International Journal for the Scholarship of Teaching and Learning

Volume 9 | Number 2 Article 11

July 2015

Supporting Graduate Student Instructors in Calculus

Daniel L. Reinholz
University of Colorado, Boulder, reinholz@berkeley.edu

Murray Cox
University of Colorado, Boulder, murray.cox@colorado.edu

Ryan Croke
University of Colorado, Boulder, ryan.croke@colorado.edu

Supporting Graduate Student Instructors in Calculus

Abstract

This paper focuses on the use of the teaching method Peer-Assisted Reflection (PAR) to help graduate student instructors (GSIs) develop as teachers. PAR engages students in analyzing the work of their peers and providing feedback to promote their abilities of communication, collaboration, and persistence. The goals of the PAR activity were taken as goals for instruction generally, and used to support the GSIs to develop student-centered pedagogies. This report provides in depth case studies of two of four GSIs involved in implementing PAR in their introductory calculus recitation sections. Two of the GSIs showed considerable changes in practices and beliefs, while the others showed little growth.

Keywords

Mathematics Education, Reflection, Teacher Learning, Graduate Student Instructors

Supporting Graduate Student Instructors in Calculus

Daniel L. Reinholz[1], Murray Cox[2] and Ryan Croke[2]
[1]Center for STEM Learning and [2]Applied Mathematics, University of Colorado, Boulder, CO, USA

(Received 24 March 2015; accepted 6 July 2015)

This paper focuses on the use of the teaching method Peer-Assisted Reflection (PAR) to help graduate student instructors (GSIs) develop as teachers. PAR engages students in analyzing the work of their peers and providing feedback to promote their abilities of communication, collaboration, and persistence. The goals of the PAR activity were taken as goals for instruction generally, and used to support the GSIs to develop student-centered pedagogies. This report provides in depth case studies of two of four GSIs involved in implementing PAR in their introductory calculus recitation sections. Two of the GSIs showed considerable changes in practices and beliefs, while the others showed little growth.

INTRODUCTION

This paper focuses on the development of graduate student instructors (GSIs) teaching introductory calculus. We focused on calculus because it continues to be a major gatekeeper preventing students from accessing STEM careers (Bressoud, Carlson, Mesa, & Rasmussen, 2013). The four GSIs in the study participated in a semester-long working group to implement Peer-Assisted Reflection (PAR) in their recitation sections (Reinholz, 2015a). The present study focuses on how PAR could be used to help GSIs develop student-centered pedagogies. Although the study takes place in calculus, our approach focuses on classroom discussions, which should be broadly applicable across disciplines. We focus on the development of two of the GSIs as a cross-case study, and address two main questions:

1. How did GSIs develop student-centered pedagogies as a result of the working group?
2. Did implementing PAR impact student outcomes in introductory calculus?

THEORETICAL FRAMING

A growing body of evidence highlights the positive impact of student-centered classrooms in post-secondary education (Freeman et al., 2014). Students in such classrooms outperform their peers in traditional classrooms and are more likely to persist in STEM majors (Kogan & Laursen, 2014). However, despite the wealth of effective, student-centered teaching methods, college mathematics classrooms are dominated by lecture (Lutzer, Rodi, Kirkman, & Maxwell, 2005). More than developing new teaching methods, there is an urgent need to increase the adoption of student-centered pedagogies and to support instructors to use them effectively (PCAST, 2012).

The majority of attrition in STEM majors takes place during the first two years of college (PCAST, 2012). Because graduate student instructors (GSIs) often play an important instructional role in introductory courses, they need support to learn to use student-centered teaching methods to improve retention. However, GSIs often receive little systematic support for teaching (Austin, 2002). Moreover, GSIs face a number of challenges as they transition into graduate studies (Hauk et al., 2009), exacerbating this lack of training. The development of GSIs is an emerging area of research, with an increasing number of reports published in refereed journals, both in mathematics (Deshler, Hauk, & Speer, 2015; Speer, Murphy, & Gutmann, 2009) and with regards to student-centered teaching in other disciplines (e.g., Miller, Brickman, & Oliver, 2014; Wright, Bergom, & Brooks, 2011). This paper adds to the emerging literature by focusing on the development of student-centered practices in introductory calculus.

In the present study, GSIs were supported to implement a student-centered teaching practice, Peer-Assisted Reflection (PAR; described below), in their recitation sections. The use of PAR was intended to help GSIs focus more on student thinking throughout all of their teaching practices. This approach, engaging GSIs in thinking about student thinking, has been suggested as an effective way to support their professional development (Kung, 2010; Kung & Speer, 2009). This approach has been used successfully with K-12 teachers, such as through the Shell Centre's Classroom Challenges (Herman et al., 2014). By incorporating these special lessons into their teaching throughout the year, instructors improved student outcomes and also grew as teachers. The present approach was also inspired by Cognitively Guided Instruction, through which instructors improved their practices by thinking deeply about children's mathematical strategies (Franke, Carpenter, Levi, & Fennema, 2001).

The area of focus for the present work is on GSI's abilities to facilitate whole-class discussions. Classroom discussions are a major source of student learning, and also help communicate to students what is valued in the classroom (Michaels, O'Connor, Hall, & Resnick, 2010). Moreover, establishing a high-functioning discourse community within the classroom is seen as one of the major challenges of enacting reform mathematics approaches (Hufferd-Ackles, Fuson, & Sherin, 2004). Accordingly, if GSIs were able to improve in their ability to facilitate discussions, it would be of great benefit to their students and it would be a powerful sign of teacher learning.

METHOD

Design

PAR. PAR involves a four-stage peer assessment cycle (Reinholz, 2015b), requiring students to: (1) work on difficult homework problems, (2) reflect on their work, (3) analyze peers' work and exchange feedback, and (4) revise their own work. Each week students completed one PAR homework problem, which was a challenging, multi-part problem that required written explanations; students completed a total of 14 PAR problems during the semester. Students completed steps (1), (2), and (4) of the PAR cycle outside of class, while step (3), the peer conferences, took place during recitation sections. For their conferences, students traded work with a peer, read over each other's solutions silently and provided written feedback for five minutes, and then had an additional five minutes to discuss their feedback with one another. For the self-reflection portion of PAR, students were provided with six yes-no questions reminding them of important aspects of a clear explanation: (1) Did you answer all questions asked, showing all steps, in the proper order?; (2) Did you label and explain all graphs, include units, etc.?; (3) Did you explain why (not just what)?; (4) Did you avoid the use of pronouns (and other ambiguous language)?; (5) Did you use terms according to their mathematical definitions?; and (6) Did you draw a diagram to support your explanations? For peer feedback, students were asked to comment on the communication and correctness of their peer's solutions.

To teach students how to give feedback to one another, students participated in whole-class discussions focused on the analysis of sample student work. Immediately after students finished exchanging peer feedback, they were given three sample solutions to part of the PAR problem. The students had a few minutes to analyze the solutions silently, and then the whole class had a discussion about how to provide useful feedback. Prior studies have shown that these activities help students improve their explanations, collaboration, and persistence (Reinholz, 2015a).

Working Group. GSIs met as a group six times during the semester, for one hour every other week. A member of the research team facilitated these meetings, and two full-time calculus faculty also attended. During the meetings, GSIs practiced student-centered teaching methods (e.g., facilitating discussions), provided feedback to one another, and discussed instruction more generally. The first three sessions emphasized how to implement PAR; the final three sessions focused on the classroom practices surrounding PAR (e.g., facilitating discussions), and their relation to the PAR goals of supporting collaboration, communication, and persistence.

Participants and Data Collection

The study took place in a fall semester of introductory college calculus. There were three types of sections associated with the course: lecture (50 minutes, 3 days per week), recitation (50 minutes, 1 day per week), and workshops (100 minutes, 1 day per week). All students were required to enroll in one lecture section and one recitation section. The lecture sections had about 100 students each, and the 35 corresponding recitation sections each had 20-30 students. Faculty ran the lecture sections while GSIs facilitated recitations. The optional workshops were based off of prior work on supporting minority learners in calculus (Treisman, 1992), but they were not a focus of the study and were not a confounding factor; students in workshop sessions tend to be distributed fairly evenly across different recitation sections. This paper focuses on the interactions between GSIs and students in the recitations. GSIs have full freedom to teach their recitation sections as they choose; typical activities involve GSIs answering homework questions, working out problems for students, and providing students with practice problems. These practices tend to be teacher-focused, with little opportunity for collaboration between students, but this can vary depending on the GSI.

The department chair recommended 6 GSIs for the project from a larger pool of 14 GSIs, all 6 of whom were contacted to confirm interest; no specific criteria were given for this choice of GSIs, simply that they were likely to be interested. This group of GSIs consisted of three male GSIs, each with multiple years of experience as GSIs, and three female GSIs, each with a year or less of experience. After interest was confirmed, four of these six GSIs were chosen using a random number generator. Amongst the chosen GSIs were two women, Beth and Amy, and two men, Adam and Wong (all names are pseudonyms). Each GSI taught 2-3 recitation sections, one of which was randomly chosen to be their PAR section. Of the four GSIs, Amy and Beth had similar growth trajectories, as did Adam and Wong. To illustrate these trajectories in depth and provide contrast between them, this paper focuses on just two GSIs. The first is Beth, a second-year graduate student with two semesters of experience as a GSI; she had facilitated recitations for calculus 1 and differential equations. The second is Wong, a third year international student who had experience as a GSI for the entire calculus sequence (including differential equations).

Each participating GSI was interviewed at the beginning and the end of the semester, but the pre-semester interviews were not recorded (only notes were taken). Each GSI was observed six times, approximately every other week. Diagnostic pretest scores and scores on the four exams were collected for the 871 students enrolled in the course.

Analysis

Student exam scores were compared by recitation section. Recitations were observed and analyzed for student-centered practices during discussions: (1) teacher probing for explanations, (2) teacher linking student responses, (3) students responding directly to one another, and (4) students presenting at the board. Each of these categories was coded for frequency. We focused on these "high-leverage" practices because: they can help improve learning for all students, they are used frequently in teaching, they can be used as a part of many different teaching approaches, and they are discrete practices that can be taught (Ball, Sleep, Boerst, & Bass, 2009). We used a practice-based approach to professional development, aimed at helping new teachers elicit, probe, and connect student ideas during discussions (Boerst, Sleep, Ball, & Bass, 2011). These particular practices are not directly related to the PAR activity, although as students improve their communication skills through PAR one might expect that such improvements would be visible during whole class discussions.

Student presentations were only coded if students came to the board and students explained their ideas to the class. If a student wrote a solution on the board but the teacher did all of the talking, then it was not coded as a student presentation, because the student was not a part of the discussion. During field observations, teacher questions and student speech were captured verbatim, while extended teacher explanations were paraphrased. Thus, the transcripts of field notes could be analyzed to see how opportunities for student interaction were created and seized.

Teacher speech was coded as probing for explanations when she asked follow up questions such as "why?," "what do you mean?," or "can you explain more?". When a teacher asked questions to the students, but not in response to their ideas (e.g., "what do you think about problem two?") it was not coded as probing, because it was not pushing students to elaborate their thinking beyond what they had already offered. Teacher speech was coded as linking responses when she explicitly asked students to respond to each other. For instance, after a student said something, the teacher might say, "what do you think?" to another student in the class. Anytime stu-

TABLE 1. First exam scores (no significant differences).

	Beth	Amy	Wong	Adam	Experimental	Comparison	Other
PAR	72.22	70.65	73.42	71.32	71.95		-
Non-PAR	68.65	69.39	71.43	68.00	69.65	72.34	70.59

TABLE 2. Final exam scores (no significant differences).

	Beth	Amy	Wong	Adam	Experimental	Comparison	Other
PAR	91.74	85.58	93.23	88.26	89.82		-
Non-PAR	85.86	92.00	91.69	71.42	87.88	86.03	87.77

dents directly responded to the speech of another student, not the teacher, it was coded as such. The unit of analysis was a single turn, whether it was a few words or multiple sentences. For each GSI, two lessons were chosen at random and coded independently by two researchers. Finally, interviews were analyzed for statements related to classroom practices.

RESULTS

Student Outcomes

There were no significant differences across sections on the diagnostic pre-calculus pretest or first exam (Table 1). This indicates that the recitation sections could be considered equivalent for the purposes of comparison. Exam scores are grouped by: individual GSIs (by name), the group of four GSIs using PAR (experimental), the two interested GSIs who did not participate (comparison), and the remaining eight GSIs (other).

Final exam scores were also compared across sections (Table 2). Numerical differences in all sections other than Amy's favored PAR (by about 5%), but differences were not significant. In her post-semester interview, Amy discussed struggling to engage her PAR section, despite using similar teaching practices in her other sections to facilitate discussions; there may have been something atypical about that group of students, accounting for the lower exam scores. Exams 2 and 3 were also analyzed, but the differences were small and not significant.

PAR sections and non-PAR sections were also compared for student retention (Table 3); PAR students were more likely to persist in the course than those in non-PAR sections.

Changes in Teaching Practices

As evidenced by classroom observations and pre-interviews, the observed GSIs initially taught primarily through lecture. Beth, Wong, and Adam had previously taken a first-semester teacher training course (Amy was currently enrolled), but they reported learning little from it. For the eight lessons double coded, there were a total of six discrepancies between the coders: five of these related to teacher probing for explanations, and one related to students responding to each other. All discrepancies were resolved with discussion, and consensus was reached on the application of 51 codes. Thus, inter-rater agreement was 88% for the coded lines. A total of 799 lines were coded, for an agreement of 99% by line.

TABLE 3. Student drop rates in the course.

	Experimental	Comparison	Other
PAR	8.20%	-	-
Non-PAR	13.72%	14.00%	12.53%

Figure 1 shows the changes in teaching practices for all of the GSIs from the beginning of the semester to the end of the semester. Each discrete instance of a specific practice being observed (e.g., probing student reasoning) was counted as a part of the GSI's score. The change in practices is the sum of the counts from the last three observations (weeks 8, 10, and 13) minus the sum of the counts from the first three observations (weeks 2, 4 and 6), which gives a change score for practices. Given that nearly all of the values are positive, it appears that all GSIs had at least some benefit from the working group. As Figure 1 indicates, both Beth and Amy made considerable changes in their teaching practices: they increased their use of probing for student ideas, they invited students to respond directly to one another, and they brought students up to the board to present their ideas. As a result of all of these practices, students began to respond directly to one another during class discussions, rather than only to the teacher.

Table 4 shows the initial teaching practices for all four GSIs. As the table highlights, the GSIs were relatively similar in their initial teaching practices. Thus, the minimal changes in Wong and Adam's teaching cannot be attributed to a different starting point in teaching practices than Amy and Beth; the data indicate that Wong and Adam both showed less change in their practices. The case studies, elaborated below, provide an in-depth look at how these quantitative changes were manifested in teaching practices.

FIGURE 1. Change in GSI teaching practices.

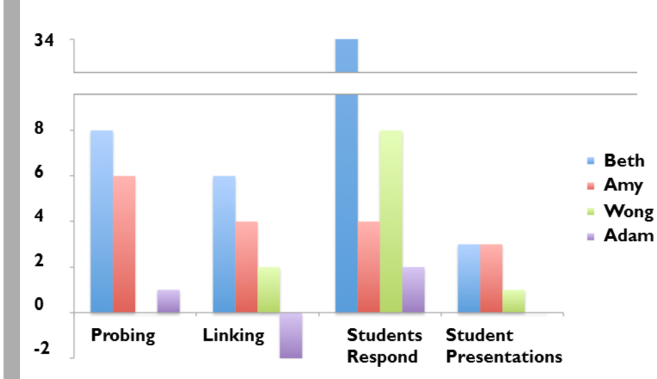

TABLE 4. Initial teaching practices (sums of weeks 2,4,6).

	Beth	Amy	Wong	Adam
Probing	11	10	8	11
Linking	0	3	0	2
Student Response	1	3	0	0
Student Presentation	0	0	0	0

Case Study: Beth (Considerable Changes)

Early in the semester, Beth mostly stood in the front of the class during student work time. Over time, she began to circulate around the room and interact with students more. In the beginning, students rarely spoke directly to each other. Large shifts in practice became evident during week 8. Beth began asking students to respond directly to one another, to come up to the board, and during week 13 she even had a student facilitate the discussion of sample student work. Because the student facilitator led the discussion, Beth did not have instances of probing for explanations during week 13. These data are now illustrated with classroom episodes.

Beginning of the Semester. Beth rarely probed students for reasoning or asked them to respond to each other; she mostly provided judgments and summaries of student responses. Consider this exchange from week 2, in which a student was describing a sample solution:

1. Student A: It uses exact terms, which you can understand
2. Beth: Yes, it's very specific. One thing we want to try to do is eliminate the use of pronouns like "it," which aren't very exact.
3. Student B: It explains why.
4. Beth: Yeah, it gives you the answer, but it also gives you a physical explanation.

Students offer two ideas (lines 1 and 3), and Beth elaborates the students' reasoning for them (lines 2 and 4). In other instances when Beth did probe for reasoning, it was at a surface level, as in the following conversation (week 2). Beth did not ask the student to probe deeper (e.g., by providing a specific example), and just accepted the response (line 4):

1. Student A: They gave an answer but not an explanation, or at least not a developed explanation.
2. Beth: How could they have made it better?
3. Student B: Give an example
4. Beth: Yeah, an example, or a counterexample.

During the week 4 instructor's meeting, the group discussed the use of wait time and how to probe deeper into student thinking. GSIs were issued the "what do you mean" challenge; they had to ask this type of question at least twice in their recitations and report back to the group. In a conversation before her week 6 lesson was observed, Beth explained that in one of her other sections, "students were motioning in the air to create a graph, and when I asked them 'what do you mean,' it forced them to come up to the board and draw it out." Because students never presented at the board during the first two observations, this indicated a shift in practice. During week 6, Beth asked students "what do you mean" to elicit reasoning (line 3):

1. Beth: Are there any ways that they could make this better?
2. Student A: Show how they got $h(x) = 1$.
3. Beth: What do you mean?
4. Student A: They need to show how they got the 0 derivative, and they need to explain more why they did that.

When Beth pushed the student to elaborate, the student articulated that the sample student work they were discussing required elaboration; namely, the sample student should show how the zero derivative was found and explain the connection between a zero derivative and a constant function. The previous two examples were both coded as probing for reasoning, but they highlight how the nature of Beth's probing developed over time. Beth explained in her interview:

> *I think this semester, compared to past semesters, I got students to come up to the board more and got them to explain their words more. I think in the past as a teacher, I would listen to someone, and say I think I know what they are saying and I think they are right, but I'm going to rephrase it, like, the right way, which has pros and cons I think. I think I pushed students more to explain their ideas, and to explain them to each other.*

Beth describes a change in practices that is consistent with our observations; rather than doing the work for the students, she gave them a chance to explain and elaborate to each other.

Later in the Semester. During the working group meeting following the week 6 observation, each GSI generated a list of techniques to improve participation. Beth's list was: (1) create a contact list; (2) facilitate students working together; (3) ask students to respond to each other; and (4) wait for 3 hands before calling on a student. To facilitate out of class collaboration, Beth wanted to create a contact list for students. The other techniques focused on giving students more opportunities to speak during class, by: (1) facilitating student collaboration, (2) having students respond directly to one another, and (3) increasing wait time. Using these practices, Beth facilitated student-student interactions. Rather than summarizing student talk, she made students respond to each other (e.g., in week 8, lines 4 and 6):

1. Student A: Drawing a horizontal line through the graph doesn't do anything.
2. Beth: What do you mean?
3. Student A: They say drawing a horizontal line changes the sign of the derivative, which it doesn't.
4. Beth: What do you think? [points to another student]
5. Student B: They are doing the wrong part of the graph, they should be looking if it is above or below the axis.
6. Beth: What were they thinking? [calls on another student]

Above, Beth directly facilitates interaction, pushing students to respond directly to one another. In other instances, students spontaneously responded to one another. For instance, during week 10, a student responded to his peer (line 3), even after Beth had already given a response (line 2).

1. Student A: What happened to this square? Should we take the implicit derivative?
2. Beth: Some people are tempted to do that, but we don't need to. We're actually only looking at one variable.
3. Student B: If you take derivative of d, and solve it all out, it ends up being that on top, with something else on bottom. So either way that is what has to go to 0, because that's what's on top.

Beth also began to create space for student discussions, as in week 8. Students were discussing the application of the mean value theorem to a particular homework problem. Beth asked students what they were thinking (line 2), and students began to respond to one another (lines 5-14). This contrasted previous discussions, in which Beth would have made a judgment about the correctness of student responses and subsequently shut down discussion:

1. Student A: Like the mean value theorem, if speeds are changing, and at some point they switch, then at some point they will have to be at the same speed when they change no matter what. And if they don't change their speeds, then they will have to change.
2. Beth: What do you think?
3. Student B: I disagree. I don't think it has to be at the same time.
4. Beth: Why?
5. Student B: They had to have the same speed at some time during the race, but they could have had it at different times. The mean value theorem just says that the secant line between the two end points of the interval has to be the same as the tangent line somewhere, it doesn't say that it has to be the same. Oh I'm not explaining it good…
6. Student C: Can I draw a graph?

…

12. Student D: I am against it, only for the same time portion, because one of the runners could have run really fast and taken a break at the friend waiting for him to catch up. Then he stops…
13. Student E: But when he decreases, he'll at some point reach his friend's speed.
14. Student F: But, then if it drops right to 0 it wouldn't be differentiable. [discussion continues]

This was the first whole-class discussion between students observed in Beth's classroom. In her post-interview, Beth discussed the value of these interactions:

> *There was one class, in my PAR section, where two students disagreed on the answer and they kind of fought it out… That was really cool. They were literally talking to each other and never even looked at me for assurance, or [to ask] is this right. That was kind of cool.*

Beth described the interaction as "really cool," indicating that she valued interactions in which students were taking ownership and agency over their own learning processes.

Another example of Beth's shifts in practice came during week 13. Rather than facilitating the sample work discussion, she had a student volunteer and facilitate. Although Beth helped elicit student thinking, she stayed out of the discussion and let the students do the thinking:

1. Facilitator (Student A): Okay. Solution 2. What do you think about solution 2?
2. Student B: It's wrong, because the area under curve isn't below the x-axis.
3. Students (choral): It is!
4. Student C: Draw a graph.
5. Facilitator: What were you thinking? [points to Student B]
6. Student B: Something else.
7. Facilitator: Because you might have been correct… [waits for student to respond, but there is no response]
8. Facilitator: Anyone else?
9. Beth: Maybe you could draw a picture?
10. Facilitator: Does anyone want to draw a graph for what the integral would look like?
11. Student D: I'll do it. [Student D comes up to the board]

Here Student A asked questions like Beth typically did as a facilitator. Rather than simply telling the answers to the students, the facilitator asked students what they thought (lines 1, 5, 7, 8, and 10). In response to these solicitations, students in the class offered up their ideas (lines 2, 4, and 11). In her interview, Beth discussed the value of exposing students to their peers' ideas:

> F*or whatever reason, I think students sometimes, they almost soak in information better from their peers than from me. Which is sad. I don't know, I think sometimes they do. Or maybe sometimes…another student can say it more easily because they are on the same level. Maybe I don't even realize that I'm saying things they might not understand.*

This quote highlights Beth's shift towards student-centered practices, and the difficulty of the shift; Beth had to come to terms with a new idea of what it meant to be a good teacher, allowing students to do more of the work. Later in the interview, Beth described her growth as follows:

> *I've definitely become a lot more reflective about my teaching, and I think about it a lot more than in the past. Just, I question, is this working. I feel like honestly when I first starting teaching, I would just do stuff, and think it's good, oh it's great, and I never really asked, is this working, do I think they are learning this way.*

Consistent with this quote, the GSIs often spoke in meetings about the benefits of sharing their ideas with their peers and having a space to think more deeply about their teaching.

Case Study: Wong (Little Change)

In contrast to Beth, Wong showed little change in his teaching practices over the semester. Although Wong appeared to genuinely try to incorporate new practices into his teaching, the way he did so often did not seem to have the desired effect. One example of this is Wong's use of questioning practices. In his post-semester interview, he noted:

> *It's really important [that students speak during recitation]. Otherwise they just keep quiet and, I think that if they speak then they really think about the problem. Otherwise they can just doze off and sit there not doing any work. They have to speak, because they have to think, and then they can speak.*

This quote illustrates that Wong is thinking about getting his students engaged with the mathematics by having them talk during recitation. However, Wong's attempts to elicit student thinking were often ineffective, and when students did not respond quickly, Wong was likely to simplify his questions or answer them himself. Consider the following exchange from Week 13. Students were discussing the PAR problem, which focused on defining a function F using an integral, where F turned out to be the natural logarithm (this was not discussed in their regular lecture sections). The class was discussing a sample explanation as to why F(0.5) is negative: We can rewrite the integral from 1 to 0.5 as the negative of an integral from 0.5 to 1. Since 1/x is positive from 0.5 to 1, we accumulate a positive area, which we multiply by negative 1 to get a negative result.

This (correct) explanation describes that because F is defined by an integral that starts at 1, when an input value less than 1 (0.5 in this case) is chosen, the accumulated area is negative because the area is accumulated backwards. This property of integrals was given in the students' textbook, but was not given a specific name. Nevertheless, in the following exchange Wong attempted to get students to name the property (line 3). When students did not, Wong began to answer his own questions (lines 6 and 8):

1. Wong: What does #3 say?

2. Student: If you multiply by -1 you flip the bounds.

3. Wong: What property of integrals does #3 use? [No response from students.]

4. Wong: If I have this [points to the original integral], and I flip the bounds, then what do I do to fix this?

5. Student: -1 on the outside.

6. Wong: So this equals the negative and you flip the bounds. So F(0.5) should be 1 to 0.5 of this [points to the integral from 0.5 to 1 written on the board], and we don't want this [points to the lower bound of the integral] to be less than this [points to the upper bound of the integral], so we flip.

7. Wong: What can you conclude about this thing? [No response from students.]

8. Wong: If you look at this graph, you can see…

In this exchange, it appeared that Wong was attempting to elicit student reasoning. He began by asking students about the sample solution (line 1), rather than explaining it himself. When the student responded that the bounds are flipped (line 2), Wong asked a follow-up question (line 3). Students likely did not know how to respond, because the integral property was not named in their textbook. When students did not respond after a few seconds, Wong asked a question that had already been answered by the student (line 4). Accordingly, the student once again said that there is a negative 1 on the outside of the integral (line 5). At this point Wong began to explain in more depth (line 6), even though his second question did not elicit any new information. After his explanation, he asked students what they could conclude (line 7). When students did not respond after a few seconds, Wong began to explain the conclusion himself (line 8).

This pattern of discourse was common in Wong's class; even though Wong attempted to elicit student thinking, he often began to answer his own questions. Later in the same lesson Wong was discussing a problem that involved bounding a complicated function using simpler functions and using the comparison property of integrals to find a bound on the value of the unknown function. In this short exchange, Wong asked six questions of the students (lines 1, 3, 4, 5, 6, and 8), but in each case the students did not respond and Wong moved on:

1. Wong: But what is the function? [No response from students]

2. Wong: The key thing is finding h and g. We need to find those.

3. Wong: Let's start with x2. x is between…what's the upper bound for x? [No response from students]

4. Wong: You can't have 1 less than ½. It should be this way, but why? [No response from students]

5. Wong: Everyone agrees with this, right? [No response from students]

6. Wong: So now you take the integral, and you have exactly the same term here. What's on the left hand side? [No response from students]

7. Wong: So you've got the bounds, and this is true.

8. Wong: Any questions about part (a)? [No response from students]

It appears that students may have realized that Wong was likely to answer his own questions. Still, Wong indicated that he did genuinely value student input. For instance, he said:

> *Previous semesters I just went through problems and gave them time to work on them and went over the problems. Now I think am more focused on their reaction. I want to let them tell me their ideas and share their ideas. It's not me showing solutions on the board. That's kinda boring if I just write down the solution, "this is right, you should do this." I don't think that works a lot.*

Wong describes his focus "on their reaction," indicating his attention to how students are following along the discussion. This is consistent with Wong's continuous questioning to elicit student ideas and check in with the students. However, it is possible that Wong did not have the awareness that his questioning was not effective, so he continued to move on and ask more questions. Wong struggled to use wait time to manage his discussions, while other GSIs such as Beth were able to incorporate it into their practices.

Another aspect of Wong's practices that showed his desire to be more student-centered was bringing students to the board to present. For instance, during his week-two recitation session, Wong had a total of three students come to the board to write their solutions to worksheet problems. However, after the students wrote their solutions, he had them sit back down without explaining their ideas, and then Wong went through the problems by explaining the students' work. This was a frequent pattern of behavior for Wong, even though he was given explicit feedback about having students explain their reasoning and having students do the talking was a major focus of the working group sessions. During week 10, Wong once again had three students come to the board to write their solutions. He explained the work of the first two students, but then had the third explain his reasoning himself. This resulted in the only major discussion that was observed throughout the semester, with at least six different students participating. In the post-semester interviews, Wong was asked explicitly about this. He responded:

> *that depends on what the problem is that they are doing. For hard problems you have a lot of things to say, so I will let them do it. Otherwise I will say it for them, because you're just following step by step to get a result.*

The interviewer followed up, asking if it mattered who does the explaining. Wong responded:

> *For the problem I want them to explain I always prepare some questions for them. Once we had a Riemann sums problem. You have to compute the limit of a Riemann sum and some student got confused about when to plug in the limit into the function. I asked him, what if you have infinity over infinity, and can you plug in the limit to the top and the bottom separately and then compute the result. Then he got confused. So I gave him the example of n over n, and the top terms goes to infinity, the bottom term goes to infinity, but the limit is just one. So it's kinda like l'Hopital's rule. I do the problem before the recitation to figure out where they will get stuck and confused.*

This episode described by Wong was shown above, in which many students participated in the whole class discussion. The quote indicates that Wong was thinking about student thinking, because he was preparing questions in advance and trying to anticipate how students will respond. Later in the interview, Wong provided more insight into why he brought students to the board:

> *I put students on the board to do the problems so that they will feel pressure and they can be more careful on the problem rather than just try something and playing with phones and doing some calculations.*

This quote explains some of the mismatch between Wong's stated desire to have students speak and his actual practices of how he brings students to the board. It appears that Wong has multiple goals for engaging students in such practices, which may not be consistent with one another. This combination of traditional and reform views was present in other aspects of Wong's interview, such as when he described the purpose of lecture and recitation:

> *I think they learn the most stuff from the lecture from the instructor, but for math they have to learn some practical stuff, that's why we have recitations.*

The interviewer followed up by asking Wong if he thought recitation or lecture was more important than the other:

> *I think it's hard to say. But if the student is a good learner, he can learn himself, then the recitation is not that important compared to the lecture. But most students are not that willing to learn, so we have to do the recitation so that they can do practice and get a taste what the real problems look like.*

This response is consistent with Wong's view that students should come to the board to "feel pressure." It appears that he viewed recitation as an opportunity to get students to do work, but has not fully embraced the notion that students need to do the talking to do the learning.

DISCUSSION

Traditionally, mathematics GSIs received little support to grow as teachers, but this trend is beginning to change (Deshler et al., 2015). As this paper highlights, all six GSIs who were contacted were interested in growing as teachers, and some GSIs were able to improve their practices with relatively little support (six one-hour meetings). Participating GSIs implemented PAR in their recitation sections, and discussions built around PAR helped the GSIs develop student-centered pedagogies more broadly. This paper extends prior work on PAR, showing that with relatively little support, even new GSIs could implement it effectively as a part of their teaching. Moreover, PAR was a promising tool for supporting the development of GSIs.

This paper focused on the development of GSIs as instructors (research question 1), evidenced both by classroom observations and interviews. The largest changes in practices were for Beth and Amy, both who were female GSIs with relatively little teaching experience. Beth was chosen as a focal instructor, to illustrate this growth. At the beginning of the semester, Beth's classroom was almost entirely teacher-focused. Over time, Beth began to elicit student reasoning and help students engage with each other; she discussed this in her post-interview, indicating alignment between beliefs and practices. The working group helped Beth reflect on her teaching and learn techniques to support students.

In contrast, the male GSIs who had more teaching experience showed less changes in their practices. Although the sample size was small, it is also possible that gender played a role in GSI receptiveness; given their historically marginalized status in STEM, the female GSIs may have been more willing to put in additional effort to adopt the new methods, but data were not collected to investigate this assertion. Despite relatively little change in their practices, both Adam and Wong noted that they were interested in pursuing academic careers and wished to improve in their teaching. In contrast to Beth, Wong did not benefit as much from the working group, likely due to a number of factors. First, Wong held a complex set of beliefs about learning, which contained a mix of traditional and reform ideas. Moreover, even when Wong did attempt to implement more student-centered practices, he sometimes had difficulty doing so. This may have been due to both a lack of awareness of the impact of his practices, and also a basic struggle with the mechanics of teaching. Because Wong had more teaching experience than Beth, he may have been more likely to fall into the routine of old habits when new practices did not work for him. This suggests that attending to GSI's well established teaching practices in addition to their beliefs may be required to support their growth. This may be achieved by allowing the GSIs to reflect on their teaching approaches more regularly throughout the semester. Another approach would be to have GSIs conduct peer observations, to give them additional models for teaching. Finally, Wong's status as an international student may have been a factor (Boman, 2013).

Although there was no significant impact on student outcomes in the course (research question 2), there was evidence of progress towards achieving the goals of PAR: students in the PAR sections were more likely to take the final exam than those in other sections (persistence), they were pushed to explain their reasoning in more depth (communication), and they began to work with their peers more in class, both in whole class discussions and in partner and group work (collaboration). The lack of impact on student outcomes contrasts with prior research, which showed significant improvements due to PAR (Reinholz, 2015a). There are a number of differences in the present study that may explain this. First, in the original study, the experimental instructors were more familiar with student-centered pedagogies to begin with, so they used PAR with greater fidelity earlier in the semester. Second, instructors in the original study taught all four days of instruction, whereas the GSIs in the present study only interacted with students once a week. This may have resulted in a mismatch for students between their more traditional lectures and more student-centered recitation sections. Finally, as Amy noted in her interview, the GSIs also began to implement student-centered teaching practices in their non-PAR sections, which may help account for the lack of significant differences.

The present study has a number of limitations. First, only four GSIs were studied, so there is a question of generalizability of the results. Second, these were GSIs recommended by the department chair, so there is a question of whether GSIs who were not interested in the project would be as successful. Third, the GSIs in the study were frequently observed by one of the researchers, which was the primary time commitment of the project. The limitations are being explored in a follow-up study with more GSIs who are required to participate as a part of their teaching assignment, and who will conduct peer observations and feedback to each other using a method similar to PAR.

Although these case studies are from calculus, the lessons learned appear to be generalizable. The use of high-leverage practices to support professional development of teachers has been used across mathematical contexts (e.g., Franke et al., 2001; Herman et al., 2014), and is likely a generalizable method for supporting teacher learning. Moreover, PAR has been used successfully in other STEM contexts (e.g., physics), and the idea of using peer review to support learning has a long history in writing. Because many disciplines use discussion as a medium for promoting learning, and PAR involves features to deepen and enhance the quality of discussions, it is likely that these methods can be adapted to a variety of contexts. The application of this approach to other disciplines, with appropriate modifications, appears to be a productive area for future research.

ACKNOWLEDGEMENTS

We thank Lina Chopra Haldar and Elissa Sato for their helpful feedback on earlier drafts. This work was funded by the Association of American Universities and the 2014-2015 President's Teaching and Learning Scholars Program at the University of Colorado, Boulder.

REFERENCES

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education, 73*(1), 94–122.

Ball, D. L., Sleep, L., Boerst, T. A., & Bass, H. (2009). Combining the Development of Practice and the Practice of Development in Teacher Education. *The Elementary School Journal, 109*(5), 458–474. http://doi.org/10.1086/596996

Boerst, T., Sleep, L., Ball, D. L., & Bass, H. (2011). Preparing teachers to lead mathematics discussions. T*eachers College Record, 113*(12), 2844–2877.

Boman, J. S. (2013). Graduate Student Teaching Development: Evaluating the Effectiveness of Training in Relation to Graduate Student Characteristics. *Canadian Journal of Higher Education, 43*(1), 100–114.

Bressoud, D. M., Carlson, M. P., Mesa, V., & Rasmussen, C. (2013). The calculus student: insights from the Mathematical Association of America national study. *International Journal of Mathematical Education in Science and Technology, 44*(4), 685–698. http://doi.org/10.1080/0020739X.2013.798874

Deshler, J. M., Hauk, S., & Speer, N. (2015). Professional Development in Teaching for Mathematics Graduate Students. *Notices of the AMS, 62*(6), 638–643.

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38*(3), 653–689.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N. , Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 201319030. http://doi.org/10.1073/pnas.1319030111

Hauk, S., Chamberlin, M., Cribari, R. D., Judd, A. B., Deon, R., Tisi, A., & Khakakhail, H. (2009). Case story: Mathematics teaching assistant. *Studies in Graduate and Professional Student Development, 12,* 39–62.

Herman, J., Epstein, S., Leon, S., Matrundola, D. L. T., Reber, S., & Choi, K. (2014). *Implementation and Effects of LDC and MDC in Kentucky Districts* (CRESST Policy Brief No. 13). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hufferd-Ackles, K., Fuson, K. C., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education*, 81–116.

Kogan, M., & Laursen, S. L. (2014). Assessing Long-Term Effects of Inquiry-Based Learning: A Case Study from College Mathematics. *Innovative Higher Education, 39*(3), 183–199. http://doi.org/10.1007/s10755-013-9269-9

Kung, D. (2010). Teaching assistants learning how students think. Research in Collegiate Mathematics Education VII, 7, 143.

Kung, D., & Speer, N. (2009). Mathematics teaching assistants learning to teach: Recasting early teaching experiences as rich learning opportunities. *Studies in Graduate & Professional Student Development, 12*, 133–152.

Lutzer, D., Rodi, S., Kirkman, E., & Maxwell, J. (2005). Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall CBMS 2005 Survey. *Providence, RI: American Mathematical Society.*

Michaels, S., O'Connor, M. C., Hall, M. W., & Resnick, L. B. (2010). *Accountable Talk® Sourcebook*. Pittsburgh, PA: Institute for Learning.

Miller, K., Brickman, P., & Oliver, J. S. (2014). Enhancing Teaching Assistants' (TAs') Inquiry Teaching by Means of Teaching Observations and Reflective Discourse. *School Science and Mathematics, 114*(4), 178–190. http://doi.org/10.1111/ssm.12065

President's Council of Advisors on Science and Technology. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. *Executive Office of the President.*

Reinholz, D. L. (2015a). Peer-Assisted Reflection: A design-based intervention for improving success in calculus. *International Journal of Research in Undergraduate Mathematics Education*. http://doi.org/10.1007/s40753-015-0005-y

Reinholz, D. L. (2015b). The assessment cycle: a model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, 1–15. http://doi.org/10.1080/02602938.2015.1008982

Speer, N., Murphy, T., & Gutmann, T. (2009). Educational research on mathematics graduate student teaching assistants: A decade of substantial progress. *Studies in Graduate and Professional Student Development, 12*, 1–10.

Treisman, U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal, 23*(5), 362–372.

Wright, M. C., Bergom, I., & Brooks, M. (2011). The Role of Teaching Assistants in Student-Centered Learning: Benefits, Costs, and Negotiations. *Innovative Higher Education, 36*(5), 331–342. http://doi.org/10.1007/s10755-011-9197-5