Paper to appear in the *Proceedings of the Eleventh Conference on Research in Undergraduate Mathematics Education* (2008).

Explaining Student Success in One PDP Calculus Section: A Progress Report

Aditya P. Adiredja , Randi A. Engle, Danielle Champney, Amy Huang, Mark Howison,

Niral Shah, and Pegah Ghaneian, UC Berkeley

In 1974 at UC Berkeley, many African American and Hispanic students were failing Calculus while most Chinese American students were successful (Treisman, 1985). Recognizing calculus as a gatekeeping course, Uri Treisman investigated the study practices of both groups. He found that while both attended the same lectures and worked equally hard on homework, the African-American and Hispanic students were academically isolated, while the Chinese-American students worked together on additional challenging problems (ibid.). In 1978, with Rose Asera, Robert Fullilove, Leon Henkin, Dick Stanley, and other members of the Professional Development Program (PDP) at UC Berkeley, he then created calculus intensive discussion sections for underrepresented students. The sections emphasized groupwork on challenging problems and a community based on shared interest in succeeding in mathematics (Asera, 2001; Fullilove & Treisman, 1990; Eric Hsu personal communication, February 2008; Treisman, 1992).

Since then, students in PDP's Intensive Discussion Sections (IDS) and similar programs elsewhere have outperformed comparable students in traditional calculus sections, and been more likely to successfully take additional STEM courses (Alexander, Burda & Millar, 1997; Chin et al., 2006; Fullilove & Treisman, 1990; Kosciuk, 1997; Moreno & Muller, 1999; Moreno et al., 1999). As a result of this success, colleges and universities around the country have sought to use the IDS model on their own campuses. Although some adaptations of the model (like those cited above) have been highly successful, others have not (Asera, 2001; Hsu, Murphy &

Treisman, 2007). We suggest that one key reason for the difficulties in scaling up the program is an insufficient understanding of the underlying causal factors that make the model work (Asera, 2001) and simply preserving surface features does not guarantee that they will serve the same functions in new educational contexts (Brown & Campione, 1996). Although we do know something about the underlying design principles that the originators of IDS-style calculus sections had in mind when creating these programs, little empirical work investigates which of these design principles actually contribute to students' success and exactly how this occurs (Herzig & Kung, 2003; Hsu, Murphy & Treisman, 2007; PDP director Steven Chin personal communication, March 2006). In this study, we begin to address that gap by reporting initial findings from an intensive study of one PDP section that was successful with respect to the traditional measures of grades and future STEM course-taking. We began our study by specifying four hypotheses consistent with the design principles espoused by the originators of the program, the existing literature, and those who currently run the program. We then elaborated on the hypotheses using relevant theories from the literature and designed a data collection plan in order to systematically investigate them. This report presents our initial findings from that study. In the next section we explain each hypothesis before presenting the findings about them.

Theoretical Framework and the Four Hypotheses

Hypothesis 1: Through working on additional challenging tasks during section, students develop fluency with mathematical language and both conceptual and procedural understanding of calculus (Asera, 2001; Fullilove & Treisman, 1990). One difference in the study practices of the African American and Hispanic students and the Chinese American students was additional practice with challenging problems (Treisman, 1985). It is not surprising that Treisman and his colleagues made challenging mathematical tasks a focus in these sections (Asera, 2001; Fullilove

& Treisman, 1990). Fullilove and Treisman (1990) distinguish PDP sections from regular ones by emphasizing the use of worksheets containing “carefully constructed, unusually difficult problems” that provide opportunities for students to “identify areas in mathematics knowledge that students must strengthen to survive and excel” (p. 468).

The idea that working on challenging tasks leads students to higher performance in mathematics has been explored in the education literature. Specifically, Silver and Stein (1996) showed that students whose teachers used challenging tasks that consistently encouraged high level thinking and reasoning in their classroom had the highest learning gains with respect to mathematical problem solving, reasoning, and communication. To determine the level of challenge the tasks in our IDS section reached, we used the Mathematics Task Framework (Stein, Grover & Henningsen, 1996, Stein, Smith, Henningsen & Silver, 2000), which classifies tasks into four increasing levels of cognitive demand: memorization, procedures without connections to concepts, procedures with connections to concepts, and doing mathematics.

Hypothesis 2: The sections support students’ self-efficacy by promoting student success on what they perceive as challenging tasks, leading to more effort and persistence with mathematics (e.g., Asera, 1988; Fullilove & Treisman, 1990). Self-efficacy refers to the “belief in one’s capabilities to organize and execute the courses of action required to manage prospective situations,” (Bandura, 1995, p. 2). Fullilove and Treisman (1990) indirectly address self-efficacy in section in their belief that, “as students achieve success in a subject that they typically find daunting, they become committed to maintaining their success” (p. 476). Support for these beliefs can be traced to Bandura’s (1977) findings that “mastery of challenging tasks conveys salient evidence of enhanced competence,” leading to higher self efficacy that then “enhance[s] intensity and persistence of effort” (pp 201, 212). As African American and

Hispanic students are still underrepresented in STEM majors (Adelman, 2006; Hsu, Murphy & Treisman, 2007), a commitment to maintain success after calculus is crucial to increase the representation of minorities in STEM fields (Treisman, 1992).

Hypothesis 3: In small groups students discuss and engage with one another's mathematical ideas to make progress on the tasks while simultaneously developing academic solidarity (Asera, 2001; Fullilove & Treisman, 1990; Chin personal communication, March 2006). The nature of this hypothesis can be best captured by quotes from Asera's (2001, p. 15) monograph "Calculus and Community." As she noted, "The intent of groupwork was academic; it was to foster conversations in which the students had to articulate their own mathematical ideas and listen to the ideas of others." However, program leaders discovered it had other functions as well: "Over the life of the program, it has become evident that this groupwork does more than just help students learn the mathematics; it helps them to learn what it is to be part of an academic community" (ibid., p. 15). Groupwork is thought to help both with students' learning of mathematics and becoming part of a community of mathematics learners.

We used the principles of student *authority* and *accountability to others* from our theoretical framework about fostering student productive disciplinary engagement (Engle & Conant, 2002) as a theoretical lens to explore this hypothesis. We did so as the notions of students articulating their ideas and listening to the ideas of others closely parallel these two principles, respectively. In addition, these principles can be thought of as norms that support productive groupwork as one enters into and becomes more deeply engaged in an academic community (cf. McClain & Cobb, 2001).

The principle of authority revolves around various incarnations of the root of the word: first, students feel "authorized" to share their ideas whether they are right or wrong (e.g., Cobb et al., 1997; Lampert, 1990a, 1990b), then they become publicly recognized as "authors" of those ideas (e.g., Lampert 1990a, 1990b; O'Connor & Michaels, 1996), and finally through repeated success at influencing others they can develop into local "authorities" within their academic community (Brown et al., 1993; Engle, McKinney de Royston & Langer-Osuna, submitted).

The basic idea of the principle of accountability to others is that students "account for" how their own ideas make sense by reference to those of others in their academic community (Michaels, O'Connor & Resnick, 2007). This involves students knowing about and understanding the ideas of their groupmates, which can develop to a point where they feel accountable for coming to a consensus on how problems can be solved and making sure that everyone understands their mathematical basis before moving on (cf. Lampert, 2001).

Hypothesis 4: The sections provide space for students to safely incorporate their personal identity in developing their academic identity (Asera, 2001; Chin, personal communication March 2006). Returning to the idea of a community based on shared interest in mathematics, we looked into development of students' mathematical identity and how that related to students' personal identities. Asera (2001) supports the idea that sections foster individual growth in the personal identities of the students, explaining "PDP was a place on a big campus where, as students frequently said, 'Someone knows who I am' …. PDP was a place where students could find things familiar and comfortable, but was also a place safe enough to explore new and unfamiliar ideas" (p. 16). Additionally, Fullilove and Treisman (1990) claim that the sections promote growth in academic identity stating that they "create academically oriented peer groups whose participants value success and academic achievement" (p. 476).

Data Collection

Our study focused on one section of 12 first-year students taught by an experienced graduate student instructor (GSI) whom PDP's directors believed would best embody their vision of the program that semester. We recorded videotapes and fieldnotes during almost all section meetings, documenting whole class discussions of lecture material and homework and also following two randomly selected groups as they worked on worksheets. We also copied worksheets and student work on the GSI's weekly quizzes and the professor's two midterms. Further, students responded to four questionnaires throughout the semester: before the semester (10 students), before the first midterm (11 students), between the midterms (9 students), and before the final exam (all 12 students). In addition, 7 of the 12 students participated in an interview at the end of the semester in which we asked them additional questions about their background and experiences, responses to the questionnaires, and insights on three video excerpts of their group interaction. Finally, we interviewed the GSI after most section meetings as well as before and after the semester. This initial report focuses on the questionnaire data supplemented by selected analyses of the worksheets, videos, and student interviews.

Analysis and Findings

Degree of Success of This Particular PDP section

Consistent with expectations, this PDP section was successful by standard measures. The students in this section scored above the other PDP section and all other sections associated with the same lecture on each midterm and final. Their scores were so high that the section was not included in determining the curve for each exam. In addition, 75% of its students indicated they were planning to take the subsequent calculus course. In the following paragraphs we discuss our analytical methods and initial findings about each hypothesis.

1) Mathematical Task Hypothesis

Analysis. For this hypothesis, we analyzed two sets of tasks: those on worksheets and those on exams. For the worksheets, we determined how often the students were in fact given opportunities to work with challenging tasks. We then analyzed the exam tasks to explore an alternate explanation, arising from comments by the GSI in interviews, that the students did well in this class as a result of the GSI's ability to predict actual exam questions.

There were 24 worksheets over the semester and each worksheet had anywhere from 1 to 5 sections, so we estimated the degree of challenge of the worksheets with a stratified random sample of 47 tasks. To represent any variation in task choice over time, we split the 24 worksheets into 12 sequential pairs and selected one worksheet from each for the sample. We then randomly chose one task per section on each worksheet.

As mentioned earlier, our primary tool for analyzing the degree of challenge of the worksheet tasks was the Mathematical Task Framework (Stein et al., 1996, 2000). It consists of nineteen statements that classify tasks into four levels of increasing cognitive demand: memorization (M), procedures without connections to concepts (PWOC), procedures with connections to concepts (PWC) and doing mathematics (DM). This coding involved making a lot of inferences and according to its authors "participants do not always agree with each other—or with us—on how tasks should be categorized, but that both agreement and disagreement can be productive" (Stein et al., 2000). Given that, we further elaborated on the framework with a more specific coding guide, and kept track of inter-rater reliability between the two raters as well as disagreements in classification that remained after discussion. The raters agreed for 77% of tasks, disagreed for 13% (though the disagreement was resolved through discussion), and disagreed on the remaining 10%. Percent agreement rose to 80% when PWOC and DM were collapsed into one "higher-level" tasks category.

In addition to analyzing tasks on the worksheets, as part of a different study by two of the authors, we used the MTF to also analyze selected videos of students working with tasks to determine whether changes occurred in their level of cognitive demand (Engle & Adiredja, 2008). The demand of higher-level tasks is often difficult to maintain while students work on problems, especially when instructors reveal the solution when pressured by students (Henningsen & Stein, 1997), so we wanted to see the extent to which this occurred.

Regarding the GSI's ability to accurately predict exam questions, we examined the degree of fit between the GSI's practice exam questions and the professor's prior and actual exam questions. We developed a coding scheme to score degrees of similarity (scale of 0-5) of questions up to their chapter, section, concept, function class, function, and sign/number. We compared all possible pairings of questions from the GSI's practice exam, the professor's prior exams, and the actual exam.

Findings. Contrary to expectations, we found that the majority (68-76%) of the worksheet tasks were lower-level PWOC tasks (see Table 1a). However, this paralleled the kinds of tasks assigned for homework and on the professor's exams. At the same time—and contrary to previous findings using the MTF—some lower-level tasks *increased* in cognitive demand during group work, which we argue occurred through the GSI problematizing them as the students worked (Engle & Adiredja, 2008).

When comparing the different pairs of exams, we found that pairs of questions had mean similarity scores ranging between 3.0 and 3.3, meaning they commonly had very similar wording but used different mathematical functions (see Table 1b). However, there was no association between the GSI and the professor's decisions of which questions from their prior exams to reuse for Midterm 1 ($\chi^2(1)$, $P > 0.99$) or for Midterm 2 ($\chi^2(1)$, $P > 0.26$). So this GSI was not unusually good at predicting the professor's exams. Another GSI armed with a copy of the prior exams would likely create a practice exam that mirrored the professor's decisions as well as this GSI did. What may have mattered, however, was that time was provided in section for students to do each practice exam, while it was an optional activity for students in non-PDP sections.

Table 1a. *Percentage of tasks at each level of cognitive demand*

Level of Cognitive Demand	Number of Problems	Percentage
M	3	6%
PWOC	32	68%
PWC	3	6%
DM	4	8%
PWOC and PWC	2	4%
PWC and DM	1	2%
PWOC and DM	2	4%

Table 1b. *Mean similarity scores for comparisons between exams*

Source exam	Target exam	Mean score (0-5 scale)
GSI's Practice Midterm 1	Prof's Actual Midterm 1	3.22
GSI's Practice Midterm 1	Prof's Prior Midterm 1	3.33
Prof's Actual Midterm 1	Prof's Prior Midterm 1	3.33
GSI's Practice Midterm 2	Prof's Actual Midterm 2	3.27
GSI's Practice Midterm 2	Prof's Prior Midterm 2	3.00
Prof's Actual Midterm 2	Prof's Prior Midterm 2	3.20

2) Self-efficacy Hypothesis

Analysis and Findings. This hypothesis relies heavily on the fact that students will find tasks challenging, providing opportunities to increase self-efficacy and persistence in mathematics. However, we found that the students rated the worksheets as only moderately challenging ($\mu = 3.22$ on a 1 to 5 scale from very easy to very difficult). Consistent with the hypothesis, students reported they were "often" able to successfully solve the problems.

To assess self-efficacy, we used five questions included in a psychometrically validated measure for the construct by Midgley et al., (2000) and found that the students in our section began and ended the semester with strong self-efficacy for calculus (see Table 2). While there were no significant changes in overall self-efficacy over the semester, students' responses on one of the five questions (with regard to doing the most difficult class work) did indicate a significant increase ($p < .05$) toward even higher self-efficacy. Three questions measured the degree to which students thought they could excel in calculus provided they work hard. Responses in this category indicated agreement that did not waver as the semester progressed. Students recognized that hard work was required, but also that success was in their reach.

Table 2. *Mean student responses to the self-efficacy measure*

Question (Increasing Scale: from "not at all true" to (5) "always true")	**Means** First	**Means** Last	**Significance of Change**
"I am certain I can figure out how to do the most difficult class work in calculus."	2.82	3.92	$p < .05$
"I am certain I can master the skills and concepts taught in calculus this semester."	3.82	4.17	$p < .10$
"I can do even the hardest work in calculus class if I try."	3.64	4.00	NS
"Even if the work is hard, I can learn it."	4.00	4.25	NS
"I can do almost all the work in calculus if I don't give up."	3.91	4.17	NS
Overall self-efficacy measure	18.19	20.51	NS

3) Groupwork Hypothesis

Analysis. To understand the effects of groupwork on the students, we included a variety of relevant items in the student surveys. First, to have an open-ended item that would not be influenced by our theoretical perspective, at the beginning of the three surveys during the course we asked students to report what they perceived to be the pros and cons of working with their groups. Later in the surveys, we asked students to report their levels of agreement with various groupwork norms based on the authority and accountability principles of our theory about fostering productive disciplinary engagement.

Findings. When asked about the advantages and disadvantages of groupwork, all 12 students mentioned ways that their peers were particularly valuable to them, with most citing more than one ($\mu = 2.9$). The most frequently cited advantages of working with peers in small groups were the ability to get unstuck on problems (67%: "if I get stuck on a problem, there is a great chance that one of my group members will be able to help me") and development of better understandings of problems (50%: "if you don't understand something, someone's always there to make sure you do"). Other common responses included getting new ideas from peers (42%: "seeing how other people go about solving a problem"), generally getting help or learning from peers (42%: "I can learn from others"), and receiving more understandable explanations from them (33%: "the explanations of topics are much more clear").

With respect to disadvantages of groupwork, half the students responding to each survey either said that there were no disadvantages or left the question blank. Of those who mentioned disadvantages, the only one mentioned by more than one student was negotiating pacing. For example, one student noted, "Not everyone is able to go through the problems at the same speed, which is not only frustrating to those who understand it, but also those who don't understand it very well." Every student that noted any disadvantages of groupwork mentioned this issue.

With respect to norms, there was also evidence that students perceived several aspects of authority and accountability to either be in place in the section or to have developed over time. With respect to authority, most students felt "safe" or "somewhat safe" ($\mu = 3.8$) sharing their solutions, whether correct or incorrect, with a trend in the direction of greater safety at the end of the course ($F(2,16) = 2.6$, $p = .10$). There was also evidence for development of the authorship aspect of authority as students reported that their ideas were recognized and acknowledged by both their classmates and the GSI, with the amount of recognition by classmates increasing from "sometimes" at the beginning ($\mu = 2.9$) to "often" at the end of the semester ($\mu = 3.8$; $F(2,14) = 4.4$, $p < .05$; HSD = .78 at $p < .05$). Finally, at the end of the semester we found that students had generally 'become authorities' in their groups and were "somewhat comfortable" or "comfortable" correcting the ideas of the GSI ($\mu = 3.7$) and their classmates (3.6). They were even more comfortable correcting the ideas of their groupmates, compared to other classmates ($\mu = 4.3$; $F(2,18) = 4.0$, $p < .05$; HSD = .63 at $p < .05$).

With respect to accountability to others, we identified strong levels of perceived accountability to others' ideas along with weaker and mixed levels of accountability for helping other students learn. Throughout the semester, students reported that it was either "important" or "very important" that they know the methods that their groupmates are using to solve problems ($\mu = 4.3$), understand the concepts behind those methods ($\mu = 4.5$), understand how their methods relate to them ($\mu = 4.4$), and eventually agree with their groupmates on at least one solution to each problem ($\mu = 4.5$). With respect to helping other students learn, most students agreed that they made sure that all group members understand before their group moved on to the next problem ($\mu = 3.9$), with a trend towards greater agreement with this norm at the end of the semester ($F(2, 14) = 2.74$, $p = .10$). With respect to whether students felt responsible for group members that were falling behind, however, there was variation among students in whether they agreed or disagreed ($\mu = 3.2$, $SD = 1.0$). This may have represented different degrees of an actual sense of responsibility or alternative interpretations of "responsible," as being either responsible for doing something to address the fact that the groupmate was falling behind or at fault for the situation, an issue we plan to explore in the student interview data.

4.) Identity Hypothesis

Analytical Methods. To explore this hypothesis, we combined survey questions and interviews of students explaining their answers to the questions. We asked two separate questions, one about personal identity and the other about academic identity. To assess the students' personal identities we asked them to rate the degree to which they agreed with the statement, "I can be myself in section." For academic identity we asked the students to rate the degree to which they agreed with the statement, "I consider myself a 'math person.'" Since both questions are vague and thus can be interpreted differently, we complemented their responses to the questions with their elaborations and explanations about them during interviews.

Findings. Related to personal identity, we found that most students agreed throughout the semester that they could "be themselves" in section ($\mu = 3.9$ on a 1 to 5 scale). From the interviews we found that five of the seven students interpreted this as meaning 'being comfortable' in section. As one student put it, the section had "definitely a comfortable feeling compared to other classes." In addition to feeling comfortable, for five out of seven students this also meant that they felt safe to make mistakes in section (e.g., "If I made a mistake it would be okay"). For three out of the seven students it meant that they did not need to play a role in section. For example, one student stated, "I didn't have to change anything about myself to suit them…I could just talk to them how I would talk to my friends."

Related to academic identity, we found that half of the students did not consider themselves "math persons" initially, but by the end of the semester most did ($\mu = 3.6$). Surprisingly, we discovered through interviews that to all seven of the students this meant enjoying mathematics. For three out of the seven it also meant being good at math.

Summary and Future Analyses

Thus far we found that in this one PDP calculus section most of the mathematical tasks that students worked with were not particularly demanding, though some may have grown in their levels of demand. The idea of increasing self-efficacy through success on challenging tasks was not relevant to our students, as they began with high self-efficacy and did not view the tasks as challenging. With respect to groupwork, students adopted several potentially productive groupwork norms and appreciated their peers' contributions. Pertaining to issues of identity, it can be argued that this section did support the simultaneous development of students' personal and academic identities, as students in our section remained comfortable being themselves as their identification with mathematics grew.

The findings for the type of mathematical tasks are inconsistent with prior expectations in that the tasks were not higher-level tasks. However, they were akin to the kinds of problems the professor was including in his exams, which is one aspect of task design mentioned by the originators (Fullilove & Treisman, 1990). Still, many would argue that PDP sections and those of similar programs around the nation are supposed to be based on challenging tasks (Asera, 2001, Hsu et al., 2007) and ours was not. While that is a principle shared by PDP more generally, despite its absence in this section, these students were still wildly successful in this course. However, it is unclear whether their grades truly reflected their understanding of calculus given the limited challenge provided by the exams. To see the longer term effects of their participation in this section, we are planning to distribute another round of surveys to these students and comparison groups to see how they performed in subsequent mathematics classes that depended on understanding of material from this one. In addition, it would be informative to study sections associated with different professors whose exams require students to understand key concepts in calculus instead of only knowing how to perform procedures.

While increasing student self-efficacy was not relevant to our group of students, the fact that a high level of self-efficacy was maintained throughout the semester is notable. While it is not surprising for students to come to UC Berkeley with high self-efficacy, its maintenance is not trivial. Many factors, such as general struggle with calculus, being away from home for the first time, multiple pressures from different courses, or simply being at UC Berkeley with other talented students can often cause decreases in students' self-efficacy. So the fact that our students maintained high self-efficacy suggests that PDP may have helped support student self-efficacy. We plan to use our interview data to identify different kinds of support that PDP provided the students that might be associated with this maintenance.

With respect to groupwork, so far we have focused our analyses mainly on survey responses. However, there is a wealth of untapped information in the videotapes that we plan to investigate. In particular, we would like to understand how the instructor and students developed the productive groupwork norms in the beginning of the semester. Video episodes in concert with relevant GSI interviews will help us explore this issue. Similarly, we will use our interviews and videotapes from early in the semester to better understand how the students figured out that this was one class where they could in fact be themselves while showing a commitment to developing as mathematics learners.

Finally, we recognize that the four hypotheses are not exhaustive and there remain other hypotheses to be explored. For example the idea that students can become a part of the PDP community is worth investigating considering the importance both Asera (2001) and Fullilove and Treisman (1990) placed on it. For now, we conclude with some implications.

Conclusion and Implications

The success of students in this particular PDP section came as a result of several factors that are probably connected (Asera, 2001, Herzig & Kung, 2003; Hsu et al., 2007). In addition, the success of this section compared to the other PDP section in the same semester, which had a GSI new to the program, demonstrates that key surface features such as additional two hours that PDP sections meet, extra attention from the program, and student self-selection into the program may be important, but they are not sufficient to explain the success of the program.

Looking at the broader picture, this study was based on the traditional way of measuring success that focuses simply on grades and continued course-taking. Given our findings we argue that other aspects of success are worthy of further research, including students' growth in conceptual understanding that might be measured by assessments like the Calculus Concept Inventory (Epstein, 2006), their progress in learning to participate in mathematical discourse (e.g., Kieran, Forman & Sfard, 2002), and the nature of their growing identification with mathematics (e.g., Boaler & Greeno, 2000).

References

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, D.C.: U.S. Department of Education.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, D.C.: U.S. Department of Education.

Alexander, B. B., Burda, A.C., and Millar, S.B. (1997). A community approach to learning calculus: Fostering success for underrepresented ethnic minorities in an emerging scholars program. *Journal of Women and Minorities in Science and Engineering, 3,* 145-159.

Asera, R. (2001, May). *Calculus and community: A history of the emerging scholars program*. Princeton, NJ: The College Board,.

Asera, R. (1988). The math workshop: A description. *CBMS Issues in Mathematics Education, 1,* 47-62.

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

Boaler, J., and Greeno, J. G. (2000). Identity, agency, and knowing in mathematics worlds. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning.* Westport, CT: Ablex.

Brown, A. L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J. C. (1993). Distributed expertise in the classroom. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 188–228). New York: Cambridge University Press.

Brown, A. L. & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.) *Innovations in learning: New environments for education* (pp. 289-325). Mahwah NJ: Erlbaum.

Chin, S., Fan, M., Imbens, G. & Perloff, J. M. (2006). *Evaluation of a program to help minorities succeed at college math: UC Berkeley's Professional Development Program.* Berkeley, CA: UC Berkeley's Professional Development Program.

Cobb, P., Gravemeijer, K., Yackel, E., McClain, K., & Whitenack, J. (1997). Mathematizing and symbolizing: The emergence of chains of signification in one first-grade classroom. In D. Kirschner & J. A. Whitson (Eds.), *Situated cognition: Social, semiotic and psychological* *perspectives* (pp. 151–233). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Engle, R. A. & Adiredja, A. P. (2008, April). *Growing the cognitive demand of mathematical tasks.* Paper presented at the National Council of Teachers in Mathematics (NCTM) 2008 Research Presession, Salt Lake City, UT.

Engle, R. A. & Conant, F. C. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, *20*(4), 399-483.

Engle, R. A., McKinney de Royston, M. & Langer-Osuna, J. (submitted). Toward a model of differential influence in discussions: Negotiating quality, authority, and access within a heated classroom argument. *Proceedings of the Annual Meeting of the Cognitive Science Society.*

Epstein, J. (2006, February). *The calculus concept inventory.* Paper presented at the Conference on Research in Undergraduate Mathematics Education, Piscataway, NJ.

Fullilove, R. E., and Treisman, P.U. (Summer 1990). Mathematics achievement among African American undergraduates at the University of California, Berkeley: An evaluation of the Mathematics Workshop Program. *Journal of Negro Education*, Vol 59.

Henningsen, M. & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, Vol 28, No. 5, 524-549.

Herzig, A. & Kung, D. T. (2003). Cooperative learning in calculus reform: What have we learned? In A. Selden, E. Dubinsky, G. Harel & F. Hitt (Eds.), *Issues in mathematics education Vol 12: Research in collegiate mathematics education: Vol. 5* (pp. 30-55).

Providence, RI: American Mathematical Society in cooperation with the Mathematical Association of America.

Hsu, E., Murphy, T. J., & Treisman, U. (2007). Supporting High Minority Achievement in Introductory Collegiate Mathematics Courses: The Emerging Scholars Program Turns 30. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics education*. Washington, DC: Mathematical Association of America.

Kieran, C., Forman, E. & Sfard, A., (Eds.) (2002). *Learning discourse: Discursive approaches to research in mathematics education.* Dordrecht, NL: Kluwer.

Kosciuk, S. (1997). *Impact of the Wisconsin Emerging Scholars first-semester calculus program on grades and retention from fall 1993-1996.* Madison, WI: University of Wisconsin LEAD Center.

Lampert, M. (1990a). Connecting inventions with conventions. In L. P. Steffe & T. Wood (Eds.), *Transforming children's mathematics education* (pp. 253–265). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lampert, M. (1990b). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29–63.

Lampert, M. (2001). Teaching to establish a classroom culture. In Teaching with problems and the problems of teaching (pp. 51-100). New Haven, CT: Yale University Press.

McClain, K., & Cobb, P. (2001).The development of sociomathematical norms in one first-grade classroom. *Journal for Research in Mathematics Education, 32*, 234–266.

Michaels, S., O'Connor, M. C. & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education.*

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales.* Ann Arbor: University of Michigan.

Moreno, S. E., Muller, C. (1999). Success and diversity: The transition through first-year calculus in the university. *American Journal of Education*, vol 108, 30-57.

Moreno, S. E., Muller, C., Asera, R., Wyatt, L., and Epperson, J. (1999). Supporting minority mathematics achievement: The Emerging Scholars Program at the University of Texas at Austin. *Journal of Women and Minorities in Science and Engineering*, 5(1), 53-66.

O'Connor, M. C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning and schooling* (pp. 63–103). New York: Cambridge University Press.

Silver, E. A. & Stein, M. K. (1996). The QUASAR project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education, 30,* 476-522.

Treisman, P. M. (1985). A study of the mathematics performance of black students at the University of California, Berkeley. Dissertation, University of California at Berkeley.

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