rd week engagement metric, which had a significant amount of missing data. The second model included the 3rd week engagement metric and did not include the background variables that were not related to key constructs and participation patterns. The state test results from the prior year were collected for semester B courses only. Therefore, analyses of state test scores excluded students who did not participate in any discourse session in semester B, and those cases were removed from both the S1 ("state" 1) and S2 models.

Model S1 classified 76.8% of the raw state test results accurately. The results showed that after controlling for all other variables, students who participated in a higher number of sessions in a semester had significantly higher odds of scoring Proficient and Advanced than the group that did not. For an added number of participated sessions, the odds of scoring Proficient and Advanced (versus scoring Basic Proficiency or below) was estimated to increase by 23% ($z = 2.801$, p-value = 0.005). A one-point increase in average confidence and self-efficacy scores was associated with 67% increase in the odds of scoring Proficient or above ($z = 5.229$, p-value = 0.000), while a one-point increase in average mindset scores was associated with a 27% decrease in the odds of scoring Proficient or above ($z = -2.285$, p-value = 0.022). The prior year final course score also had a significant and positive effect (OR = 1.08, $z = 5.125$, p-value = 0.000). Students' participation in semester B only was associated with 76% higher odds of scoring Proficient or above ($z = 2.470$, p-value = 0.014) than students' participation in semester A. The students at location B also had 75% higher odds of scoring Proficient or above ($z = 2.415$, p-value = 0.016) than those at location A. Effects that were not statistically significant included differences in state assessment results between high participation (six or more sessions in a semester) and low participation groups, differences between participating in semester B only and semesters A and B, and differences between grade levels.

Table 2

Logistic Regression Results for 2016-2017 State Assessment Result.

	Model S1			Model S2		
	Estimate	Std. Error	OR	Estimate	Std. Error	OR
(Intercept)	-10.550	(1.427) ***	0.00	-7.396	(2.159) ***	0.00
numPart	0.204	(0.073) **	1.23	0.144	(0.125)	1.15
highPart	-0.336	(0.501)	0.71	-0.642	(0.890)	0.53
semPartB	2.657	(2.425)	14.26	2.090	(1.854)	8.09
avgCS	0.516	(0.099) ***	1.67	0.472	(0.176) **	1.60
avgMD	-0.309	(0.135) *	0.73	-0.098	(0.284)	0.91
score1516	0.076	(0.015) ***	1.08	0.050	(0.022) *	1.05
semesterB	0.568	(0.230) *	1.76			
locationB	0.557	(0.231) *	1.75			
grade4	-0.012	(0.241)	0.99			
grade5	-0.412	(0.256)	0.66			
engagement3rdWeek				0.062	(0.104)	1.06
AIC, BIC	661.37, 710.85			190.07, 214.21		
McFadden's R^2	0.617			0.896		
Classification Rate	76.8%			70.2%		
Chi-square (DF, p-value)	144.42 (10, 0.000) ***			21.64 (7, 0.003) **		
N	665			151		

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1; OR: odds ratio.

Similar to the results from the final course score models, in Model S2, we did not find sufficient evidence that the 3rd week engagement level had a significant effect on the state assessment results, after controlling for discourse participation and all other variables. Meanwhile, the effect of the confidence and self-efficacy score on the state assessment results continued to be statistically significant following the results from Model S1. However, contrary to Model S1 results, after controlling for all other variables, the more students participated in the discourse sessions, the odds of scoring Proficient or above was higher (15% increase) but not significantly so (z = 1.151, p-value = 0.250). Higher average mindset score was not associated significantly with higher state test results.

From Model S1, along with the significant effect of number of participated discourse sessions, we found that the effects of prior year final course scores and average confidence/self-efficacy scores on state assessment results are significant covariates. We thus fitted two simpler logistic models with only two variables as independent variables; the first model included prior final course scores and number of participated sessions, and the second model included average confidence/self-efficacy score and number of participated sessions as predictors for scoring at or

above Proficient on state assessments. The results are shown in Figure 5, highlighting two patterns: 1) that the more sessions students participated in, the higher the probability of scoring at or above Proficient on state assessments are, after holding the prior achievement consistent, and 2) that a higher frequency of discourse participation is associated with a higher probability of scoring at or above Proficient on state assessments, regardless of the confidence and self-efficacy level. Overall, high participation in discourse sessions was associated with increased probability of scoring at or above Proficient in state assessments, after controlling for other covariates.
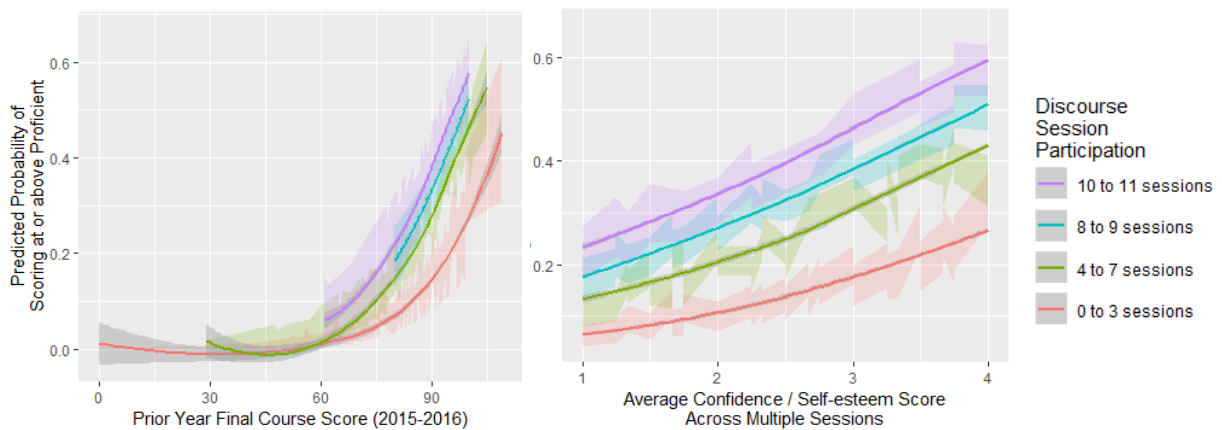

Figure 5. Probability of scoring at or above proficient in state assessment (2016-2017), by prior year final course score (left) and by average confidence/self-efficacy scores (right).

Discussion

Due in part to their high mobility, students served by fully online K-12 virtual schools have shown lower mathematics performance (Gatti, 2018). In an effort to identify effective supports for online students, we analyzed empirical data to examine whether participation in synchronous discourse sessions offered in asynchronous online math courses would affect math learning outcomes. In the 2016-2017 school year, we designed and embedded synchronous discourse sessions in math courses at two fully online virtual elementary schools. The results showed that students who participated in more discourse sessions had both higher final course scores and higher odds of scoring at or above Proficient on the state assessments, after controlling for multiple covariates, including prior math achievement level and the levels of confidence, self-efficacy, and math mindset.

Confidence, self-efficacy, and mindset, however, did not increase with participation in math discourse. This finding may have been due to students' already high ratings on these metrics at the start of the sessions. Also, without access to all the interactions and activity within the math discourse sessions, it is unclear whether facilitators adequately and consistently reinforced these student constructs. Future work should address changes in attitude more directly through intentional observation and possible manipulation of the discourse environment.

This study filled in the gap in literature by describing the design of a small-group synchronous discourse intervention aimed at improving math learning for fully online virtual school students. Consistent with prior work that encouraged students to use language to reason

through problem solving (Mercer & Sams, 2006), this expert teacher-student math discourse led to improved math learning.

What is new is that we now have evidence that discourse can be implemented effectively within a fully online learning environment and with students who do not typically have opportunities to engage in this way. Hence, the results of this study imply that neither experience with this mode of interaction nor physical proximity with peers and teachers is necessary to the success of math discourse. More broadly, given the importance of meaningful synchronous interaction and collaboration in increasing online student learning (Kang & Im, 2005; Hiltz et al., 2000), this work offers evidence of a specific, synchronous approach to target and improve math performance.

Our analyses, however, did not tell us why students had higher outcomes. Even though math specialists were provided with the same trainings that included guidelines for facilitating the sessions, we were unable to observe consistently whether they were implemented with fidelity. What elements of the activities that students and facilitators encounter in the sessions were really related to the better outcomes? How might facilitators integrate Lampert's (1990) work so that the online social interaction facilitates problem-solving over solutions? Future studies might examine directly the implementation dynamics and fidelity of the discourse sessions, which will contribute to our understanding of the most important components of the effective implementation of math discourse in online math learning of fully online K-12 students and beyond. Specific research topics may include size and homogeneity of students, optimal participation level, task type and difficulty, facilitator prompts and guiding questions as well as facilitator experience level and background. Close examination of the interactions will allow for a more direct test of the impact of math discourse on mindset, self-confidence, and self-efficacy.

Also, findings from this study are limited to a set of grade levels, math topics, and certain types of math problems. By expanding the participants and content, a future study can explore additional applications for this study. Finally, integrating a control group would further strengthen the comparison and our ability to predict learning outcomes.

This study has important implications for the future design and implementation of online math instruction as well as the training and development of fully online math teachers. As courses are designed to increase students' opportunities to engage in synchronous student-to-student discourse, future studies will need to determine the extent to which discourse can be effectively led by existing teachers or whether a math specialist or other supports are necessary. This question speaks to the feasibility of applying this work more broadly across a variety of blended and fully online settings.

References

Ahn, J. (2016). *Enrollment and Achievement in Ohio's Virtual Charter Schools*. Thomas B. Fordham Institute. Retrieved from https://edexcellence.net/publications/enrollment-and-achievement-in-ohios-virtual-charter-schools

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*(5), 181-185. https://doi.org/10.1111/1467-8721.00196

Choi, J., Belenky, D., DiCerbo, K., Lai, E., & Wardlow, L. (2016). Learners and learning in virtual schools: A literature review. Pearson Efficacy & Research. London: Pearson.

Choi, J., Walters, A., & Hoge, P. (2017). Self-reflection and math performance in an online learning environment. *Online Learning Journal, 21*(4). http://dx.doi.org/10.24059/olj.v21i4.1249

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. *Self-efficacy in Changing Societies, 15,* 334.

Boaler, J. (1999). Participation, knowledge and beliefs: A community perspective on mathematics learning. *Educational Studies in Mathematics,* 40(3), 259-281. https://doi.org/10.1023/A:1003880012282

Diener, C. I., & Dweck, C. S. (1980). An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology*, 39(5), 940. Retrieved from https://pdfs.semanticscholar.org/d8a8/1dd37b0c47bacf3beb36a01d17f9fd50712e.pdf

Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. Random House Incorporated.

_____. (2017). *Mindset: Changing the Way You Think to Fulfil Your Potential*. Hachette, UK.

Fabert, N. S. (2014). *Growth Mindset Training to Increase Women's Self-Efficacy in Science and Engineering: A Randomized-Controlled Trial*. Arizona State University. Retrieved from https://repository.asu.edu/attachments/137308/content/Fabert_asu_0010E_14151.pdf

Gatti, G. (2018). *A Comparison Study of Connections Academy Schools to Matched Brick and Mortar and Virtual Schools, Examining the Types of Students Who Attend K–12 Virtual School and the Effects on Performance of a Highly Mobile Student Body*. Efficacy & Research Report. London: Pearson.

Hiltz, S. R., Coppola, N., Rotter, N., Turoff, M., & Benbunan-Fich, R. (2000). Measuring the importance of collaborative learning for the effectiveness of ALN: A multi-measure, multi-method approach. *Journal of Asynchronous Learning Networks, 4*(2), 103–125

Kang, M. & Im, T. (2005). Factors of learner-instructor interaction which predict perceived learning outcomes in online learning environment. *Journal of Computer Assisted Learning, 29*, 292-301.

Kim, C., Park, S. W., & Cozart, J. (2014). Affective and motivational factors of learning in online mathematics courses. *British Journal of Educational Technology, 45*(1), 171-185.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27*(1), 29-63. https://doi.org/10.3102/00028312027001029

Lampert, M., Rittenhouse, P., & Crumbaugh, C. (1996). Agreeing to disagree: Developing sociable mathematical discourse. In D. Olson & N. Torrance (Eds.), *Handbook of Education and Human Development* (pp. 731–764). Oxford: Blackwell. Retrieved from http://www-personal.umich.edu/~mlampert/lampert%20pdfs/Lampert_etal_1996.pdf

Lehrer, R., & Schauble, L. (2005). Developing modeling and argument in the elementary grades. In T. A. Romberg, T. P. Carpenter, & F. Dremock (Eds.), *Understanding Mathematics and Science Matters (Part II: Learning with Understanding)*. Mahwah: Lawrence Erlbaum Associates.

Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education*, *20*(6), 507-528. https://doi.org/10.2167/le678.0

Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283-297. https://doi.org/10.1007/s11217-007-9071-1

National Council of Teachers of Mathematics (2014). *Principles to Action: Ensuring Mathematical Success for All.* Reston, VA: National Council of Teachers of Mathematics.

Resnick, L. B., Bill, V., & Lesgold, S. (1992). Developing thinking abilities in arithmetic class. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-Piagetian Theories of Cognitive Development: Implications and Applications for Education* (pp. 210–230). London: Routledge.

Rumberger, R. W. (2015). *Student Mobility: Causes, Consequences, and Solutions.* National Education Policy Center. Retrieved from https://files.eric.ed.gov/fulltext/ED574695.pdf

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), A Vol. in the educational psychology series. Development of achievement motivation (pp. 15-31). San Diego, CA, US: Academic Press.

Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W., & Van Donge, C. (2015). *Online Charter School Study.* Stanford, CA: Center for Research on Education Outcomes. Retrieved from https://credo.stanford.edu/pdfs/OnlineCharterStudyFinal2015.pdf

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477. https://doi.org/10.2307/749877

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302-314. https://doi.org/10.1080/00461520.2012.722805

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82-91. https://doi.org/10.1006/ceps.1999.1016