Educational Studies in Mathematics (2019) 101:35–50
https://doi.org/10.1007/s10649-019-9877-0

Forged in failure: engagement patterns for successful students repeating calculus

Rebecca Dibbs[1]

Published online: 19 January 2019

Abstract
Although there is extensive research on attrition in gatekeeper courses and students' cognition about calculus concepts, there is one population in introductory calculus that remains understudied: those who failed their initial course and chose to repeat it rather than change majors. These students can provide insight into overcoming poor mathematics affect and major persistence. This case study follows eight students repeating calculus from their second try at undergraduate calculus until they graduated or left the university; six graduated with either a mathematics major or mathematics minor. While participants identified several reasons for their success in the repeated course (processing their initial failure, having a better instructor in the repeated course, and participating regularly in the formative assessments), only participation in formative assessment led to the long-term cognitive and behavioral engagement required for long-term success.

Keywords Calculus · Leading event · Patterns of engagement

Forged in failureEngagement patternsCalculus

1 Introduction

Since it is the gatekeeper to many high-paying jobs, mathematics is highly valued worldwide (Andersson, Valero, & Meaney, 2015; Esmonde, 2009; Quintos & Civil, 2008); however, students who have poor experiences in mathematics classes often change their educational goals (Braathe & Solomon, 2015), and problems with negative feelings about mathematics (negative mathematics affect) are common (Brown, Brown, & Bibby, 2008; Martínez-Sierra &

✉ Rebecca Dibbs
Rebecca.dibbs@tamuc.edu

[1] Department of Mathematics, Texas A&M University-Commerce, P. O. Box 3011, Commerce, TX 75429-3011, USA

García González, 2014; Martínez-Sierra & García González, 2016). In the USA, introductory calculus is the single biggest leak in the science, technology, engineering, and mathematics (STEM) major undergraduate pipeline. Regardless of school type, student preparedness, or class size, students who leave STEM most often do so after the entry-level calculus course (Ellis, Kelton, & Rasmussen, 2014).

Although many students who are unsuccessful in their first attempt at introductory calculus change their majors, prior research has not examined the group of introductory calculus students who fail introductory calculus and repeat the course rather than change their majors. These students may offer insight about the non-cognitive learning factors that affect students' learning of calculus concepts and their retention in calculus courses and in STEM majors. Instructors often perceive mathematics in purely content terms, and we understand quite well how students learn calculus content (Fernández-Plaza & Simpson, 2016; Roh, 2010). However, students see mathematics as a mixture of mathematical content, social norms, and affective reactions (Hardy, 2009; Imada & Ellsworth, 2011), and there are significant relationships between students' emotional perceptions of content and their motivation, problem-solving, and achievement (Frenzel, Pekrun, & Goetz, 2007; Goldin, Epstein, Schorr, & Warner, 2011; Hannula, 2006; Ingram, 2012). To improve STEM retention, mathematics instructors must also attend to the affective dimensions of student experience: when students' affect is supported and when they increase their sense of ownership of course material, students are more likely to be successful at learning mathematical content.

Students often struggle to transition to the university environment, where calculus is often their first undergraduate mathematics class. Although most students in introductory calculus believe that they are well prepared (Bressoud, Carlson, Mesa, & Rasmussen, 2013), recent educational policy in the USA has emphasized standardized testing, which promotes surface learning, and many students are unready to make connections between concepts (Gueudet, 2008; Selden, 2005; Selden & Selden, 2002). Students are thus likely to struggle with the advanced mathematical thinking needed to be successful in courses beyond calculus (Kajander & Louric, 2005; Selden & Selden, 2002; Tall, 2008).

However, there is a considerable body of research into helping students succeed in calculus, and many proposed strategies take student affect into account. The remainder of this literature review will focus on the benefits of formative assessment for students with negative mathematics affect, students with low self-efficacy, and students with poor calibration. The sample of participants in this study, who all failed calculus the first time they took it at university, offers useful case studies in how formative assessment in particular, and teaching strategies that address students' mathematics affect (including self-efficacy) more broadly, can help students improve their mathematical reasoning and prepare them for further university-level mathematics study. This study's findings may be applied to other student populations to increase student success in college mathematics courses.

Academic and social support for learning, including formative assessment, has been found to increase student success; formative assessment, which uses many teaching "best practices," increases student achievement on summative assessments (Black & Wiliam, 2009; Clark, 2011). Formative assessment is thus an effective pedagogy to help students transition to undergraduate thinking (Windschitl, Thompson, Braaten, & Stroupe, 2012)—an especially promising finding, given the relative lack of positive results with technology-infused and flipped calculus classes (Sonnert, Sadler, Sadler, & Bressoud, 2015). No matter what is being taught or how old students are, formative assessment has an effect size of around 0.5 in most quantitative studies—a moderate effect that is achievable with minimal instructor effort (Karpinski & D'Agostino, 2012).

Formative assessment helps instructors teach at students' actual level by collecting data on students' current proficiency; formative assessment also helps students understand what successful answers look like and how to produce them (Pekrun, 2006). Formative assessment may be especially powerful for students in underrepresented groups, helping to close achievement gaps (Stiggins & Chappuis, 2005). Although nearly all studies on formative assessment have been quantitative and quasi-experimental (Black & McCormick, 2010; Black & Wiliam, 1998; Clark, 2011), two qualitative studies have examined the effect of participation in formative assessment on students struggling with difficult material. In one study, math students who scored low on an aptitude pre-test were taught using formative assessment; these students outperformed high-ability students who were taught using general lesson plans from the textbook on a common unit test (Chiesa & Robertson, 2000). In fact, while using formative assessment to inform teaching decisions raises all students' achievement levels, low-achieving mathematics students show the most gains with formative assessment (Gallagher, Bones, & Lambe, 2006).

Self-efficacy, the belief a learner has about whether or not s/he can do a task, has a profound effect on achievement (Bandura, 1977). Perceived self-efficacy is correlated with many other affective variables, such as persistence, self-concept, and confidence, so lowered self-efficacy can have drastic effects on a learner's performance (Bandura, 1993). Increased self-efficacy is correlated with increases in achievement and persistence (Malmivuori, 2006). Recent qualitative studies indicate that flexible pedagogy—meeting students where they are—can help students overcome low self-efficacy (Wyatt, 2011), increasing student success and helping to retain at-risk students (Elliot & Gillen, 2013). Formative assessments increase students' ownership of their learning (Black & Wiliam, 2009), providing support for positive changes in students' self-efficacy, calibration, and motivation to learn (Black & Wiliam, 2009) and giving students a sense of control over and investment in their academic success. Formative assessment also improves students' mathematical affect by opening lines of communication between instructor and student, helping students feel that their instructor cares about their success (Ellis et al., 2014).

Formative assessment also increases the accuracy of students' beliefs about their performance, helping overconfident students accurately calibrate their abilities and efforts. Calibration is a measure of how well self-perceived achievement levels align with reality (Klassen, 2007). Most American students tend to exhibit moderate calibration, tending toward slight overconfidence (Chen & Zimmerman, 2007; Chiu & Klassen, 2010; Klassen, 2007). Poor calibration and overconfidence can be devastating to achievement. This is due to a psychological phenomenon known as the Dunning-Kruger effect, which posits that people with low abilities tend to overrate their probable performance on any quantity that is difficult to measure (Kruger & Dunning, 1999); this type of overestimation produces a false belief that no additional work is needed to master the material (Kruger & Dunning, 1999). By giving students a false sense of security, the Dunning-Kruger effect also reduces students' affective, cognitive, and behavioral engagement (Halverson, Woodfield-Porter, Graham, Hernrie, & Borup, 2013).

This study was grounded in engagement theory. Engagement theory includes affective, cognitive, and behavioral components. Affective engagement involves positive and negative reactions to content, classmates, and teachers (Fredricks, Blumenfeld, & Paris, 2004). An increase of affective engagement, by creating emotional ties between people and institutions, is thought to increase commitment and effort (Mahatmya, Lohman, Matjasko, & Farb, 2012), thus improving cognitive engagement. Cognitive engagement is when the learner invests the effort needed to understand complex concepts and master difficult skills (Mahatmya et al., 2012); affective engagement is generally required before cognitive engagement occurs. Cognitive engagement is strongly correlated with self-regulation (Halverson et al., 2013). Behavioral engagement occurs when a learner takes concrete steps to support their learning, and generally requires that a learner is already exhibiting affective and cognitive engagement (Mahatmya et al., 2012).

The purpose of this study was to understand which events for students repeating calculus fostered changes in students' affective, cognitive, and behavioral engagement. The study was guided by two questions: (1) to what degree did aspects of students' prior failing experience with college calculus support affective, cognitive, or behavioral engagement changes in the repeated course? (2) To what degree did the increased affective engagement students made in response to failure benefit them in their repeated calculus course and in subsequent semesters? This study begins to link the existing literature on transitioning to college, formative assessment, and belief change to repeating students, an understudied and at-risk group, and sketches out some ways that non-repeating students' mathematics affect may be made more positive.

2 Methods

The eight participants in this study were students at a mid-sized doctoral granting public university in a rural region of the western part of the USA who were enrolled in one of three offered sections of Calculus I in the spring of 2010. Approximately 33% of the students enrolled at the institution either speak a native language other than English or are the first student in their family to attend a university. Five of the participants (Andrew, Jim, Elizabeth, Sharon, and Corwin) were taught by other instructors, and three participants (Amanda, David, and Kendra) were drawn from the section taught by the author; both instructors responsive for the two sections of calculus had comparable experience, course schedule, and syllabus. Both courses were closely coordinated by a senior faculty member, both instructors used identical formative pre-lecture worksheets before introducing a new section of the book, and both spent approximately 25% of instructional time on small-group student work and the other 75% of the instruction time in lecture. Both instructors used the same pacing guide used by both classes and wrote common exams.

After obtaining approval from the Institutional Review Board, all students enrolled in introductory calculus were invited to participate in the study; during the third week of class, all students were given a short survey that asked for basic information about demographics and their history with mathematics. Of the 83 students that completed the survey, 17 were repeating calculus and consented to be interviewed.

Consenting participants were selected to maximize the variation in responses. When participants gave the same reason for repeating calculus (e.g., multiple students who failed the same section of the course in the fall), the participant that was chosen for the interview was selected to maximize participant variation in age and major. Seven participants failed the standardized AP Calculus exam or an introductory college calculus course (Amanda did both). The eighth participant, Andrew, was repeating introductory calculus despite earning an A; he had failed the second semester of calculus three times at previous institutions, and decided to repeat introductory calculus before attempting Calc II for the fourth time. All participants were required to take calculus I for the major declared in the table, and everyone except Sharon and Elizabeth was required to take at least two additional semesters of calculus, linear algebra, and differential equations. The remainder of Sharon and Elizabeth's mathematics courses were specialized content courses for pre-service elementary school teachers. A table of the participants in the study with a brief description appears below (Table 1).

The three sections from which the participants were drawn were four-credit courses that met 4 days a week (MTWF) for 50-min classes; each class had approximately 35 students at the beginning of the semester. Three of the class days (MTW) were lecture. Friday was reserved for tests and group work. Before each class period, students completed a formative assessment: students wrote down major definitions and theorems, tried a simple problem, and asked questions about what they did not understand. A typical formative assessment appears in Fig. 1. The three lectures were based on these formative assessments; instructors noted the two to three most common questions and discussed them in class (students with other questions were encouraged to come to office hours). On Fridays without tests scheduled, students worked in small groups on problems designed to give them extra practice and conceptual insight on the material covered earlier in the week. Attendance was not a part of the course grade, but students were expected to attend class every day. Students' grades were determined by a weighted average of their three tests/final exam (60%), homework (25%), labs (12%), and formative assessments (lecture preparation activities; 3%). The course covered the first five chapters of the sixth edition of *Calculus* (Stewart, 2006).

I interviewed eight students in the fifth week of the 15-week semester; all participants answered the same initial questions in the semi-structured interviews, but probing questions

Table 1 Participants

Pseudonym and description	Major at time of retake	Reason for retaking	Long-term outcome
Kendra, 18, female	Elementary Mathematics Education	Took calculus at community college that failed to transfer	Completed Elementary Math Education BS
David, 20, male	Meteorology	Failed calculus first semester of college	Left university one semester after repeating calculus
Andrew, 28, male	Secondary Mathematics Education	Non-traditional student, calculus credit expired; failed Calculus II three times at three institutions	Completed Secondary Mathematics Education BS
Jim, 21, male	Secondary Mathematics Education	Failed AP Exam, then failed business calculus in college	Completed Applied Mathematics BS
Elizabeth, 19, female	Elementary Mathematics Education	Failed calculus first semester of college	Completed Elementary Math Education BS
Amanda, 18, female	Secondary Mathematics Education	Failed calculus first semester of college and failed AP Calculus exam; C+ in high school Calculus	Completed Mathematics minor (BS in Physics Education)
Sharon, 18, female	Elementary Mathematics Education	Failed AP Exam, traumatic brain injury	Completed Elementary Math Education BS
Corwin, 19, male	Secondary Mathematics Education	Failed calculus first semester of college	Left university at end of the semester

Section 3.4: The Chain Rule

Directions: Answer all of the following questions to the best of your ability. **All questions must be attempted to receive full credit.**

1. When should the chain rule be used?
2. What are the two forms of the chain rule?
3. For the following functions, state whether the product rule, chain rule, both or neither rule is appropriate to take the derivative of the function.
    a. $f(x) = 2\sin(x)$
    b. $g(x) = 2\sin(2x)$
    c. $h(x) = 2xsin(x)$
4. What questions do you have about this section?

Fig. 1 Typical pre-lecture formative assessment

differed slightly, depending on their responses (Merriam, 1998). I wrote observation notes during the interviews and recorded the interviews for later transcription. I journaled after each interview, recording impressions to maintain an audit trail and increase the trustworthiness of my findings (Merriam, 1998). I observed the students throughout the semester in their classroom setting to triangulate their interview responses; a graduate student twice observed the participants in the course I taught. The three participants whose grades and frequency of participation increased the most throughout the semester were re-interviewed at the end of the semester. Ungraded written journal assignments, asking students to reflect on the class, their feelings about it, and what could be improved, were assigned to all students at midterm and at the end of the semester. For all interviewed students, the number of formative assessments completed and test grades were obtained; these documents were used to triangulate students' interview statements. I also obtained the final grades for any mathematics courses the interview participants took for the next academic year. In addition, Andrew, Amanda, and Elizabeth consented to be interviewed once a year until they graduated. The other participants allowed me to use their transcripts to track their progress through their academic career.

Once transcription was complete, I coded the data using codes and standards of evidence derived from the literature. Based on participant feedback, I modified my original coding scheme to the one presented in Table 2 and coded the data based upon the new standards.

There were several steps taken to ensure the trustworthiness of the analysis. An expert in qualitative research performed an external consistency check on my coding. I performed member checks for each interview by writing a one-page explanation of the themes in each interview; this document was then emailed to the participant with a request for feedback. A graduate student also performed spot checks on 20% of the codes; the inter-rater agreement was 86% (acceptable for qualitative research). In follow-up interviews, interview participants read over and critiqued the analysis of their previous interviews, which provided triangulation.

Table 2 Coding scheme

Code	Definition	Standards of evidence
Affective engagement: processing failure	A student's emotional ties to either a major or an institution that led them to choose to repeat calculus	Direct statements made within the initial interview
Affective engagement: "better" instructor	Emotional ties to instructor of the repeated course	Direct statement made in interview by a student indicating their belief their instructor cared about their success, accompanied by supporting evidence for the claim
Behavioral engagement: participation in formative assessment	The value students assigned to the formative assessments used in the class	Direct statement made in interviews or during classroom observation
Behavioral engagement: calibration	How well an individual's perception of their performance aligned with their actual performance	High calibration: Students are able to identify within 5% of a grade where they stand in the course Moderate calibration: Students are able to identify within 10% where they stand in the course
Magical thinking	Thinking characterized by a belief that stating/wishing for something will cause it to occur	Direct statements from students which indicated predictions unlikely to occur under present circumstances
Cognitive engagement: problem recognition	Student's ability to recognize and correctly apply the appropriate calculus technique to a particular problem	Students' abilities to distinguish problems similar to those in Fig. 1

3 Findings

There were two activities that all participants said were crucial to successfully repeating their introductory calculus course: understanding their previous failure and having a "better" instructor in the second course. In both cases, the identified events indicated an increase in affective engagement through connections to a particular department/major, and in the perception of a greater emotional connection with their new instructor. Students who successfully repeated calculus reported that participation in the formative assessments was a leading activity (Leontiev 1981) since it brought about cognitive and behavioral engagement changes leading to success. By participating in the formative assessments (which they initially did only because of their affective engagement with their instructor), students practiced their calibration and self-monitoring skills and communicated their misconceptions to their instructors before the unit exams so that they could be addressed. These students also had a lower observed rate of magical thinking; all of these changes indicate increased behavioral engagement.

Each interview began by asking students to discuss their previous course and why they were repeating calculus. Participants blamed bad teaching (4), outside forces beyond their control (2), or themselves (2). Corwin explained, "It was my fault that I failed in the fall. I didn't study hard enough, and I should have done some of the optional homework." Understanding their failure was very important to all participants, but it was not a leading activity; participants' motivation for repeating the course—to pass the course and move on to the next required course in their major—remained unchanged after justifying why repeating the course was necessary. Furthermore, the reasons students gave had no relationship to their ultimate success in the class: some students who blamed their teachers passed, and some students who blamed themselves failed the course again.

However, when asked why they chose to repeat, both successful and unsuccessful repeaters discussed affective engagement with their major. “All I’ve ever wanted to do was teach sixth grade math,” Kendra explained, “If that means I have to take calculus again, so be it.” David, an unsuccessful repeater, had a similar response, “I’ve always wanted to be in weather. Not on TV or anything, but working for the National Weather Service. I have to take a lot of math to do that, and that starts with calc I.” Six of the other participants cited their major leading to a career path, but Amanda also cited family pressure in addition to a desired career path as a reason not to change her major, “My father doesn’t even know I failed in the fall,” she confessed, “My mother is covering for me. My parents are immigrants, and they will only pay for school if I am in a STEM major. I also want to coach high school, and my best way to get there is as a math teacher.” In all cases, we see at least some emotional connection between participants and their envisioned future selves. This increased level of affective engagement may explain why they chose to repeat the course rather than change majors.

In the initial interview, participants did not show any particular cognitive or behavioral engagement increases; however, all nine participants reported increased levels of affective engagement in the form of a perceived increase in an emotional connection with their current instruction. This increase was characterized as having a better instructor. Even students that saw their failure as their own fault did not discuss changes they needed to make in cognitive or behavioral realms. Participants gave four qualities of a good teacher, but generally defined a good teacher as one who does not do the things that a bad teacher does, as Kendra explained in her first interview:

> I failed calculus for a couple reasons, but mostly because my teacher was bad…Well, he didn’t care about us. He did the same thing in class, and never changed, even if we got it or if we were totally lost. And he seemed more concerned about finishing all the material, especially after Thanksgiving, than whether or not we learned it. If we asked a question he thought was dumb, he would sometimes make this face when he explained it, or if it was a really dumb question, he'd tell us to go to the math lab. I was afraid to ask questions… I think he only cared about the smart students, the ones that were getting it. Not students like me. (Kendra, first interview)

All participants agreed with Kendra that good teachers cared about students’ learning, made it safe to ask questions, and avoided rushing through the material; the characteristics of a good teacher common to all participants provided emotional support to participants and increased affective engagement. Six of the participants also reported that good teachers returned graded homework before the test over that homework, an action which supported participants’ behavioral engagement. Participants saw their current instructors as good teachers because they returned work promptly and exhibited none of the indicators of poor teaching they had experienced in their unsuccessful attempt of the course. Perceiving their current instructor as better than their previous instructor led to increased behavioral engagement for students who repeated calculus. Further, their perception shifted students’ motivation from avoiding failure and staying on a desired career path to that of pleasing a caring authority figure:

> I think that the biggest thing is that I work harder because I don’t wanna get written off again. I mean, last time, after I failed the first test and asked for help and [my previous

> professor] told me to get out of the class. I had no motivation to go to a class with a professor who paid no attention to me and had no confidence in me. This semester, from day one I was expected to come to class and actually learn calculus. [Knowing that] helped me a lot. (Amanda, second interview)

Although it is likely that there was a gendered component to participants' perceptions of instructor caring, since both instructors teaching the course in the spring of the initial interviews were female (Bellas, 1999), there was one concrete action the instructors in the repeated course did that the instructors in the initial course did not: the spring classes used daily formative assessments. Students who participated in these formative assessment activities took them as evidence that their new instructors were good teachers who cared about student success; this improved their affective engagement, and helped them connect to the material through their feelings about their teacher. In other words, the formative assessments were pivotal events for students who completed them:

> I've been to three other schools before this one, trying to get past calculus. One thing about this course that's different (well, besides actually passing this time) is that there are these assignments we do before something is talked about in class… The reading sheets are extra grading, which is something none of my other teachers did. So even [though] they are extra work for [my instructor], I think they [the reading sheets] can be really helpful for some students… taking on that extra grading makes [my instructor] different from all the ones I've had before. They wouldn't do that [extra grading]. (Andrew, first interview)

For five students, the initial motivation to complete the formative assessments was a desire to avoid disappointing their current instructors. However, as Jim points out, the more he did the sheets, the more he saw his understanding improve, and he felt that he was developing better study habits and improved behavioral engagement in the course:

> I have to be a good student this time, so I have to make sure that I don't miss any assignments. I usually also actually do a little calc every day. Since I have to open the book anyway to do the formative assessment for the next class, I kind of do a couple problems from the day's homework while I'm at it. I still do a lot on Sunday [the night before it is due], but it seems easier to remember this time. I think I am a better student because of them; I know more about what I don't know before the test. (Jim, first interview)

Participation in the formative assessments helped shift students' engagement for completing the formative assessments from affective (pleasing the instructor) to cognitive and behavioral (taking ownership of learning and completing steps toward learning the material) (Black et al., 2010). In other words, the more formative assessments the students completed, the more they began to see that mastering the material was up to them and their efforts, not dependent only on their instructor. The formative assessments actually helped students develop a sense of self-efficacy around learning the material and concepts. Elizabeth, who asked the most explicit questions on her assignments, articulated the shift from affective to behavioral engagement caused by the formative assessments:

> You know what actually helped? The formative assessments. I didn't think they would. At first, I thought they were just busywork. But then, after, like, a few weeks, I realized that anything I asked about a question about on a formative assessment got answered in

> class or on my sheet when I got it back. Then I knew that [my instructor] cared about my learning. I'm not even sure [previous instructor] knew my name. The other thing that helped was that when there was something on the formative assessment I didn't get, I knew I needed to pay more attention in class that day. (Elizabeth, first interview)

It is clear in Elizabeth's response that her feeling that the instructor cared about her success mattered. It is also clear that she began to develop, through the formative assessments, a sense of what she herself could do to improve her performance; the sheets helped her see where she was weak, and she addressed these weaknesses by paying attention and by asking explicit questions in class.

Andrew and Jim both mentioned that the sheets improved their behavioral engagement in an unexpected way; since they could ask questions in writing, they could ask for help without the embarrassment of admitting they needed help in front of their friends (Jim) or the much younger students in the class (Andrew):

> I know what I don't know, but [my two fraternity brothers] are in this class. They know I had this before, and I am not going to admit what I don't get in front of them. But because of the formative assessment it's like I have an out, you know? I can ask for help on the stuff I need help on without anyone noticing that I asked for help. (Jim, first interview)

The increased affective engagement and improved sense of self-efficacy were not the only benefits of the formative assessments; students who regularly participated in formative assessment showed improved calibration and increased behavioral engagement. First, students began tracking their grades—a proxy for performance and understanding and a type of behavioral engagement. Students who did the formative assessments believed that not missing assignments was important to their success; five of the six students regularly participating in formative assessment began to track which assignments they had turned in and what grade they had received. Andrew, who earned the lowest grade, an A−, was the only participant who did not actively track his grade throughout the semester either by logging on to the online course management software or on a paper sheet the students created based on the syllabus. Even without formal tracking, Andrew still was able to accurately describe his performance and showed self-awareness about his understanding of the material:

> I expect to get an A… unless I get lazy, which, um, is totally possible. What I guess is different this time is that I deserve my grade. Since I got that note card [in the prior calculus courses to use on exams] I never had to learn anything. Sure, I got a grade, but that grade didn't mean [anything], since I've never been able to pass Calculus II. Since I get it this time, I think I can make it through next time. (Andrew, second interview)

The formative assessments, particularly those on Wednesdays, also helped students to regularly *self*-evaluate their own understandings, and this improved both self-efficacy and calibration. Every Wednesday, students were asked one additional final question as part of the prelecture formative assessment: "Thinking back on the material we have discussed so far this unit, what do you feel that you still need to practice more?" The responses to those questions were used to create the lab assignments that students would complete in small groups of three to five students during Friday's class. For the students regularly completing the formative assessments, answering this question every week encouraged regular self-evaluations of their understanding of the recent course content as well as targeted re-teaching of the material they found the most difficult.

Completing the formative assessments not only improved students' behavioral engagement; it also seemed to reduce students' magical thinking that can lead to reduced cognitive and behavioral engagement. No participant who completed the formative assessments made a statement coded as magical thinking in their interviews. David and Corwin, two students who completed less than a third of the formative pre-lecture formative assessments and did not see them as valuable, were the only students who gave responses coded as magical thinking.[1] These two students, who did not develop their self-efficacy or calibration by completing the assignments, both believed that they were earning at least a B in the course despite failing every test. Their calibration was poor mostly due to magical thinking: they believed that if their grade over the material in the review chapter improved by X%, their final grade would also automatically increase by X%. Corwin explained this thinking about grades in his first interview, which took place shortly after the first test was returned:

> My first test showed me that I am going to pass. Even though it was a high D, it was still 20% better than I did the first time in the first semester. That means I should get 20% better this time, so a C. (Corwin, first interview)

David also showed signs of poor calibration when he thanked me after learning that he failed the class a second time:

> Why did I thank you? Well, the first time I took this class, I got a 44%. This time I got a 67%. This means that I will get a B, B+ the next time I take Calc I, so I won't have to change my [meteorology] major. I can definitely handle the math. (David, final exam period)

Poor calibration also tended to reduce motivation and hence cognitive engagement; since David and Corwin believed, incorrectly, that their current level of performance was satisfactory, they thought there was no need to seek help from their instructors or peers. This is likely a manifestation of the Dunning-Kruger effect; without the feedback of the formative assessments, Corwin and David never received evidence challenging their belief that they were performing passably. For repeating students, overconfidence may be a coping mechanism for the academic trauma of failure (Burks, Carpenter, Götte, & Rustichini, 2001; Hoffman, 2010); students repeating a course that now seems familiar to them may be overconfident about success.

The benefits of participation in formative assessment for mathematics affect, self-efficacy, and calibration became evident during the derivative unit of the course. While the participants in the study had sufficient algebra skills to use the basic derivative formulas, they struggled to recognize when it was appropriate to apply particular formulas (Fig. 1). Only 38% of the students who completed the formative assessment in Fig. 1 correctly identified the appropriate derivate rules needed for each function. Because their responses on the sheet revealed this problem, the instructor led discussions in class that revealed the general student confusion

[1] Corwin said in his first interview, "The assignments are not worth very many points, so there really are not any consequences for skipping them." David knew that the assignments were graded, but forgot to do them, and agreed with Corwin that the small weight of the assignments meant that they were not very important. Although this shows a type of strategic thinking and behavioral engagement, regularly missing even low stakes assignments is unlikely to lead to long term success in any mathematics course.

about function composition and recognizing when to use the chain rule, which then became the topics of the Friday lab session.

Interestingly, although all students received the same instruction and participated in the same lab, only the students participating in the formative assessments seemed to benefit from the specific response to their confusion about this important calculus concept. These students saw the discussion as particular evidence of instructor caring; the extra emphasis on this specific misconception helped four out of five successful repeaters realize they needed to take more ownership of their learning (increased behavioral engagement) in order to be successful, and it helped their calibration by showing them what they were getting wrong. In other words, this specific discussion was pivotal for students completing the formative assessment. Andrew explained it this way:

> About halfway through the semester I started going to the math lab. When we did the chain rule, I realized that I didn't really get what a function was, which is why I thought sin(2x) was a product rule. [My instructor] can't always take the time to fix my misunderstandings in class, so if I really want to be a math teacher, I need to patch the holes in my knowledge myself if I am ever going to go further. (Andrew, second interview)

Participation in formative assessment made a lasting impression on two of the three students who participated in the long-term follow-up interviews. In the final interview, conducted 1 month before graduation, Amanda said that the most important lessons she learned from repeating the course were the importance of personal effort and of knowing when to ask for help, but she did not bring up the formative assessments nor remember them when prompted. Elizabeth and Andrew both brought up the formative assessments without prompting. Elizabeth felt that completing the formative assessments was helpful in her job as a tutor throughout her undergraduate career and in her approach to learning in future classes:

> I learned a lot in that class, especially from those formative assessments. Not [because I learned the math in the section from doing them], but because they made me look at the book. I didn't get everything, but I got some things. Then I'd get more in class. Over time, I began to believe that even though I failed the first time, I could learn calculus. In my other math classes, I used to make myself do a formative assessment before class. I'd have to ask my own questions in class, but since I had them written down, I know what to say if the teacher didn't answer the question in class. I tell the kids I tutor all the time to do that if they are having trouble. (Elizabeth, final undergraduate semester)

Andrew, who had accepted a teaching job at a local high school, planned to incorporate formative assessments in his future classes because of the way they supported his affective, cognitive, and behavioral engagement:

> In that first semester here, I did finally learn Calc I well enough to pass Calc II, but what I really learned was how to be a math student. No one is going to open up my brain and magically pour in understanding. I have to work every day on more things than are assigned for points in order to be successful in a math class. That class was the first time that I realized I needed to suck it up and ask for help if I needed it; since I could do that on the formative assessments, it kind of made me realize that nothing terrible would happen if I asked for help. I think formative assessments helped me be a better student. I hope to inflict them upon my students now that I'm done student teaching. (Andrew, final undergraduate semester)

All students that participated regularly in the formative assessments passed the class and successfully completed a science or mathematics undergraduate degree without repeating another mathematics class. Elizabeth, Kendra, and Sharon all graduated with elementary mathematics education degrees with 4.0 GPAs in their mathematics classes; Andrew completed a secondary mathematics education degree with a 3.7 GPA in his mathematics courses. Amanda went on to teach high school physics, and Jim became an actuary instead of a teacher, but both completed their degrees. Corwin and David failed out of the university.

4 Discussion

There were three important events identified by participants in the experience of failing and repeating a course successfully; only two of these events were leading activities. The first was processing failure and choosing to repeat the course. This was not a leading activity, because student motivations remained unchanged. The second was getting a better instructor, which was a leading activity in the sense that students' motivations shifted from passing the course to being seen as competent by a caring authority figure; this increased affective engagement affected student motivation and mathematics affect and gave them the initial nudge needed to change their learning behaviors. For the successful repeaters, participating in the formative assessments was also a leading activity that improved students' self-efficacy and allowed students to practice self-monitoring (calibration) skills; participation in this activity appeared to lessen incidents of magical thinking, and shifted the motivation for learning from extrinsic (being seen as competent, passing the course) to intrinsic (meeting future career goals, understanding the material).

Perceiving their instructor as better than their previous instructor was important during the first half of the course, and this perception is important for student success (Di Martino, Coppola, Mollo, Pacelli, & Sabena, 2013; Wheeler & Montgomery, 2009), in part because it activates positive affective factors that impact student mathematics affect and student learning. Beyond that, students needed some sort of scaffolding to help them confront their deficiencies in study skills and their misconceptions, and to give them the confidence to engage in classroom discourse (Andersson et al., 2015). The formative assessments provided an early warning system for both students and instructors. Students could signal their instructors that they needed help; the act of identifying problems and asking for help encouraged students' ownership of their learning and helped them develop self-efficacy, and the act of reflecting on their understanding helped students improve their calibration. Successful repeaters acknowledged that becoming better at these skills made a difference in their success in the course.

Justifying failure and seeing your instructor as better both seemed important to students' decision to repeat the course, but did not substantially change these participants' actions in the repeated course. Furthermore, the actual justification used for failure did not appear to matter; some students who blamed poor instruction successfully completed the course, and some students who blamed their own lack of effort failed the course again. However, participation in the formative assessments scaffolded missing skills to repeaters, and identified misunderstandings early enough for the instructor to respond; students also responded to these identified misunderstandings by changing their own study and learning behaviors. The benefits of formative assessment are both easily achieved and significant. Such assignments are worth implementing in entry-level classes, particularly at institutions like the one described in this study, which had a large population of first-generation and low socioeconomic status students who needed additional support in these areas.

Leading activities are those activities that produce psychological development; learners encounter many activities, but only some are significant enough to cause a shift in behaviors or beliefs, and these significant events are called leading activities (Leontiev, 1981). University education is a transition between childhood and adulthood, and a finer-grained analysis of students' openness to learning during this transitional period is warranted (Black et al., 2010). Leading activities can be brief, discrete events that lead to a change in students' motivations. For instance, students with high levels of mathematics anxiety often report a specific event where they felt ashamed or upset, often an academic failure, that led to the development of their math anxiety (Hembree, 1990). After the inciting event, students develop ego-protecting behaviors (such as self-sabotage) to avoid a repetition of this incident. The formative assessments, which served as a leading event (Leontiev, 1981) for those students that completed them, appear to provide students with critical emotional support leading to increased affective engagement during their first few weeks of the repeated course. Such emotional support is vital for student success in the course (Tainio & Laine, 2015), for changing students' identities as learners (Braathe & Solomon, 2015; Tsay, Judd, Hauk, & Davis, 2011), and for helping students change beliefs about learning that may have interfered with prior success (Pekrun, 2006). Although this study cannot make any causal conclusions, there was a significant correlation between the number of formative assessments completed and the final course grade using the Spearman's rank correlation ($p < 0.01$), which suggests that further inquiry into this relationship is warranted.

There are two potential limitations in this study. While I was clear with participants that their responses would not have any effect on their course grade, only one participant spent any significant time criticizing their current class, so my dual role as a calculus instructor and a researcher may have degraded the quality of the data. There are two mitigating factors to this dual role. First, five of the eight participants were not enrolled in a class I taught. Second, in longitudinal interviews, three participants (only one of whom was enrolled in any course I taught subsequently) were as candid in follow-up interviews as in the original interviews. Also, repeating students could not repeat the course with the instructor from the failed class; although all students explicitly said that they wanted a different instructor, there is no way to tell how much effect was from instructor difference.

The results of the study suggest two potential areas of further research. Although the students in this study passed the course at the same rate as their peers taking calculus for the first time, further qualitative research should be conducted on whether and how regular participation in formative assessments engenders positive affective changes in students taking calculus for the first time. There should also be research into which types of formative assessment optimally increase students' positive mathematics affect, self-efficacy, and calibration; while the formative assessments in this study were effective for the repeating students, other types of formative assessments might be even more effective for undergraduate mathematics students.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Andersson, A., Valero, P., & Meaney, T. (2015). “I am [not always] a maths hater”: Shifting students’ identity narratives in context. *Educational Studies in Mathematics, 90*(2), 143–161.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.

Bandura, A. (1993). Perceived self-efficacy development and functioning. *Educational Psychologist, 28*, 117–148.

Bellas, M. L. (1999). Emotional labor in academia: The case of professors. *The Annals of the American Academy of Political and Social Science, 561*(1), 96–110.

Black, L., Williams, J., Hernandez-Martinez, P., Davis, P., Pampaka, M., & Wake, G. (2010). Developing a ‘leading identity’: The relationship between students’ mathematical identities and their career and higher education aspirations. *Educational Studies in Mathematics, 73*(1), 55–72.

Black, P., & McCormick, R. (2010). Reflections and new directions. *Assessment & Evaluation in Higher Education, 35*(5), 493–499.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 139–148.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*(1), 5–31.

Braathe, H. J., & Solomon, Y. (2015). Choosing mathematics: The narrative of the self as a site of agency. *Educational Studies in Mathematics, 89*(2), 151–166.

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*, 1–15.

Brown, M., Brown, P., & Bibby, T. (2008). “I would rather die”: Reasons given by 16-year-olds for not continuing their study of mathematics. *Research in Mathematics Education, 10*(1), 3–18.

Burks, S. V., Carpenter, J., Götte, L., & Rustichini, A. (2001). *Overconfidence is a Social Signalling Bias. Discussion Paper # 4840* (pp. 1–43). Bonn: Institute for the Study of Labor.

Chen, P., & Zimmerman, B. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs in middle school mathematics students. *The Journal of Experimental Education, 75*(3), 221–244.

Chiesa, M., & Robertson, A. (2000). Precision teaching and fluency training: Making maths easier for pupils and teachers. *Educational Psychology in Practice, 16*(3), 297–310.

Chiu, M., & Klassen, R. (2010). Relations of mathematics self-concept and its calibration with mathematics achievement: Cultural differences among fifteen-year-olds in 34 countries. *Learning and Instruction, 20*(1), 2–17.

Clark, I. (2011). Formative assessment: Policy, perspectives and practice. *Florida Journal of Education Administration and Policy, 4*(2), 158–180.

Di Martino, P., Coppola, C., Mollo, M., Pacelli, T., & Sabena, C. (2013). Preservice primary teachers’ emotions: The math-redemption phenomenon. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 225–232). Kiel, Germany: PME.

Elliot, D. L., & Gillen, A. (2013). Images and stories: Through the eyes of at-risk college learners. *International Journal of Qualitative Studies in Education, 26*(7), 912–931.

Ellis, J., Kelton, M. L., & Rasmussen, C. (2014). Student perceptions of pedagogy and associated persistence in calculus. *ZDM*, 1–13.

Esmonde, I. (2009). Ideas and identities: Supporting equity in cooperative mathematics learning. *Review of Educational Research, 79*(2), 1008–1043.

Fernández-Plaza, J. A., & Simpson, A. (2016). Three concepts or one? Students’ understanding of basic limit concepts. *Educational Studies in Mathematics, 93*(3), 315–332.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. http://rer.sagepub.com/content/74/1/59. Accessed 15 Nov 2017.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*(5), 478–493.

Gallagher, E., Bones, R., & Lambe, J. (2006). Precision teaching and education: Is fluency the missing link between success and failure? *Irish Educational Studies, 25*(1), 93–113.

Goldin, G. A., Epstein, Y. M., Schorr, R. Y., & Warner, L. B. (2011). Beliefs and engagement structures: Behind the affective dimension of mathematical learning. *ZDM, 43*(4), 547–560.

Gueudet, G. (2008). Investigating the secondary-tertiary transition. *Educational Studies in Mathematics, 67*, 237–254.

Halverson, L. R., Woodfield-Porter, W., Graham, C. R., Hernrie, C., & Borup, J. (2013). Measuring cognitive and affective engagement in blended and online settings. In *Paper presented at the 10th Annual Sloan Consortium, Milwaukee, WI.*
Hannula, M. S. (2006). Motivation in mathematics: Goals reflected in emotions. *Educational Studies in Mathematics, 63*(2), 165–178.
Hardy, N. (2009). Students' perceptions of institutional practices: The case of limits of functions in college level calculus courses. *Educational Studies in Mathematics, 72*(3), 341–358.
Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*(1), 33–46.
Hoffman, B. (2010). I think I can but I'm afraid to try: The role of self-efficacy beliefs and mathematics anxiety in problem solving efficiency. *Learning and Individual Differences, 20*(3), 276–283.
Imada, T., & Ellsworth, P. C. (2011). Proud Americans and lucky Japanese: Cultural differences in appraisal and corresponding emotion. *Emotion, 11*(2), 329–345.
Ingram, J. (2012). *Whole class interaction in the mathematics classroom: A conversation analytic approach* (Unpublished doctoral dissertation, University of Warwick).
Kajander, A., & Louric, M. (2005). Transition from secondary to tertiary mathematics: McMaster University experience. *International Journal of Mathematical Education in Science and Technology, 36*(2–3), 149–160.
Karpinski, A. C., & D'Agostino, J. V. (2012). The role of formative assessment in student achievement. In *International Guide to Student Achievement*, (p 202).
Klassen, R. (2007). Using predictions to learn about the self-efficacy of early adolescents with and without learning disabilities. *Contemporary Educational Psychology, 32*(2), 173–187.
Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134.
Leontiev, A. N. (1981). Problems of psychic development. Moscow: Moscow St. Un. Publ.
Mahatmya, D., Lohman, B. J., Matjasko, J. L., & Farb, A. F. (2012). Engagement across developmental periods. In *Handbook of research on student engagement* (pp. 45–63). Springer US.
Malmivuori, M. (2006). Affect and self-regulation. *Educational Studies in Mathematics, 63*, 149–164.
Martínez-Sierra, G., & García González, M. D. S. (2016). Undergraduate mathematics students' emotional experiences in linear algebra courses. *Educational Studies in Mathematics, 91*(1), 87–106.
Martínez-Sierra, G., & García González, M. D. S. (2014). High school students' emotional experiences in mathematics classes. *Research in Mathematics Education, 16*(3), 234–250.
Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
Quintos, B., & Civil, M. (2008). Parental engagement in a classroom Community of Practice: Boundary practices as part of a culturally responsive pedagogy. *Adults Learning Mathematics, 3*(n2a), 59–71.
Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341.
Roh, K. H. (2010). An empirical study of students' understanding of a logical structure in the definition of limit via the ε-strip activity. *Educational Studies in Mathematics, 73*(3), 263–279.
Selden, A. (2005). New developments and trends in tertiary mathematics education: Or, more of the same? *International Journal of Mathematical Education in Science and Technology, 36*(2–3), 131–147.
Selden, A., & Selden, J. (2002). Tertiary mathematics education research and its future. In D. Holton (Ed.), *The teaching and learning of mathematics at university level* (pp. 237–254). The Netherlands: Springer.
Sonnert, G., Sadler, P. M., Sadler, S. M., & Bressoud, D. M. (2015). The impact of instructor pedagogy on college calculus students' attitude toward mathematics. *International Journal of Mathematical Education in Science and Technology, 46*(3), 370–387.
Stewart, J. (2006). *Calculus early transcendentals, 6e*. Belmont: Thompson Brooks/Cole.
Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice, 44*, 11–18.
Tainio, L., & Laine, A. (2015). Emotion work and affective stance in the mathematics classroom: The case of IRE sequences in Finnish classroom interaction. *Educational Studies in Mathematics, 89*(1), 67–87.
Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal, 20*, 5–24.
Tsay, J. J., Judd, A. B., Hauk, S., & Davis, M. K. (2011). Case study of a college mathematics instructor: Patterns of classroom discourse. *Educational Studies in Mathematics, 78*(2), 205–229.
Wheeler, D. L., & Montgomery, D. (2009). Community college students' views on learning mathematics in terms of their epistemological beliefs: A Q method study. *Educational Studies in Mathematics, 72*(3), 289–306.
Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education, 96*(5), 878–903.
Wyatt, M. (2011). Overcoming low self-efficacy beliefs in teaching English to young learners. *International Journal of Qualitative Studies in Education, 26*, 238–255. https://doi.org/10.1080/09518398.2011.605082