Writing can play an important role in the teaching and learning of mathematics; the field itself comprises vocabulary, concepts, and symbols, the understanding of which provide a foundation for grasping the thinking processes involved in mathematics. Student understanding of math can be probed through writing, not just so that students can successfully solve problems but so that they can also understand underlying theorems, definitions, and proofs. Kittleson and Southerland have argued that constructing sound arguments and justifying reasoning in mathematics is a key to providing well-articulated solutions (verbal or written) and that communication, whether in writing or verbal, shapes understanding and plays a role in knowledge construction. The building of mathematical understanding has stretched to K-12 levels, as suggested by the Michigan Department of Education’s endorsement of low-stakes write-to-learn (henceforth WTL) approaches such as journaling, imagery and visualization techniques, concept mapping, and vocabulary documentation (“WAC”). Similarly, Kelly McCormick, of the University of Southern Maine, reports that Maine pre-service mathematics teachers are urged to explore the instructional value of writing mathematical explanations. The organization Mid-Continent Research for Education Learning (Urquhart) has indicated that writing can be an essential aid in meeting a key objective of the National Council of Teachers of Mathematics—the deepening of students’ mathematical understanding.

At the post-secondary level, Reynolds, Thaiss, and others, in an article published in 2011, reviewed the research on writing as a pedagogical tool in STEM disciplines. They posit that the relevance of writing to STEM education has been under-studied perhaps due to other priorities but also due to the reluctance of STEM disciplines to stake a claim in writing, an area they view as outside their expertise. Despite this disciplinary reluctance, Reynolds, Thaiss et al. were able to construct a database of over two hundred articles pertaining specifically to WTL pedagogy in STEM disciplines at the university level; they report that writing provides pedagogical opportunities that extend beyond communicating or performing knowledge in polished term papers. Among the studies they reviewed was work by Fleron and Hotchkiss, who
argue for seminars involving writing at both the introductory and capstone level of mathematics instruction; such courses, they claim, help students conceptualize and unify mathematical knowledge, making them better contributors to the mathematics community. Another study, by Ganguli, examined twenty-five remedial mathematics students who made substantial strides in their mathematical thinking as a result of doing short in-class writing about mathematical concepts. Reynolds and Thaiss argue especially for the value of experimental studies of WTL in STEM, such as that provided by Cummings et al., who found that introductory physics students, regardless of writing ability, showed greater gains in understanding physics concepts than did peers from a comparison classroom where WTL was not used. In psychology, which in the local setting is considered more STEM than social science, Gingerich et al. (2014) found that performance on examinations was not only improved but retained six weeks after the course was over in classrooms where short writing was done. As these examples suggest, both descriptive accounts and experimental studies point to the value of writing in STEM disciplines.

The use of narrative, as a subset of writing integration in disciplinary contexts, has been a subject of particular interest to STEM scholars. However, narrative has often been positioned less as tool for learning than as a functionally useful way of expanding practitioner knowledge in applied settings. Sorrell, for instance, examined instruction in narrative in a healthcare context as a means for developing nurse empathy for patients. Others have undertaken rhetorical analysis in order to describe the types of narrative employed by STEM disciplines. For instance, Stockton examined the genre conventions and tacit expectations for narrative approaches in biology, and Heckelman and Dunn analyzed the “grammar” of mathematical narrative, saying that “algebraic notation is a form of argumentation. It is not just a representational but a persuasive exercise” (76). Dieteker perhaps comes closer to the aims of our study when she argues that mathematics instruction, as seen in textbooks, involves a layering of stories, which “allows new questions to be pursued, such as, What propels this mathematical story forward? How does this mathematical story build curiosity and desire to learn what will happen? What different (and new) types of mathematical stories can we find or design?” (19).

Our study extends this discussion by suggesting the use of narrative and expository approaches in the mathematics classroom as a means for building conceptual understanding. We align with Urquhart, who describes writing about math this way:

Until I read what I have written, I don’t see the holes in my logic, the missing steps, or the rambling thoughts. Writing informs me that I only have a cursory knowledge of the content when I need a deep one. Simply put, it doesn’t let me cut corners. (4)
Similarly, Meier and Rishel describe the importance of writing in the teaching and learning of mathematics in this way:

. . . to get students to absorb mathematics, or any other subject, better, you need to have them think about, then write about, that subject. Let students tell you their thoughts, their confusions, their half-formed ideas, their frustrations and triumphs. From this, they will understand better, and we instructor[s] will, too, what the process of learning is all about. (5)

Meier and Rishel further state that “Our experience has shown that, when using short writing assignments, what we learn about the students and their understanding of the topic at hand is extremely useful in the day-to-day structuring of lectures, homework, and worksheets” (7). They argue especially for the value of narrative, distinguishing it from exposition, because narrative works “. . . in the interstices of papers, between the theorems and examples [that] tell us a story of the paper, of the proof” (89). They go on to say that when

. . . students write about mathematics, they are placing the subject in a context [that] makes sense to them. If they are going into the field, the narrative they learn is the story they will carry into their subsequent courses to inform them as to why they need to know the definition, the theorems, the proof techniques. If they are not going to continue in mathematics, they will probably forget the body of material they have studied; but, if they have written about the course they will be much more likely to hold that narrative in their memory as their record of what that area of mathematics is about. (90)

A Bridge to Our Study

For our purposes, working in the context of a large Research I institution in the Rocky Mountain West, we focus primarily on the narration of mathematical knowledge as a rational, problem-solving use of discourse that may be helpful to a broad range of students coming from varied levels of mathematics preparation. In this context we focus on student writing of conceptual understanding as might be obtained through narrations of knowledge. In particular, we look at narrative explanations, in which students write brief stories demonstrating that they understand a mathematical concept, and then we provide additional, sometimes expository, opportunities for students to demonstrate insight into the ways that they understand math concepts. With Manouchehri and St. John, as well as the National Council of Teachers of Mathematics, we argue that storytelling, as a form of discourse around mathematics, can be said to foster knowledge building and can play a key role in what students learn about mathematics as a field of study.
Our project acknowledges and draws upon the long pedagogical project associated with writing across the curriculum and distinguishing WTL from writing to communicate (WTC) (Thaiss and McLeod). WTL has been yoked with various kinds of thinking skills, going back to Janet Emig in 1977, who defined writing as a unique mode of learning, and forward to John Bean in 2011, who defined WTL as generative, exploratory, and contributory to critical thinking. Our work with WTL draws as well upon the classroom assessment work of D’Angelo and Cross as a mechanism for informing instruction. WTL, we argue for STEM program leaders, fosters the development of metacognition and the enhancement of problem-solving skills (Bangert-Drowns, Hurley, & Wilkinson; Bicer, Capraro & Capraro; Flores and Brittain; Pugalee) and deepens understanding and retention of information (Cavdar & Doe; Gingerich et al.). Writing acts as a method for internalization of and reflection upon ideas and is “self-rhythmed,” thus allowing for learning to take place at the pace of the student (Emig 96), while also providing a durable, visible record (Emig 91; Urquhart 4) that is “epigenetic” or demonstrative of the evolution of thought (Emig 96).

As this short history suggests, it has long been held that students can develop, reflect upon, and revise their ideas about a subject through writing. For math instruction at the college level, however, there is increasing awareness that writing may also provide a mechanism for constructing and deepening the knowledge of undergraduates who arrive with varying levels of readiness but aspire to STEM disciplines as majors. Writing can provide valuable formative assessment information prior to graded (summative) assessment and as such allows for a more fully pedagogical, or interventionist, approach to mathematics instruction than has been the norm in university settings where, historically, students have mostly had to sink or swim on their own.

Indeed in the context of introductory calculus, where we situate our work, students are often greeted by unapologetically rigorous university courses focused on exam-based problem-solving that some might describe as unforgiving. In such courses, it is often the case that more students “wash out” than is necessary. Our study therefore sought to address the needs of students for whom such courses do not come naturally but for whom learning is possible through hard work, engaged instruction, and student narrations of understanding. Such narrations can influence instruction by helping faculty tailor teaching to more completely meet student needs. Such efforts, when accompanied by substantial feedback and coaching from college math teachers, may help to prevent many students from becoming D, F, or W (Withdrawal) statistics.

To this end, our study involved not only the discursive processes of students but also the conversations among mathematics faculty in a faculty development setting in which writing was the trigger but faculty discussion of curricular goals and classroom approaches was the target. This effort involved a half dozen mathematics and STEM
faculty (notably physics), an assistant dean of undergraduate programs, and two faculty from the writing program who had extensive WAC experience. We undertook this collaborative effort with the hope of having positive effects upon student performance and with the aim that instruction might also be improved. This meant that not only did we envision that writing might become a persistent feature of mathematics instruction at the introductory calculus level but also that faculty, in their collaborative efforts around determining the best, most salient writing prompts, would identify their central instructional objectives and align their instruction accordingly, including the use of substantial feedback on student performance. We had the distinct advantage in this project of working toward a shared, unified purpose; as a group, we were dedicated to seeing more (and more diverse) students succeed.

Our efforts from the writing consultant side of the house involved a WAC approach that aspired to what Dr. Pamela Flash, director of the WAC program at the University of Minnesota, has described as anthropological. By this she means an approach that engages in ethnographic approaches such as listening to members of the community, observing teaching practices, and analyzing documents such as assignments and exams, rather than swooping in as writing “experts.” The type of approach Flash recommends involves digging into the curricular and cultural dimensions of the learning context, which in our case meant engaging in an effort to understand a mathematics department and its calculus program. This involved talking extensively with the faculty and administrators in order to understand program objectives, internal pressures and constraints, motivations for integrating writing, and the politics around doing so. In this view of WAC, the writing consultants for this project served as, in Flash’s terms, “external interlocutors,” whom she describes as being “academic anthropologists engaged as participant observers in the study of a field and its pedagogical objectives.”

The WAC faculty, both rhetoric and composition faculty, teamed with three math faculty, one physics faculty member, a program chair from math, an assistant dean of undergraduate programs, and an assessment expert. The group met once a week in the fall and every other week in the spring. The project was funded by a course redesign grant from the university’s Institute for Teaching and Learning. As an interdisciplinary effort, our study suggests the value of this kind of prolonged study as one mechanism for achieving extended discussions among faculty; given our expansive timeline, we were able to talk at length across disciplines about instructional objectives and how writing might be put to work for achieving those goals. Since it was possible to have ideas discussed on a generous timeline, we were able to develop writing prompts that were tried, revised, and retried. Even then, the project could have used another year because we were unable to study the next and arguably more difficult phase, which would involve inexperienced graduate teaching assistants (henceforth GTAs) as the
designated instructors of record for introductory calculus and hence as the key integrators of the initiative.

In this paper, we report not only on the findings of our writing integration in terms of student outcomes but also on the nature of our processes and their potential long-term value to participating faculty. We also discuss, although only in an introductory way, the second phase of the project where the needs of GTA preparation are addressed. While we know that our introductory efforts with lead faculty were essential, we acknowledge that all other factors being equal, the “problem” of GTA implementation of the writing-enhanced curriculum remains a crucial and difficult facet of this effort as it moves forward; we hope to offer reflection on emerging discoveries relating to the GTA factor in a future paper. For now, we offer only the slightest of insights into the importance of GTA preparation and ongoing support in mathematics instruction.

The Context and the Goal: To Get Students to Think More Deeply About Mathematics

While it is often the case that academic departments at colleges and universities are siloed from each other so that meaningful collaboration is hindered, perhaps particularly at large Research 1 institutions like our own, in our setting an energy had developed in recent years around efforts to have more conversations across academic units—conversations and activities tied specifically to the scholarship of teaching and learning. These conversations, under the rubric of High Impact Practices, had begun to partner disciplines in ways familiar to WAC scholars, linking, for instance, writing to psychology, mathematics, and physics and focusing on faculty development. To support such efforts, the provost’s office had funded an institutional summer workshop organized by the university’s Institute for Learning and Teaching. Specific faculty from across the university who shared a strong interest in pedagogy were invited to participate. The emphasis was on engaging students in active pedagogy in foundational courses with particular focus on courses with high D, F, and W rates.

Upon seeing presentations from various departments, faculty from the college of natural sciences (CNS) began to think about what they might derive from a cross-pollination of ideas with the university writing program, which was located within the English department. Facilitated by an associate dean for instruction, these participants began to have conversations, which resulted in a white paper, written jointly by STEM and writing faculty. This white paper evolved into an internal grant proposal, the purpose of which was to develop a WAC initiative focusing on WTL processes in natural science courses (mathematics, physics, chemistry, etc.) and teaming natural sciences and writing faculty. The ultimate goal of the project was to address the student experience in gateway courses across the CNS. The WTL Proposal, as it came to
be known, eventually landed on a two-pronged approach focused on developing 1) WTL activities for students and 2) faculty understanding of WTL. It was hoped that WTL would enhance student learning by:

- Deepening learning and challenging misconceptions
- Fostering critical thinking, synthesis of ideas, and transfer of knowledge
- Shifting student focus from a short-term problem-solution orientation towards a long-term grasp of concepts, including grappling with complexity and counter-example
- Developing learning skills that transcend subject areas including the development of metacognitive abilities and self-regulation of learning

Beyond this, it was hoped that the project would support a similarly motivated learning community of STEM faculty as they worked to adopt robust teaching techniques and build teaching capacity. They would:

1. Learn how to implement writing assignments in science courses
2. Better their understanding of student comprehension in gateway courses
3. Balance their expectations of student maturity and the reality of the student experience
4. Build a sustainable course redesign initiative that could be implemented across departments within the college (CNS)

To this end, faculty from the department of mathematics and, in particular, the core introductory course in calculus took the lead. This group worked with writing program faculty who were known to the math faculty because of their successful implementation of a WTL initiative in the university’s introductory psychology class, which was also part of the CNS and was a high-enrollment core curriculum course much like the introductory calculus course. The psychology initiative had shown not only that the design of writing prompts was crucial (Gingerich et al.) but also that the quality of feedback to students was essential to effective teaching through writing; the findings on feedback, in particular, provided a strong case for the professional development of faculty and GTAs (Doe et al.). The objective for the math project was similar to that of the psychology project: to develop writing prompts that would probe and deepen student knowledge on carefully scaffolded learning objectives and to develop feedback/response capabilities that would focus and prioritize teaching efforts. It was different in that writing in psychology had focused on discreet learning modules, such as student understanding of the differences between positive and negative reinforcement. For the math curriculum the goal was to deepen student understanding of foundational mathematics concepts, such as the notion of limits, so that as students matriculated they would have a solid foundation in essential areas of math knowledge.
From the beginning, the involved mathematics faculty understood that there was much work to be done to unpack long-term “baggage” about the course and legacies regarding instruction and students. There were assumptions about the readiness and work ethic of calculus students and there were also assumptions about the readiness and time constraints of GTAs. While these were challenges that had to be admitted, there were also distinct assets associated with the project: a math program leader and assistant dean for instruction, who were deeply committed to the project and were involved in high impact teaching and learning discussions at the highest levels of the university, and a pair of dedicated calculus faculty, who led the calculus program and identified as teaching and learning teacher-scholars. Along with the WAC consultants they asked, “What do we want to accomplish in Introductory Calculus?” and “Can writing help?”

Early on, Math faculty acknowledged student weaknesses as well as the challenge that a first semester calculus course poses for students. Students frequently arrive with gaps or shallow understanding of prerequisite knowledge, which makes the task of learning calculus difficult, as it relies heavily on the use of both algebra and trigonometry. In addition, poor experiences in a foundational STEM course, such as Introductory Calculus, play a significant role in the retention of STEM majors (Seymour and Hewitt). These challenges were further compounded by high $D$, $F$, and $W$ rates. But circumstances had changed locally, and where once it might have been a sign of rigor to have many students failing the course or dropping out, now such outcomes were seen as a sign of weak teaching. Therefore, the key was to develop writing prompts in conjunction with a feedback loop that would address these teaching and learning shortfalls, a task much easier said than done.

In time it also became clear that changes were needed in terms of the preparation and supervision of the GTA instructors, for whom the course director for introductory calculus was also responsible. In the past the GTAs had received little to no training in the use of evidence-based pedagogy, such as WTL, nor did they have much experience in giving feedback on assessments or activities beyond counting things right or wrong. Faculty mentors of the GTAs saw their role as protecting GTA time so that they might pursue their graduate work to greatest effect. As a practical and creative solution to these competing demands, the course director connected the WTL initiative with an established and respected mechanism in the program, a notion called “DARTs” where GTA course instructors describe student answers on math homework problems as either 1) a bullseye (completely correct), 2) on the board (partially correct), or 3) off the board (completely wrong). This mechanism became a way for her to demonstrate to internal audiences that there was a meaningful range of student responses that could be discerned through writing. Furthermore, by using
a method that was already understood by GTAs and faculty mentors, the potential problem of demanding too much GTA time was also averted.

Key Math Concepts in Introductory Calculus

The faculty engaged in this WAC conversation wanted students to derive definitions for mathematical concepts, rather than start with definitions and “solve problems.” For example, the notion of “continuity” involves a definition that students can develop on their own if given appropriate prompts and opportunity to explore. Other key concepts, such as limits, differentiability, and rates of change, were identified as topics that beginning calculus students struggle to understand. Therefore, math faculty felt students would benefit from exploring through writing rather than through the more typical mimicking of procedures, copying of definitions, and routine testing (i.e., learning methods that address topics more superficially).

Having settled on some key concepts, almost immediately the faculty began to see how teaching might be done more effectively. In part, this willingness to make improvements to instruction was reflective of initiatives already sponsored by the university and getting wide attention. The notion of the “flipped classroom,” for instance, had become an established idea where faculty were being provided with incentives for abandoning lecture in favor of creating engaged, active classrooms of hundreds of students. In the flipped classroom, lecture notes and slides were assigned for student review before class while class time was spent on applications, facilitation of student inquiry, small group projects, and elaboration into new terrain or novel problems. In this context, the long-established pedagogies of writing instruction not only fit favorably but offered variation and substance. Writing pedagogy was able to offer more than simple checking in and could be used to probe student understanding by asking them to write; in turn, the evidence deriving from such efforts became assessment information demonstrating students’ level of understanding long before the time of examination.

In this context, the CNS faculty discussed the idea of GTAs, as inexperienced instructors-of-record, positioning themselves in more of a facilitation role and doing less lecturing and modeling of problem-solving at the board than had been the norm. It was hoped that a “facilitation role” might relieve them of some of the burden of “teaching” while also deepening student learning, as was the aspiration with all flipped classrooms. The idea was settled on that GTAs would typically provide brief information in a lecture format at the beginning of class and then would spend the majority of time circulating as students worked toward their own definitions. A physics faculty member, who was sitting in on this project alongside math faculty, pointed out that in his department GTAs were not instructors of record, but were instead strictly teaching in recitations. So he began to reimagine the recitations where his GTAs were
involved. Those GTAs, he decided, might talk for just ten minutes and then turn it over to students to do problems, which would be followed by cold calling on students and whole room voting on answers. The physics professor had also imagined doing “think-alouds” wherein students would rehearse their strategies by speaking concepts aloud to classmates. In such ways did both the math and physics classrooms become increasingly dynamic locations.

As these examples suggest, from the beginning, participating faculty in this writing intervention were discussing not only content but method. Faculty explained at one point that they were concerned that GTAs were likely to teach in ways that would fall back on what they had themselves experienced. WAC faculty pointed out that math is not alone in witnessing this so-called “apprenticeship of observation” (Lortie; Grossman) which holds that no amount of professional development is ever as strong in developing the teacher as are the years of experience and observation (as a student) that have preceded the teacher’s entrance into the classroom as teacher. In fact, WAC faculty warned, given demands on GTA time, they were more likely than faculty to take shortcuts towards that which was familiar, and this might also translate into low effectiveness in giving feedback to students. The goal, therefore, was to counter that which was nearly inevitable and to realize that this would not be a simple, one-time fix. Ongoing professional development would be needed.

At the same time, the faculty themselves were re-examining their own methods. For instance, the physics professor, an award-winning teacher, started this conversation:

Physics P: I asked students to describe a notion in words and not use numbers.
Math P: Maybe the challenge is more with the intentionality of approaches and activities than with what it used.
Physics P: Still, we need the students to do the talking and report back to the group.
WAC Consultant: This is sometimes called “Write-Pair-Share.”
Physics P: I’m going to try that. I think you should try it.

As this example suggests, conversations were anecdotal but also moved toward the drawing of general conclusions across disciplines.

**Settling on Prompts**

As we approached November, the prompts were starting to come together and math faculty had begun doing a kind of pre-pilot of the questions within their own classrooms. As had been suggested by the assistant dean for instruction, the writing
The idea was to focus on big concepts and break them down into manageable parts, as with, for instance, the important concept of limits. The group discussed that the prompts might therefore be sequenced to establish the principle and then working toward greater difficulty and variety, first referring to the typical case and then moving toward elaboration into new or novel contexts, and ultimately exploring counter-examples. This process, it was hoped, would avail students to multiple iterations of similar kinds of tasks—variations on a theme—so that students would become increasingly masterful and develop flexibility as they saw how a principle works in new, even contradictory, situations.

Two types of tasks in particular stood out for 1) being simple and easy to implement, 2) providing opportunity for rich group and whole-class discussion, and 3) highlighting student misconceptions. The first task involved graphs and stories. For example, students worked with the graph of a function (fig. 1) multiple times during the semester, but in different contexts. Each time students were asked to identify quantitative characteristics of the graph (e.g., largest position, maximum velocity, etc.). In the first iteration, the function represented the position of an object with respect to time, but in the second, the scenario changed to velocity. In both contexts, the students were asked to write a story to match the graph.

Fig. 1. The graph of a function changing with respect to time.

This prompt, it was determined, could be given to students at various points during the semester. When given at the beginning of the semester, student misconceptions could be quickly highlighted, which it was hoped would motivate rich discussions and early intervention. This prompt could then also be revised such that the graph represented velocity rather than position, or the graph could be modified so that it dropped below the x-axis, offering a new level of complexity. Typical student responses to the first prompt (not all correct) are below.
Selected student responses (when posed as a position graph) included:

- A ball is rolling down a hill, then it goes up a smaller hill then it rolls down some more.
- A skydiver jumps out of a plane and opens the parachute.
- I drive home but hit road construction so have to turn back and take an alternate route, then continue home.

To explain the differing quality markers of these responses, note that while the first response is correct, it does not indicate whether or not the student is correctly interpreting the horizontal axis as time or incorrectly perceiving the horizontal axis as horizontal distance. The scenario of a ball rolling down a hill is ambiguous and highlights a misconception that the student may or may not have. The third response, however, distinguishes itself from the first one because it illustrates a clear, correct interpretation of the horizontal axis. Such misconceptions can be teased out when the scenario changes to velocity with respect to time.

Selected student responses (when posed as a velocity graph):

- A ball is rolling down a hill, then it goes up a smaller hill then it rolls down some more.
- I was running away from zombies, and I was getting tired so I was running slower and slower. Then I noticed that the zombies were getting closer, so I started running faster. Then a zombie caught me and bit my leg off, and I fell down and crawled until I could crawl no longer.

In these explanations, faculty would note that a ball rolling down a hill does not match what is illustrated in the given graph. For it to be true, the graph of the velocity of the ball would need to be below the horizontal (and therefore depict negative velocity values). The zombie story, on the other hand, does demonstrate a clear understanding of velocity with respect to time—indicating the appropriate changes of running slower and faster—and it is also a creative story that demonstrates that the student could generate a fun correlative to the principle in question, thus suggesting confidence in the student’s knowledge.

To determine if students truly understood the relationship between object movement and time, this graph task was reversed. Students were given a story and then asked to draw the corresponding graph representing the position or velocity of an object with respect to time. For example, students were given the following prompt and then asked to draw the graph representing the scenario:
It took Ryan 15 minutes to walk from his house into campus on Tuesday. He needed to get to the Weber building. On his way, he walked past a coffee shop. He decided to turn around and buy a cup of coffee. After buying a cup of coffee he continued on toward campus, but was stopped by the train for three minutes. After the train passed by, he began walking again into campus and arrived at the Weber building. In the axes below, draw the graph that represents Ryan’s position relative to the Weber building at time $t$.

Fig. 2 is a sample student response that correctly represents Ryan's position relative to the Weber building at time $t$.

![Graph](image)

A student who can correctly transition between different representations of mathematics (between words and graphs in this case) has a deeper understanding of concepts than a student who cannot. Further, such knowledge transfer illustrates that a student can make connections between different forms of mathematical concepts.

Additionally, in mathematics, understanding what does not “fit” or “work” demonstrates a deeper understanding of a mathematical concept. This notion is often referred to as the counterexample. To get students to think about counterexamples, we utilized true/false questions, which was the second type of writing task we employed and which involved more exposition than narration. Specifically, we asked students to consider statements, and if a statement was true, they had to explain why it was true. If
a statement was false, they had to provide a counterexample. If utilized in a meaningful way, math faculty hoped that such tasks might promote thoughtful discussion and deeper knowledge.

This second approach also probed students’ increasing sophistication with broad topics and variation and counterexample. We experimented with synthesis prompts that took the ideas that students had been exposed to and queried about via narrative prompts and brought them together through unifying, synthetic questions that took the conceptual level up a notch. For example:

1. Draw a concept map illustrating the ideas and concepts presented in this course. Explain how these concepts are connected.

2. Derivatives and integrals are two overarching topics in this course. However, we began the course with limits, which could be argued to be the theme of the course. Explain how limits connect to derivatives and integrals.

By the end of the first semester, we had a set of questions such as these that we believed we could pilot in the spring and then revise for fall instruction of GTAs.

**Pilot Study and Results:**

We piloted the prompts in two sections of Calculus I, and the sizes of these sections as well as the times of day each was offered were comparable. For purposes of the formal pilot, the classes were taught by two experienced GTAs, one of whom had six years of teaching experience. The other held a master's degree in mathematics and had three years of teaching experience. Both worked directly and substantially with the course director. Both GTAs met their classes four days a week, gave the same exams, assigned the same weekly homework, and achieved similar retention rates.

In the experimental group, referred to as the “Writing Section” (N = 36), students were provided minimal lecture time, and class consisted mainly of group activities with writing and discussion. In terms of writing, students in this section were given WTL activities. For group activities, groups were carefully constructed to achieve balanced ability groups. In the control group referred to as the “Standard Section” (N = 33), students were exposed primarily to traditional lecture with some worksheet activities provided that could be done individually or in groups, consistent with the course's traditional delivery.

**Sample Writing Prompts and Responses**

Students enrolled in the experimental section engaged in several discussions and writing activities. Students also participated in oral assessments before each of the four exams (three midterms and one final exam).
The prompts asking about graphs and stories as well as the true-false statements were among several of the prompts used for writing and discussion. Oral assessments were given to students in groups of three to four and were comprised of a variety of questions intended to prepare students for upcoming exams. Student responses to oral assessments provided additional insight into students’ misconceptions and knowledge gaps, which would provide guidance in developing review content for exam preparation. Questions for oral assessments ranged from the procedural, such as taking the derivative of a function, to the conceptual, wherein students were asked to discuss the connection between continuity and differentiability. Students responded to questions as a group. For example, prior to exam two, students were given a graph (see fig. 3) and a list of stories. Students had to determine if a story represented position, velocity, or acceleration. One of the stories was: “Tatiana is jumping on a trampoline until her foot slips and she falls to the ground.” Here is the conversation that followed:

Fig. 3. The graph of a function with respect to time, $t$.

Facilitator: Does the graph represent position?
Student 1: It’s not position because she doesn’t hit zero when she falls down.
Facilitator: Could it be velocity?
Student 2: No. Just because she breaks her foot and immediately stops.
Student 3: It’d be oscillating.
Facilitator: For velocity, should the graph drop below the x-axis?
Discussions such as this provided us with insight into students’ knowledge of the physical quantities’ position, velocity, and acceleration. The discussion also provided opportunity for a misconception to be resolved. By this point in the study, students had become comfortable weaving among the writing of stories to depict mathematical phenomena, the depiction of stories onto graphs, and the verbal narration of graphs. Discourse in varied forms had become mixed and complementary.

**Findings**

Students in the writing (experimental) section outperformed students in the standard (control) section for every exam (see table 1). The differences were at a level of statistical significance on the first exam and on the final exam, and the effect sizes (meaning the actual difference in performance as measured in terms of real world effect) were in the middle range, meaning that student grades were meaningfully influenced by these efforts and were not trivial differences—say, for instance, moving from a C to a C+ performance.

<table>
<thead>
<tr>
<th></th>
<th>Writing Section N = 36</th>
<th>Standard Section N = 33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam 1:</td>
<td>82.8 ( p = 0.03, d = 0.53 )</td>
<td>76.1</td>
</tr>
<tr>
<td>Exam 2:</td>
<td>67.9</td>
<td>63.3</td>
</tr>
<tr>
<td>Exam 3:</td>
<td>64.2</td>
<td>60.8</td>
</tr>
<tr>
<td>Final:</td>
<td>71.6 ( p = 0.03, d = 0.56 )</td>
<td>62.1</td>
</tr>
</tbody>
</table>

One concern that had been voiced about this language-based intervention was that students might sacrifice “procedural skills” or fail to demonstrate mathematical processes and computational skills if their attention was focused on writing narratives or explanations. However, findings demonstrated that there was no loss of procedural skill. Instead, there were gains not only in conceptual understanding, at a statistically significant level, but also in procedural knowledge, although the latter was not at a statistically significant level (see table 2).
While students in the writing section did not demonstrate the full extent of deep understanding that we aspired to obtain through the intervention, they nonetheless had responses that were better than those from students in the standard section.

Written responses were coded based on how students used mathematical ideas, vocabulary, and notation. So additionally, it was noted if a response contained pronouns in such a way that the corresponding noun could not be identified. Examples of correct and incorrect use of ideas, vocabulary, and notation are provided in table 3.

Table 3. Examples of qualitative data coding.

<table>
<thead>
<tr>
<th></th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideas</td>
<td>“The equation [for] ( f'(a) ) is the slope of the tangent line at a given point.”</td>
<td>“The equation is for tangent lines and secant lines at any given point on the graph.”</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>“The equation describes the slope of the tangent line at point ( a ).”</td>
<td>“A secant line is the slope between ( a ) and ( b ).”</td>
</tr>
<tr>
<td>Notation</td>
<td>((a,f(a))) (in reference to an ordered pair)</td>
<td>( f(a) ) (in reference to an ordered pair)</td>
</tr>
</tbody>
</table>

Our findings from coding student responses show that the writing section had a higher percentage in the use of correct ideas, correct vocabulary, and correct notation than the standard section. In addition, the writing section had a lower percentage of incorrect ideas, incorrect vocabulary, and incorrect notation than the standard section.

The writing section also had a lower percentage use of pronouns. An overuse of pronouns often indicates that students are trying to use terminology but do not yet know how to link together ideas correctly. For example, a student may state, “The equation is the slope. The limit goes to 0, so it approaches the tangent line.” While \( \text{limit} \), \( \text{slope} \), and \( \text{tangent line} \) may be correct vocabulary words to include in a response involving the discussion of the definition of the derivative, the use of the word \( \text{it} \) is unclear and does not indicate that the student understands the meaning of the use of a limit in terms of slopes of secant lines and a tangent line. Using correct vocabulary in
a meaningful way rather than making vague use of pronouns indicates that a student has a clear understanding.

**Coda: Implications for GTA Training**

Our discussion of findings thus far has focused on student performance and faculty development. Like Grawe and Rutz, “Our experience has convinced us that engaging faculty directly in the assessment of student work provides the impetus for curricular change” (14), but we acknowledge significant constraints moving forward as implementation of the WTL integration moves into its next, full implementation phase. We hold that the effectiveness of this WTL initiative on student performance was likely caused by a combination of both deepened student engagement with calculus material and enhanced teaching efficacy. However, imaginings about a writing-enhanced math curriculum must take into account who will do the implementing of such efforts once the approach is settled upon. In other words, if the experienced and highly motivated faculty who undertook this project will not be its deliverers in classrooms, does the initiative have much hope of succeeding? What problems might be anticipated? And what might be done to address these challenges?

As stated earlier, our investigation was undertaken by experienced faculty with high interest in the scholarship on teaching and with deep knowledge of student strengths and shortcomings. These faculty went into the project with high motivation and intention. They knew there were problems in student learning and they had already spent some considerable time improving the curriculum, even investing in teaching expertise by hiring a tenure-track pedagogue whose main research interest was in thinking about improvement to student outcomes. Even then, the project of identifying foundational objectives for the introductory calculus course took the better part of a semester, with faculty meeting weekly. Once the group had settled on objectives, they then devoted weeks to clarifying the prompts that would be used to probe understanding of both small and large (synthetic) concepts. As the prompts were being developed, these experienced and highly motivated faculty tested and refined them. In addition, these faculty had already embraced active teaching methods which included the use of active (flipped) classroom techniques such as paired conversations, discussions, and collaborative, small-group problem-solving. Faculty were also including more and more student verbalization in pairs and small groups as well as instructor facilitation and they had become conscious facilitators who floated around the room to listen and intervene with particular students and groups of students. They were equally engaged with ongoing assessment that allowed them to adjust their instruction. They learned through the WTL intervention that they also had an obligation to provide ongoing formative feedback to students.
In moving forward with this initiative, it is important to note that most of the instruction was scheduled to be provided by GTAs, as is often the case in large R1 institutions. A key question then became, to what degree could GTAs be expected to deliver this kind of curriculum? We posited that while it is convenient and even seductive to assume that GTAs are ready for this work, generally speaking they are not ready for it and need a great deal of mentoring and supervision. Too often it is assumed that GTAs arrive ready to teach and to provide feedback on student work, having even been described by some scholars as expert graders (Pare and Joordens). Yet most GTAs lack preparation for teaching (McKeachie) and are generally unfamiliar with the scholarship on their own development as teachers and scholars (Nyquist, Abbott, Wulff, and Sprague). They are also quite early in career development (Eble), may be challenged by their proximity to undergraduates in terms of age and authority (although, as Doug Hesse suggests, proximity in age can also be a benefit), and are likely to experience interference between their roles as teachers and graduate students (Duba-Biedermann; Doe “Lived Experiences”). Furthermore, most are still developing a teaching identity (Schempp, Sparkes, and Templin), and all are still developing disciplinary expertise. The literature further establishes that there is generally a strong need for GTAs to be socialized into their roles and responsibilities as faculty—skill sets that are by no means tacitly understood (Slevin; Braxton, Lambert and Clark).

None of these are new challenges. In fact, in 1987, at a conference about revising calculus instruction for a new century, organized by the National Academy of Sciences and the National Academy of Engineering, Bettye Anne Case and Allan C. Cochran observed that “The role of teaching assistants . . . in the teaching of calculus . . . is of serious concern” (76). Their preparation, they argued, requires considerable expense in terms of time and resources and requires ongoing training and the preparation of demonstrations and materials (77). Similarly, Eison and Vanderford argue for ongoing pedagogical instruction of GTAs, including observations of teaching, engagement with teaching theory, and faculty development in regard to the unique features of each discipline’s pedagogical traditions. On the bright side, research by Doe and Gingerich suggests that while GTA grading and responding remains short of the standard set by experienced faculty, even when accompanied by rigorous training, GTAs do show impressive growth with just one semester of carefully planned pedagogical instruction.

In the current study, a first semester of piloting was conducted by faculty and a second by experienced GTAs with substantial teaching experience. It seems reasonable to assume that different problems will surface when new GTAs are attempting to integrate writing for purposes of learning mathematics. Given the challenges facing GTAs in every setting, it is likely that a quite substantial task remains to be addressed at the close of this project in regard to preparing calculus GTAs for the work that lies
ahead. We therefore conclude our paper by providing some early observations about how GTAs fared in this context during the first year.

**GTAs in the First Semesters of Implementation**

GTA training improvements were first implemented prior to the start of the fall 2015 semester and have continued since. The work with GTAs began with the assumption that they need a realistic understanding of the students who will be in their classrooms and have the ability to provide instructional support for students who are learning to communicate and write mathematics at the university level. We find that GTAs typically begin with the assumption that their Calculus I students are just like them and can easily understand mathematical concepts and have the natural ability to connect mathematical ideas. However, this is not the case, of course, and GTAs typically discover the actual situation when grading their students’ first exams. Unfortunately, though, this is often too late for students who can quickly become discouraged by early failure and wash out prematurely. To address this problem, we sought to help GTAs develop a better understanding of their Calculus I students and therefore developed WTL activities with corresponding sample student responses as part of training activities.

GTAs worked through a variety of student activities, including WTL prompts. They then discussed potential student responses as well as implications for the classroom. This exercise stimulated an extensive discussion around pedagogy with questions such as “How should a task be implemented?”; “What are good facilitating questions?”; and “What are possible student misconceptions?” Following a rich discussion about how students might respond to questions and prompts, GTAs were provided with student writing that showed how students actually respond. This led to a deeper discussion about the students that take Calculus I—their typical knowledge gaps, mathematical misconceptions, and weaknesses in notation and vocabulary. GTAs were predictably surprised by the examples that illustrated the range and content of student writing, and these made them, we believe, more mindful of their instruction starting on the first days of classes rather than only after the first examination.

Of course, we were hopeful that such engaging conversations with GTA instructors-of-record would lead to immediate improvements in instruction that would in turn have a positive impact on student success. However, we also realized that another factor was at play: turnover in Calculus I GTAs occurs every semester, so sustaining pedagogical change immediately presented as a challenge. To address this issue, the tenure-track calculus course director successfully argued for creation of a calculus center, which very much resembled a writing center built for mathematicians. Opening in the fall of 2016, the CSU Calculus Center is positioned to lead efforts to improve instruction in all calculus courses, through GTA and faculty training, and is intended to build upon the recently developed efforts of this writing integration.
project. Specifically, training will focus on the use of evidence-based practices, such as WTL, and will include student data and responses as a way to give meaning to such practices. Students will lead the conversations, just as they do in writing center work, and additionally, the CSU Calculus Center is envisioned to provide GTAs and faculty with a repository of rich tasks and activities. Envisioned as a working pedagogical research center through which data can be collected in an ongoing way to assess the impact of these efforts, the CSU Calculus Center’s primary purpose is to help the student—to develop better mathematicians, not just better math exams—echoing Stephen North’s charge that writing centers develop better writers, not just better writing (438). It is hoped that research and scholarship on teaching and learning coming out of the center will also affect faculty instruction, much as writing center pedagogy has influenced writing classrooms.

One important outcome of this project is that a small professional learning community has begun to flourish between faculty in the university writing program and faculty in mathematics. Collaborative efforts between these faculty will continue with support and facilitation from the calculus center and from WAC-interested rhetoric and composition faculty, perhaps even connecting the CSU Calculus Center with the CSU Writing Center for cross-disciplinary professional development. Some joint GTA training efforts are already planned, and WTL workshops for mathematics will be developed and implemented. The Calculus Center will also act as a vehicle for getting other STEM faculty involved in writing and learning activities, similar to what has occurred in Carleton College’s QuIRK program (Carlton College). Our collaboration offers an opportunity to raise awareness about and provide education on the power of using writing in STEM disciplines through faculty professional development in the Calculus Center, led by both English and Mathematics faculty.

Special Acknowledgments

The authors would like to acknowledge the following individuals and groups at Colorado State University for their involvement in this project: Kathleen (Kate) Kiefer, Professor Emeritus of English and University Distinguished Teaching Scholar; Lisa Dysleski, Assistant Dean of Undergraduate Programs, College of Natural Sciences; Hilary Freeman, Senior Teaching Faculty in Mathematics; Brian Jones, Senior Teaching Faculty in Physics and Director of Little Shop of Physics; The Institute for Learning and Teaching (TILT); Simon Tavener and the College of Natural Sciences, Professor of Mathematics and Associate Dean for Academics; Anton Betten, Associate Professor of Mathematics; Julie Maertens, Research Associate in Psychology and Evaluator for the CSU STEM Center.
References


Flash, Pamela. Email with co-author, Sue Doe. 7 July 2016.


