Journal for Research in Mathematics Education
2010, Vol. 41, No. 4, 351–382

Enacting Proof-Related Tasks in Middle School Mathematics: Challenges and Opportunities

Kristen N. Bieda
Michigan State University

Discussions about school mathematics often address the importance of reasoning and proving for building students' understanding of mathematics. However, there is little research examining how teachers enact tasks designed to engage students in justifying and proving in the classroom. This article presents results of a study investigating the processes and outcomes of implementing proof-related tasks in the classroom. Data collection consisted of observations of 7 middle school classrooms during implementation of proof-related tasks—tasks providing opportunities for students to produce generalizations, conjectures, or proofs—in the Connected Mathematics Project (CMP) curriculum by teachers experienced in using the materials. The findings suggest that students' experiences with such tasks are insufficient for developing an understanding of what constitutes valid mathematical justification.

Key words: Classroom interaction; Curriculum; Discourse analysis; Middle grades, 5–8; Proof; Reasoning; Teaching practice

Despite the wealth of existing literature on students' abilities to generate and understand proof, we still know very little about how skills related to justifying and proving are taught in school mathematics—particularly in mathematics courses outside of high school geometry. We do know that of the many functions of proof in mathematics, which consist of "verification, explanation, systematization, discovery, communication, construction of empirical theory, exploration of definition and of the consequences of assumptions, and incorporation of a well-known fact into a new framework" (Yackel & Hanna, 2003, p. 228), proof in the high school geometry curriculum has evolved to emphasize verification (Herbst, 2002b). Further, research has not examined students' opportunities to develop deductive reasoning and to learn skills for evaluating the validity of others' mathematical arguments. Without such opportunities, students are inadequately prepared to participate in meaningful discussions about mathematical proofs and to explore the variety of roles a proof can play in doing mathematics.

This research was completed to fulfill the dissertation requirement for a doctoral degree at the University of Wisconsin–Madison, under the advisement of Eric Knuth. The author wishes to thank Amy Ellis, Eric Knuth, Raven McCrory, Mitchell Nathan, Simone Schweber, Michael Steele, and three anonymous reviewers for their feedback on prior drafts of this article.

In the United States, calls for reform declare that reasoning and proof cannot be activities that students do in some mathematics classes but not in others. Rather, reasoning and proving should be an everyday occurrence in K–12 mathematics classrooms (NCTM, 2000, 2009). However, mathematics education researchers in other countries with long-standing commitment to proof throughout the curriculum, such as Italy, France, and Japan, have found that teachers find it challenging to include activities related to reasoning and proving as a regular part of their practice. This is primarily due to difficulties students face when generating justifications and proofs (Mariotti, 2006). To understand how to change the ways in which students learn to prove in school mathematics communities, we need to understand how teachers, students, and the curricula they use—elements existing at the nexus of school mathematics communities (Cohen, Raudenbush, & Ball, 2003)—interact in classroom settings when students are discussing and developing justifications and proofs. The goal of this research is to investigate how tasks designed to engage students in justifying their reasoning are implemented in middle school classrooms.

The corpus of existing literature on students' conceptions of proof indicates that students at all grade levels have great difficulty with proof (Knuth, Choppin, & Bieda, 2009; Chazan, 1993; Hoyles, 1997; Hoyles & Küchemann, 2002). Students, even undergraduate mathematics majors, tend to generate empirically based justifications instead of constructing deductive proofs and to supplement valid deductive proofs with diagrams or examples (Chazan, 1993; Hoyles, 1997; Knuth, Choppin, Slaughter, & Sutherland, 2002; Weber, 2001). The nature of the didactical situation of teaching students to do proofs in school mathematics, however, may unintentionally support a dichotomy between the mathematically rigorous proofs that teachers expect from students and the empirically based arguments that students believe are sufficient for proof (Harel & Sowder, 1998; Hoyles, 1997; Weber, 2001). Students' prior experience with what serves as proof in contexts outside of mathematics, such as the notion of providing proof that a criminal has committed a crime, conflicts with the construct of mathematical proof, particularly its emphasis on deductive reasoning with general arguments, that is unique to the discipline (Fischbein, 1982). Mathematics teachers and the textbooks they use are representatives of the discipline and serve as guides in helping students understand what counts as valid mathematical proof. This role requires teachers to strike a precarious balance between illustrating norms of proving as characteristic of the discipline and leaving room for students to discover and learn the necessity and utility of those norms (Herbst, 2002a).

Whereas teachers may face pragmatic difficulties when engaging students in proving in classroom settings, teachers' knowledge of proof and their beliefs about teaching proof may also constrain their ability to teach proof effectively. Examining high school (Grades 9–12) mathematics teachers' beliefs about pedagogy and proof, Knuth (2002a) found that many of the teachers did not believe that either formal or less formal proofs were suitable for all secondary mathematics students. The teachers struggled to understand, mathematically, the role of proof and the various forms of proof that can be employed to justify an argument (Knuth, 2002b). Given that middle school (Grades 6–8) teachers may have taken fewer proof-related advanced mathematics courses than their high school counterparts, Knuth's findings about the knowledge and beliefs of high school mathematics teachers regarding proof suggest that middle school teachers are not likely to be prepared for implementing activities to develop students' deductive reasoning and support students in communicating their reasoning through proof.

Existing literature also suggests weaknesses in the potential of curriculum to facilitate engaging students in justifying and proving. Although U.S. students' initial encounter with proof traditionally occurs during a high school geometry course, recently developed series of NCTM Standards-based middle-grades curricula (e.g., Mathematics in Context [MiC], Connected Mathematics Project [CMP]), funded by the U.S. National Science Foundation (NSF), were designed to include activities that more explicitly address reasoning and proof. One analysis of the CMP curriculum has shown that although the materials contain a variety of opportunities for students to engage in reasoning and proof, these opportunities are not consistent in sixth through eighth grades (G. J. Stylianides, 2007, 2009). Stylianides found that out of a sample of 4,855 tasks [1] from units focused on algebra, number theory, and geometry, 40% were designed to offer at least one opportunity for students to engage in reasoning and proving activities. The distribution of these tasks across the curriculum, however, was uneven. Stylianides determined that the first unit in sixth grade offered more reasoning-and-proving tasks (tasks that engaged students in pattern recognition, pattern generalization, conjecture development, and/or proof production) than any other unit in the series for sixth through eighth grades. The inconsistency of the CMP curriculum's inclusion of reasoning-and-proving tasks may result in a lack of curricular support for the development of students' knowledge and skills to reason and prove throughout middle school. Further, inconsistent experiences with proving may leave students unprepared to successfully learn more rigorous, formal proof in high school mathematics courses. To achieve NCTM's (2000) recommendations that reasoning and proof be a regular, everyday part of doing mathematics for all students, it is vitally important that the curriculum materials provide consistent, progressively advanced opportunities to justify and prove.

One way to gain a better understanding of the possibilities for improving the teaching of proof is to explore a best-case scenario: classrooms using a curriculum with rich opportunities for students to justify and prove, taught by teachers who are experienced with the curriculum and attending ongoing professional development. In such a case, the question is how proof-related tasks are enacted and how those aspects of teaching proof can apply to other settings. The aim of this study is to examine the process of learning to prove in middle school classrooms by studying the enactment of proof-related tasks in settings such as those described previously as best-case scenarios. This research has implications not only for middle school

[1] *Tasks* here refer to "any exercise, problem, activity, or parts thereof that have a separate marker in the students' textbook" (G. J. Stylianides, 2009, p. 270).

teachers' pedagogies with regard to proof but also for ways in which curricula can support preparation of students for mathematics beyond the middle grades.

CONCEPTUAL FRAMEWORK

Applying a Unifying Definition of Proof in School Mathematics

Scholarly work examining students' and teachers' proving activity has typically utilized proof frameworks that originate with a definition of proof as mathematicians use proof. One broadly disseminated definition, from *Principles and Standards for School Mathematics* (NCTM, 2000), states that proofs are "arguments consisting of logically rigorous deductions of conclusions from hypotheses" (p. 55). This definition is adequate in describing the results of proving in a mathematically appropriate way but fails to describe the core features of a proof and the social ingredients necessary to generate proof.

A. J. Stylianides' (2007) definition of proof aligns the conceptions of proof in the discipline with those in school mathematics, emphasizing sociocultural aspects of students' proving practice:

> Proof is a mathematical argument, a connected sequence of assertions for or against a mathematical claim, with the following characteristics:
>
> 1. It uses statements accepted by the classroom community (*set of accepted statements*) that are true and available without further justification;
>
> 2. It employs forms of reasoning (*modes of argumentation*) that are valid and known to, or within the conceptual reach of, the classroom community; and
>
> 3. It is communicated with forms of expression (*modes of argument representation*) that are appropriate and known to, or within the conceptual reach of, the classroom community. (p. 291)

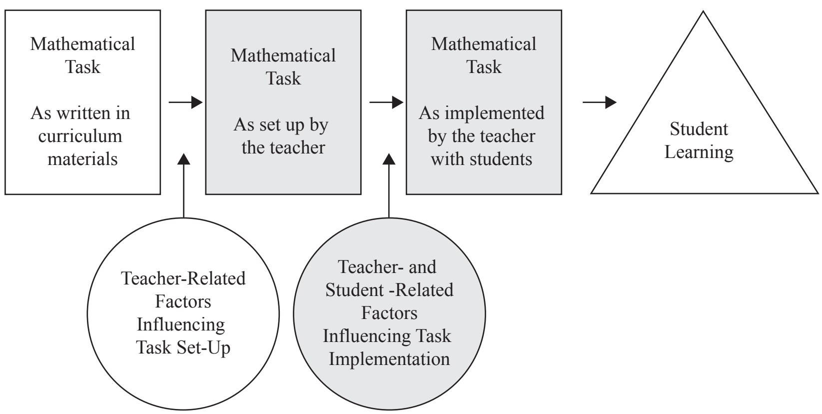
One critical feature of Stylianides' definition is the attention to the role of the classroom community. Although mathematicians rigorously scrutinize their colleagues' proofs, teachers typically serve as arbiters of what counts as valid proof in the classroom. The *didactical contract*[2] between teacher and students in the classroom can constrain students' experiences with authentic proving activity (Herbst, 2006) and may lead unintentionally to the development of authoritarian or ritualistic proof schemes (Harel & Sowder, 1998). However, if proof, as defined by A. J. Stylianides, emerges as a product of discussions between the teacher and

[2]The *didactical contract* is a set of implicit expectations between teachers and learners in instructional settings, established by institutional goals that "link the roles of teacher and students and shape their work in such a way to meet these goals" (Martin, McCrone, Bower, & Dindyal, 2005, p. 99, citing Brousseau, 1984).

students about what is required to justify a conjecture for all cases and how various forms of argumentation influence a justification's proof potential, students learn that proof goes beyond convincing oneself and convincing a friend in that it must convince the broader audience of the classroom community. Further, proof also becomes a vehicle for establishing stores of mathematical knowledge developed by, and accessible to, members of the classroom community. To understand how classroom communities engage in proving activities, attention must be paid to the teacher's and students' moves, or the various kinds of turns in discourse, following the path of a written proof-related task to the resulting justification or proof.

Defining a Process of Enacting Proof-Related Tasks

The design of this research draws from a framework of the mathematical instructional task (Henningsen & Stein, 1997) to understand the implementation of proof-related tasks. The framework (see Figure 1) illustrates that a task as written is not considered the same as the task as implemented; mathematics tasks evolve from their form in curriculum materials to a new form as teacher and students interact with the tasks. Applying this framework to a study of task implementation, Henningsen and Stein found that tasks set up by the teacher to be "doing mathematics" tasks—tasks of the highest cognitive demand that include activities such as developing and proving conjectures—often devolve into tasks of lower cognitive demand. Their research identified several factors related to a reduction in cognitive demand during implementation, including reduction of task demands by the teacher, inappropriateness of task for the students, and a focus on correct answers instead of process-oriented outcomes.



Note. The shaded elements are the focus of this study.

Figure 1. Elements of focus in studying the implementation of proof-related tasks (adapted from Henningsen & Stein, 1997, p. 528).

This study extends existing work of Henningsen and Stein (1997) by distinguishing proof-related tasks from other "doing mathematics" tasks such as nonroutine problem solving, focusing data collection and analysis on the set-up and the implementation of proof-related tasks, and attending to aspects of the teacher's and students' interactions during whole-class discussions about proof-related tasks. The particular features of interest in the framework (shaded in Figure 1) are the teacher- and student-related factors influencing the evolution of a task as intended by the teacher to the task as enacted with students. Such factors include the teacher's and students' discursive moves as they discussed proof-related tasks in the classroom, focusing not only on how the teacher guides students to work on such tasks but also on how students respond and how the teacher attempts to develop and address students' responses. This research shows the nature of teacher and student interactions as a teacher enacts proof-related tasks in a middle school classroom.

RESEARCH QUESTION

Considering the existing work about students' competencies in constructing mathematical proofs, teachers' beliefs and knowledge about proof, and the opportunities as written in curriculum for students to engage in producing justifications and proofs, the guiding question framing this research addresses the kinds of experiences students encounter in the classroom when justifying and proving their mathematical claims, namely:

> What is the nature of the students' and teacher's actions and discourse during the enactment of tasks written to engage middle school students in justifying and proving?

METHOD

The author conducted a multiple case study of classrooms engaged in doing proof-related tasks with experienced teachers familiar with the curricular materials that contained those tasks. Multiple case studies are used when addressing either descriptive or explanatory research questions for which the units of analysis are multiple cases chosen "as replications of each other, contrasting comparisons, or as hypothesized variations" (Yin, 2006, p.112). In this research, the case is the enactment of proof-related tasks in middle school mathematics, for which the units of analysis are implementations by experienced sixth- through eighth-grade teachers trained to implement curricular materials containing proof-related tasks. Given that the research question seeks descriptions of the students' and teacher's involvement in the enactment of proof-related tasks across the middle school level, the multiple case study design supports data collection to address the intent of the research. Classroom observations, teachers' responses to preobservation writing prompts, and teachers' responses during postobservation interviews served as the sources of data for these case studies. The data presented in this manuscript come from the multiple case study, primarily from observations of several lesson implementations over the course of the school year intended to generate rich descriptions of the nature of teachers' and students' engagement with proof-related tasks.

Participants

Data were collected in seven middle schools from the same school district. Among these seven schools, 54% of the entire student population declared race as White, whereas 46% reported NonWhite (either African American, Hispanic, Asian American, or Native American). The percentage of students among these seven schools eligible for the free/reduced lunch program was 42% (range: 17.5%–85%). All seven middle schools from this district had begun to implement the Connected Mathematics Project (CMP) curriculum 6 years prior to data collection. Authors of the CMP curriculum wrote the series to align with *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989). One stated goal of the CMP curriculum is to provide students with opportunities to reason and prove throughout the series, and, as indicated earlier, an independent curriculum analysis reported that 40% of the tasks, as written, engage students in reasoning and proving (G. J. Stylianides, 2007, 2009). Although a new version of CMP (CMP 2) was launched in 2005, only three of the seven schools were in the process of transitioning to CMP 2 during the period of data collection. The participating teachers from these three schools taught the lessons observed from the most recent version.

Recruiting participants began with determining that, for the purpose of sampling across grade levels and different demographic contexts, the sample would consist of two sixth-grade teachers, two seventh-grade teachers, and two eighth-grade teachers from different schools within the district. No two teachers participating in the study would teach in the same building. The district middle school mathematics resource personnel recommended 22 teachers for participation based on the criteria that each of the potential participants had taught from CMP for at least 3 years and had participated in district-mandated professional development for the curriculum. After initially categorizing this list of 22 teachers by grade level and school, the author selected one teacher at random in each of the seven schools to be contacted to participate, with the goal of obtaining two teachers at each grade level for a total of six participants. The seven teachers initially recruited received e-mail messages informing them about the study and what they would be asked to do as a participant. Two of the seventh-grade teachers in the initial pool declined to participate, so two other seventh-grade teachers, selected at random, were recruited. A total of six teachers—two sixth-grade teachers, two seventh-grade teachers, and two eighth-grade teachers—volunteered to participate in the study from the nine teachers recruited. One sixth-grade teacher, who was a part of the initial set of seven teachers recruited to participate in the study, indicated a willingness to participate late in the recruitment process. Although this resulted in three sixth-grade teacher participants, the teacher was included in the sample so that all seven middle schools that had participated in CMP implementation six years earlier would have representation in the participant pool.

Participants had taught, on average, a total of 17 years, with 9 years being the least amount of teaching experience for any teacher in the group. Additionally, six of the seven teachers belonged to the district's CMP Leadership Academy, which required members to attend daylong seminars at least once per quarter to explore problems with implementing CMP and to discuss research-based articles covering current issues in mathematics education. It is important to note that measures of knowledge about, or interest in, reasoning and proof were not used to select participants. Instead, participant recruitment was designed to obtain a sample representative of highly qualified middle school mathematics instruction—experienced teachers who participate in ongoing professional development using the curriculum. Studying such a sample of teachers would develop an understanding of how teachers, without necessarily having obtained professional development on, or having professional interests in, issues of teaching and learning reasoning and proof, would implement tasks with rich potential for engaging students in justifying and proving.

Curriculum Analysis

Preceding collection of data in the classrooms, the author conducted a targeted curriculum analysis of CMP textbooks to determine lessons (called *Investigations* in CMP) that contained proof-related tasks (called *Problems* in CMP). From this point forward, the terms *lessons* and *investigations*, as well as *tasks* and *problems*, will be used interchangeably. In this study, proof-related tasks are defined as those written to engage students in generalizing patterns and developing conjectures[3] from such patterns or problems presenting conjectures to be justified. This curriculum analysis enabled the author to schedule classroom observations that maximized the likelihood that the class periods observed would include activity involving proof-related tasks. As such, the analysis aimed to identify neither every proof-related task as written in CMP (because not all investigations containing proof-related problems were taught during the data collection period) nor every task that could potentially result in *proving activity*.[4]

Unit of Analysis

The analysis included only problems labeled as "Problem" or "Follow-up" (for example, "Problem 1.1" or "1.1 Follow-Up") within unit investigations.[5] Homework problems, known in CMP as "Applications, Connections, Extensions" (ACE) problems, or "Mathematical Reflection" questions were not included because the role of homework in the classroom and the assignment of extension-type questions

[3] Following G. J. Stylianides (2009), the analysis employed a definition of *conjecture* as "a reasoned hypothesis about a general mathematical relation [based] on incomplete evidence" (p. 11), in which the modifier "reasoned" distinguishes hypotheses that were conjectures from arbitrary claims.

[4] For a complete curriculum analysis of the CMP texts with regard to reasoning-and-proving, see G. J. Stylianides (2009).

typically varies widely from teacher to teacher. The only materials examined when determining which investigations contained proof-related problems were the student text materials; however, teacher support materials were also referenced to determine the support offered by the curriculum in implementing both the entire lesson and particular proof-related problems.

Method Used to Classify Problems

Problems were initially selected based on whether or not they contained one or more of the following key verbs indicative of proof-related problems: *prove, conjecture, justify, verify, show*, and *convince*. Once a problem had been selected, the complete lesson containing the problem was reviewed to assess whether the lesson offered potential for the problem to engage students in developing conjectures when it was implemented in the classroom. Based on the nature of the activity and subsequent questions, additional problems found in lessons with problems containing key verbs were identified as having potential to engage students in developing justifications, despite not containing key verbs. For instance, if a problem required students to generalize patterns and the teacher support materials suggested possible justifications for a generalization, the problem was selected based on its potential for students to generate justifications as the teacher and students discussed the patterns found. Because the aim of the study was to learn more about how proof-related problems are implemented in the classroom rather than study the implementation of all proof-related problems in CMP, the curriculum analysis targeted the most salient opportunities at the risk of missing some problems with proving potential. Nonetheless, each of the lessons selected for observation contained at least one proof-related problem.

Data Collection

The author observed 6 or 7 investigations that contained proof-related problems (identified as described in the previous section) per classroom, resulting in 49 observations[6] overall. Due to the occasional difficulty in scheduling observations, 6 of these investigations were taught in multiple classrooms, resulting in a total of 43 different investigations observed. The investigations observed spanned a wide range of mathematical content areas, from topics in number theory, such as factors and multiples, to statistical topics, such as measures of center. Three data sources provided information about the practices of implementing proof-related problems:

[5] CMP uses the term *investigations* to name lessons. Within an investigation, there may be one or more problems. These are the tasks that engage students in the mathematics of each investigation. These problems often have several parts, and there are also follow-ups that can be used to assess student understanding as well as provide an extension of the mathematics in the problems.

[6] Some investigations took more than one class period to implement, requiring two observations. Consequently, there is not a one-to-one correspondence between investigations and observations.

(a) teachers' responses to preobservation writing prompts, (b) classroom observations, and (c) teachers' responses in postobservation interviews.

Preobservation Writing Prompts

At least one day before each observation, the author e-mailed three questions to participants to obtain data about their preparation for the lesson. Participants e-mailed their responses back to the author prior to the beginning of each observation. The questions were as follows:

Q1. What changes have you made to the investigation as written in CMP? Why?

Q2. What are your goals for the investigation?

Q3. What problem(s) from the investigation is (are) particularly important? What are ideal responses you hope students will provide to this (these) problem(s)?

Because the preobservation questions were the same regardless of the lesson being observed, e-mailed responses were as accurate and more time efficient for the participants than responses that might have been obtained in preobservation interviews.

Responses to the preobservation questions enriched data collected in observations by identifying possible changes in the written curriculum prior to task implementation (Q1), revealing teachers' interpretations of the mathematical goals of the lesson and whether developing students' competencies in justification and proof were explicit goals (Q2), and determining whether teachers identified proof-related tasks as important in the CMP lessons prior to lesson enactment (Q3).

Classroom Observations

During observations, the author constructed field notes guided by an observation protocol. The observation protocol consisted of questions designed to capture aspects of implementation along the following dimensions motivated by the mathematical task framework (Henningsen & Stein, 1997): task features as set up by the teacher, cognitive demand of the task as set up by the teacher, classroom norms evident during implementation (e.g., expectations for group work during task), task conditions during implementation (e.g., appropriateness of time allotted for task completion), teachers' and students' habits and dispositions during implementation (e.g., a teacher's willingness to allow students to struggle during a difficult task, students' perseverance in finding a solution to a difficult task), task conditions as enacted, and cognitive demands of tasks as enacted. After each observation, the author wrote concise narratives of what happened during the observations by answering questions from the protocol using field notes and digital audio recordings of the observed lesson. Questions whose responses provided documentation of teachers' and students' instructional habits during implementation in narratives included:

1. What questions did students have about the investigation, and what were the teacher's responses?
2. In what ways, if any, did students not follow instructions?
3. How did the teacher assess student progress on the problems in the investigation?
4. How did students respond to the teacher's guidance during the investigation?
5. What changes did the teacher make to the investigation as it was enacted?
6. What questions/responses from students prompted changes?
7. Did she or he restate the explanation, ask for feedback from class, or ask for other justifications to the proof-related problem?
8. When did the teacher ask students to explain their justification further?
9. What kinds of information were assumed/"taken-as-shared" during the discussion of the proof-related problem?
10. Were any explanations given for responses to the proof-related problem highlighted and in what way?
11. What aspects of the lesson discussion seemed to be difficult for the teacher to facilitate?

The field notes did not include analytic comments in an effort to approximate an objective account of the events observed, whereas the narratives included comments and inferences from the researcher.

Teacher Interviews

The author interviewed each participant after having conducted at least four observations. The author decided to interview after an observation if the observation contained episodes of classroom discussion that were particularly pertinent to the goals of the research and needed both clarification and affirmation from a teacher's point of view. In all cases, every attempt was made to conduct the interview as close in time to the actual implementation as possible. Whereas some teachers had a planning period after the class that was observed, other teachers could be interviewed only at the end of the school day. As a result, some interviews were face-to-face, and others were conducted by telephone. The author audio recorded each interview conversation.

The goal of the interviews was to obtain the teacher's perspective on the investigation implementation, as well as to learn about why the teacher responded in particular ways to student thinking during the enactment of proof-related problems. Thus, although some questions such as "What did you think about how the investigation went today?" were standard in all interviews, other questions such as "Why did you decide to highlight [student's name]'s justification?" were more specific to the dialogue during the lesson. By obtaining the teacher's perspective on the investigation implementation, the postobservation interview responses informed interpretations of the teacher's discourse moves after analysis of the field notes.

Data Analysis

Data analysis primarily involved coding classroom observation field notes for teachers' and students' verbal responses and any written work for display to the entire class during the implementation and enactment of the proof-related problems, and then identifying patterns in the responses and written work across and within the observations at each grade level. Once patterns were identified, both the preobservation responses and postobservation interviews were consulted to provide insight into inferences made to determine the patterns.

The analytic process consisted of four stages, 0 to 3. During Stage 0, the author segmented field-note records by changes in either participant structure (e.g., teacher-led, TL, to individual work time, IW) or activity structure (e.g., warm-up, WU, to homework review, HW) as the records were created. This in-the-moment analysis was necessary so that the time spent either in a particular participant structure or activity structure could be noted in the absence of anything audible on the recording. Stage 1 was the first step in analysis after all observations were completed in a teacher's classroom. During Stage 1, the author identified units of analysis at a relatively large grain size, separating events named as *opportunities to prove*—generalizations or conjectures students offered in the context of proof-related tasks—from other kinds of classroom events. The following example shows the kinds of statements students generated during events coded as opportunities to prove:

Seventh-grade teacher:	Look at the boxes that you build. Look at the dimensions and the surface area and draw conclusions. Why are those surface areas so different?
Student:	If the length is small, the surface area might be small.
Sixth-grade student:	Even numbers are composite numbers, and the odd ones are prime numbers.

To establish the reliability of the author's coding of events as opportunities to prove, the author's codes for Stage 1 were compared with codes for Stage 1 analysis generated by a second coder from six complete observations (two observations at each grade level). Between the two independent analyses of these data sets, the coding of events as opportunities to prove was aligned 86% of the time.

In Stage 2, the analytic process distinguished segments of the field notes containing opportunities to prove from instances coded as *proving events*, events during which students provided justifications for the generalizations or conjectures coded as opportunities to prove. The author classified justifications into one of two types: (a) justification consisting of arguments treating a general case or (b) justification consisting of nonproof arguments, for example, using examples or restating a conjecture instead of a justification.

Finally, in Stage 3, utterances from the teacher and the students were coded using a constant comparative method of coding (Glaser & Strauss, 1967), focusing on the teacher's involvement and the students' involvement (teacher's and students' moves) in the discourse during proving events. Elements of the field notes were coded at the turn level; thus, individual student and teacher moves related to each part of the proving process were coded as separate instances. This article features results derived from analyses completed during Stage 3.

RESULTS

The primary goal of this study was to determine the nature of the teacher's and students' participation in the enactment of proof-related problems from the CMP curriculum in middle school classrooms. Observations of the implementation of proof-related problems not only revealed how students demonstrated proving-related competencies but also defined characteristics of the pedagogy of proving that emerged during implementation. This section presents descriptive statistics about the opportunities to prove, and presents analyses of episodes from the observations to illustrate various teacher and student moves in response to the justifications during the process of enacting proof-related problems (see Figure 2, p. 367). Figure 2 presents the various paths that the proof-related problems followed when implemented in the classroom. The author developed these trajectories with the assumption that the classroom community either enacted a task, meaning that students provided a justification to a generalization that served as a response to the task, or did not enact a task. In either case, Figure 2 shows the teacher's and students' feedback to a generalization after task enactment took one of three paths.

Implementation of Proof-Related Problems

Two conditions, determined prior to the observations, were used to identify a proof-related problem as being implemented: (a) the teacher provided students with instructions to work on the proof-related problem(s) at the start of the activity without changing the problem(s) as written in the text; and (b) the teacher elicited students' responses to the problem(s) during whole-class discussion. Although it was possible for the teacher to modify a problem written in the text as a proof-related problem to an alternative proof-related problem, this did not happen in any of the observations.

Overall, 71% (52 of 73) of the problems identified as proof-related prior to the observations were proof-related as implemented (see Table 1). Of the remaining 21 problems, two factors contributed to nonimplementation. For 15 of the 21 problems, lack of time was the primary factor that prevented the problems from being implemented. Although teachers instructed students to work on the proof-related problems during the investigation, there was not enough time after the activity to discuss answers to some or all of the problems. A related but less common factor was the teacher's choice not to have students work on the problems in the moment of task implementation (6 out of 21 problems). When asked in postobservation interviews why they decided not to assign those problems during class work time, in all instances teachers cited a concern that there would not be enough time to get

Table 1
Problems Implemented Compared to Problems Written

Gr.	No. of different investigations observed	No. of different problems used in observed investigations that were "proof-related as written"[a]	No. (percentage) of different "proof-related as written" problems observed in at least one of the different investigations that were "proof-related as implemented"
6	19	40	31 (78%)
7	12	16	13 (81%)
8	12	17	8 (47%)
Total	43	73	49.94 (71%)

[a]Some, but not all, investigations were observed in multiple teachers' classrooms. To minimize the unintended sampling effects of observing one lesson in multiple classrooms, a proof-related problem was counted once as implemented if it was observed to be implemented in at least one classroom, as shown in the third column of the table.

through the material in the lesson. Another potential way that a proof-related problem might not be implemented is if the teacher modified instructions when assigning a proof-related problem to change the cognitive demand of the problem as written. However, this was not observed in the sample of observations conducted.

As evident in Table 1, the sixth-grade investigations observed contained more proof-related problems as written, and the data suggest the possibility that the sixth- and seventh-grade classroom communities observed in this study supported the implementation of more proof-related problems than the eighth-grade classrooms. However, it is worth noting the limitations of the data collection methods that affect cross-grade comparisons. First, there were differences in the sampling of teachers and lessons across the three grade levels. The sample consisted of three sixth-grade teachers, compared to only two teachers in each of seventh and eighth grades. Also, there were 19 lessons observed in sixth-grade classrooms compared to only 12 lessons each in seventh and eighth grades. Most important, the author sampled from the set of CMP investigations containing proof-related problems to determine the investigations to be observed; this sample is not a representative sample of all lessons in the participating classrooms. As such, the sample cannot be considered representative of the proof-related instruction in the participating classrooms.

Despite these limitations, the data appear to support G. J. Stylianides' (2007, 2009) conclusions about the irregularity of reasoning and proving opportunities as written in the CMP curriculum. In the sixth-grade curriculum, there were more proof-related problems as written in the lessons observed (40 problems from 19 lessons [see Table 1], or roughly two problems per investigation) than in seventh grade (16 problems from 12 lessons, or roughly one problem per investigation) or eighth grade (17 problems from 12 lessons, or roughly one problem per investigation). The data also appear to suggest that there were more proof-related problems implemented per investigation in sixth grade (31 problems from 19 lessons, at a rate of 1.6 problems per investigation) than in seventh grade (13 problems from 12 lessons, at a rate of 1.1 problems per investigation) or in eighth grade (8 problems from 12 lessons, at a rate of 0.67 problems per investigation).

Opportunities to Prove

Opportunities to prove were possible outcomes of the implementation of proof-related problems, as shown in the trajectory of proof-related tasks as observed provided in Figure 2 (on p. 367). Each proof-related problem in the lessons observed was written, at minimum, to elicit conjecture development; therefore, students' expected responses to these problems were conjectures. An opportunity to prove in the classroom surfaced once a student offered a response in the form of a conjecture. Even if the proof-related problem as intended did not explicitly ask students to provide a justification to support the conjecture, the emergence of a conjecture opened the door for the teacher or other students to ask for a justification before the generalization was accepted. Table 2 shows the extent to which proof-related problems generated these opportunities to prove.

From the 52 proof-related problems implemented in observations across the seven classrooms, students' responses presented 109 opportunities to prove. Of the 109 opportunities to prove (conjectures) that emerged, students did not justify in nearly half of the opportunities to prove at each grade level. Because CMP emphasizes whole-class discussion of students' work, it is not surprising that there are

Table 2
Opportunities to Prove Generated From Problems Implemented

1	2	3	4	5
Gr.	No. of problems proof-related as implemented (see Table 1)	No. of opportunities to prove generated from proof-related problems as implemented[a]	No. (%) of opportunities to prove supported with a student justification	No. (%) of opportunities to prove not supported with a student justification
6	31	65	41 (63%)	24 (37%)
7	13	28	11 (39%)	17 (61%)
8	8	16	7 (44%)	9 (56%)
Total	52	109	59 (54%)	50 (46%)

[a]As each problem was implemented, more than one opportunity to prove could be generated. As such, the totals reflect situations in which more than one opportunity to prove emerged during the implementation.

more opportunities to prove than proof-related problems implemented, because more than one student would likely have the opportunity to share a conjecture in response to a problem. It is important to note that the opportunities to prove as reported in Table 2 are not counts of conjectures that make different claims. For example, if, at distinct times during the discussion, two students both conjectured that the sum of two odd numbers is even, two separate opportunities to prove were counted. Whereas multiple opportunities to prove (conjectures) could be generated from one proof-related problem, students provided either no justification for an opportunity to prove or one justification for an opportunity to prove.

The following results shed light upon the teachers' and students' moves in these situations in which opportunities to prove emerged as a result of proof-related problem implementation. The results first examine teachers' and students' moves during the episodes in which opportunities to prove were not justified and resulted in proof-related problems not being enacted—nearly half (50 of 109, or 46%) of the cases. These results are followed by a presentation of patterns in teachers' and students' moves during the episodes in which justifications accompanied opportunities to prove, resulting in the enactment of those proof-related problems—episodes referred to as *proving events*.

Generalizations Without Justifications: Proof-Related Problems Not Enacted

As shown in Figure 2, teachers responded in one of three ways when students did not provide justifications for their conjectures: (a) teachers gave no feedback to elicit a justification, (b) teachers sanctioned the conjecture as valid without justification, or (c) teachers asked other students to state whether or not they agreed with a student's conjecture. Data presented in Table 3 shows the types of moves carried out by the classroom community—the teacher and students—in response to opportunities to prove that students did not justify. These results have not been segregated by grade level as the total number of opportunities to prove that were not justified was small at each grade level. Of the 50 opportunities to prove, 21 (42%) that were not justified did not receive feedback by either the teacher or the students; in essence, the classroom community ignored these emergent opportunities to prove. In contrast, the remaining 29 (of the 50) opportunities to prove that

Table 3
Teacher Moves in Response to Opportunities to Prove That Were Not Justified by Students (From Table 2, Column 5)

Type of move	Frequency	Percentage
No response given	21	42%
Teacher sanctions conjecture	17	34%
Teacher asks for class input	12	24%
Total	50	100%

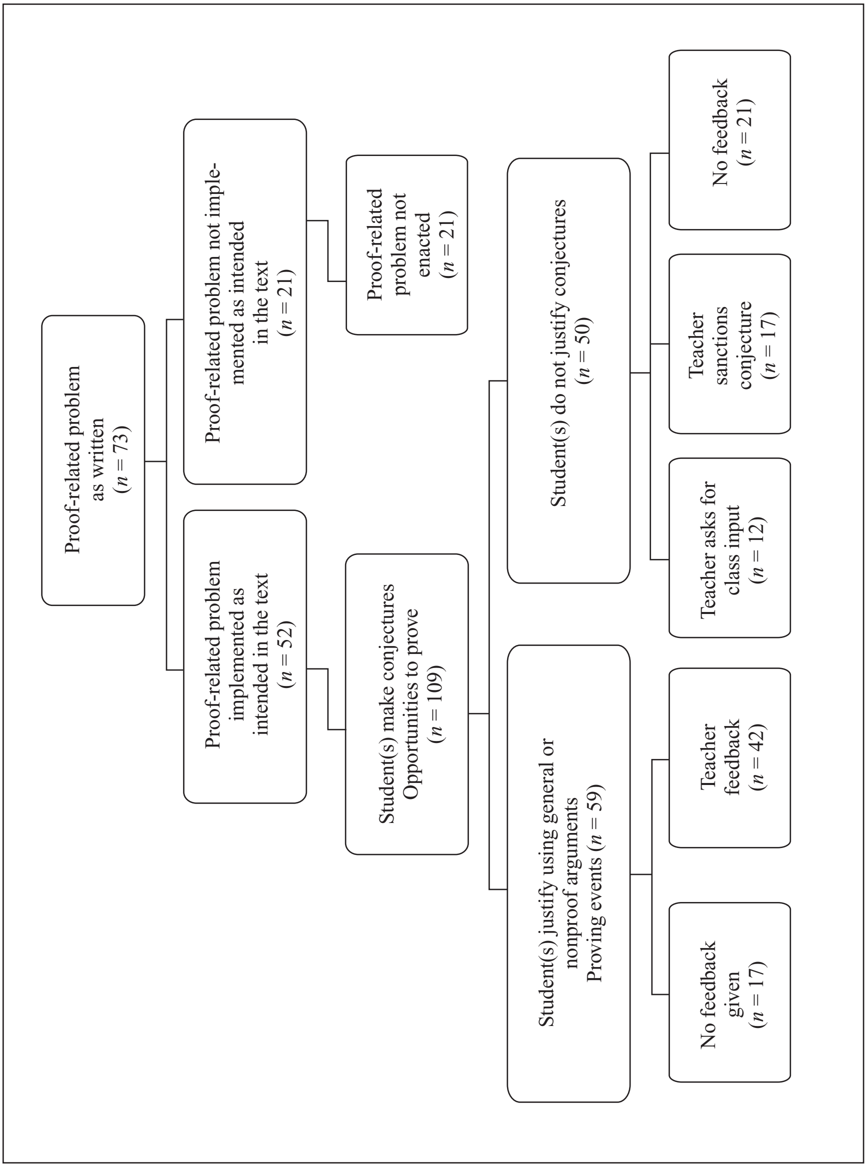


Figure 2. Trajectory of proof-related tasks as observed.

were not justified received some sort of feedback from the classroom community. The feedback consisted of either (a) teachers sanctioning the use of the generalization without obtaining sufficient justification (17 events or 59% of the 29 opportunities to prove not justified that received feedback) or (b) teachers requesting input from other students but not providing a response or subsequent follow-up (12 events or 41% of the 29 opportunities to prove not justified that received feedback). There was no student-initiated feedback to opportunities to prove that were not justified. That is, if a student provided a generalization, but the student or other members of the classroom community did not ultimately justify the generalization, no student response was noted to these unjustified generalizations.

The following excerpts from the observations illustrate the types of moves accompanying opportunities to prove that were not justified by students. An excerpt from Ms. Clinton's sixth-grade class as they developed conjectures about the parity of the product of odd and even numbers exemplifies the classroom discourse in instances in which the teacher sanctions, or approves, a student's conjecture that is not justified:

Ms. Clinton: Josh, you said even?
Josh: Two times three equals six. Six times three equals eighteen.
Ms. Clinton: No matter what, it comes up?
Josh: Even.
Ms. Clinton: Even. Easy, easy.

Note that after Josh's presentation of two confirming examples, Ms. Clinton reconfirms Josh's generalization that "No matter what, it comes up even," a move coded as sanctioning Josh's conjecture. Her satisfaction with Josh's method of justification implicitly approves the use of confirming examples to show the truth of statements about sums and products of odd and even numbers. Further evidence from a postobservation interview with Ms. Clinton confirmed that she believed examples-based justifications to be satisfactory for sixth graders:

Interviewer: When you asked them to "prove it" to you . . . what explanations did you anticipate students would provide?
Ms. Clinton: Proving it to me using numbers or using blocks or drawing on paper. Some of the brighter kids will use numbers, go beyond the blocks. I want them to use different numbers with the odd ones.

Her responses indicate how she values the different ways that students justify their reasoning, but she does not explain what mathematical content she is expecting them to use in their justifications. There is little indication that she desires a certain type of response indicative of correct reasoning about sums and products of odd and even numbers, aside from a hope that they will use different "numbers," or examples, to support their conjectures about sums and products of odd numbers. More problematic, however, is her belief that using numbers, as opposed to alternative representations such as the "blocks," is indicative of more sophisticated mathematical reasoning. The belief that justifications using numerical examples represent more sophisticated mathematical reasoning than justifications using statements about the blocks contradicts the ideas presented in the textbook. In the lesson,[7] tile

[7] Investigation 2.2 from CMP sixth-grade unit *Prime Time* (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2002)

models, or unit blocks as used in Ms. Clinton's lesson, are intended to embody definitions of odd and even numbers that should be used in general arguments about sums and products of odd and even numbers, such as the definition that an odd number is a even number with an extra tile (or block). Further, this belief may explain why Ms. Clinton's responses to Josh appear to sanction the use of empirical argumentation or examples-based reasoning. Given the wealth of literature on students' preference for empirical arguments, Ms. Clinton's move may have only reinforced her students' naïve beliefs about mathematical proof.

The following excerpt is from one of the 12 events in which teachers asked students to provide input regarding a student's conjecture. In this case, the teacher polled the students to assess whether the class believed the student's conjecture to be correct or incorrect and then ultimately sanctioned the conjecture without supplying or requesting justification. The excerpt comes from a discussion in Ms. Kelly's sixth-grade class of how to use prime factors to determine the least common multiple (LCM) of two numbers. Although Ms. Kelly introduced the lesson as an "unusual" activity in which the students would learn a procedure she was going to "give" to them and then apply it when solving problems, one student, Max, arrived at a conjecture early into the discussion of the procedure:

Max: [Using the provided example of 12 and 10, both of which are factored into their prime factorizations on the overhead projetion screen] So you have the 12. Then you multiply the numbers that weren't circled [distinct prime factors not common to 12 and 10], that's 30. Then you multiply by [the common prime factor] 2.

Ms. Kelly: Do you think his answer is going to work for all numbers?

Class: No.

Ms. Kelly: The authors want you to know this is the way you do it. Max is right. Those of you who said it wasn't going to work all the time, that's good, because in math we have things that aren't always right.

Max's initial generalization in lines 3 to 6, that the least common multiple of two numbers can be found by finding the product whose factors consist of one copy of each of the common prime factors and the distinct prime factors, is called into question by the class in line 8. However, Ms. Kelly makes the decision to sanction Max's method in lines 9 and 10 without eliciting a justification from Max or other students or providing one herself.

An interesting aspect of the classroom discourse during these episodes, in situations in which opportunities to prove did not materialize into proving events, was the abbreviated duration of the discussions. In all but 2 of the 29 episodes in which feedback was given during an opportunity to prove not justified, the dialogue between the teacher and students regarding the opportunity to prove ended after the teacher's move (as exemplified in the excerpt presented earlier from Ms.

Clinton's class). In 12 events in which the teacher requested feedback from other students, the teacher used the format of a "class poll" to elicit the feedback. The teacher would ask students questions such as, "Do you think they explained things?" and, given either no response from the class or some nodding heads, the teacher moved on to the next problem.

Justifications Emerging From Opportunities to Prove: Proof-Related Tasks Enacted

A proving event consisted of a classroom episode during which a justification accompanied an opportunity to prove. Analysis of these events began by classifying the justifications into one of two categories: (a) the justification was sufficient as a mathematical proof or treated the general case but fell short of being a valid proof, or; (b) the justification relied solely upon confirming examples, represented a rationale (G. J. Stylianides, 2007, 2009), or failed to address the conjecture. In the first category, students' justifications are, at minimum, on the right track to achieving the generality needed for proof, whereas the second category includes justifications that confirm that the conjecture works for certain cases or responses that suggested misunderstanding of the intent of the task. Table 4 shows how many of the opportunities to prove that students generated were supported with justifications—turning opportunities to prove into proving events—as well as how many of the justifications were general in nature compared to those that relied upon examples (nonproof arguments).

Table 4
Number of Proving Events (Opportunities to Prove Supported With Justifications per Grade Level)

Gr.	No. of opportunities to prove[a]	No. of proving events[b] ($n = B + C$)	No. of proving events with justifications using general arguments (B)	No. of proving events with justifications using nonproof arguments (C)
6	65	41	19	22
7	28	11	6	5
8	16	7	3	4
Total	109	59	28	31

[a]From Table 2, Column 3. [b]From Table 2, Column 4, No. of opportunities to prove supported with at least one justification.

As shown in Table 4, regardless of grade level, roughly half of the justifications students produced in proving events were nonproof arguments, whereas the others attempted to treat the general case. Because there were no clear differences by grade level, and the total number of proving events is small, the results that follow regarding patterns within sequences of teacher and student moves are not separated by grade level.

Justifications receiving no feedback. As shown in Table 4, students provided justifications for 59 of the initial 109 (54%) opportunities to prove across all observations of the seven classrooms. However, almost 30% (17) of these 59 proving events consisted of justifications that received no feedback—from the teacher or the students—as to the sufficiency of the argument. In these cases, justifications that used general arguments (10 out of 17 events) were more likely not to receive feedback from the teacher and/or students than those arguments that relied solely upon confirming examples (7 out of 17 events).

As illustrated in the following excerpt, these proving events in which students' justifications received no feedback from the teacher or students typically contained no more than three turns. A turn was counted as either the teacher's or a student's verbal input into class discussion. The following excerpt comes from a portion of the discussion in Ms. North's seventh-grade class, as students discuss why the product of two negative numbers results in a positive number.

Ms. North: Yesterday, you saw those two negative signs make it a positive . . . today's goal is why does that work. Adrienne?

Adrienne: I counted up the negatives so I know that the answer is negative then 3 times 4 times 1 equals 12.

Ms. North: Are people in agreement? Brittany?

Brittany: Since there is a positive and a negative, then it would be negative automatically.

Class: Woo hoo . . .oh . . . oh.

Ms. North: Hey, no trash talking. This is math. Time me, I'm going to give directions in less than a minute.

At the beginning of the excerpt, Ms. North charges the class with understanding why two negative numbers multiply to make a positive, given that the students saw confirming examples of the rule during class the day before. In the first turn, Adrienne's justification involved a recounting of the procedure she used supported with an example. In the second turn, Ms. North asks for class input and calls on Brittany, a student with her hand raised. In the third turn, Brittany's response provided a generalization of the parity of positive and negative products. Ms. North follows with instructions to start the next phase of the lesson, a move that sidesteps further discussion of Brittany's or Adrienne's justifications. As a result, Ms. North missed an opportunity to discuss what it means to validate that a rule is true in mathematics, by discussing how both arguments simply restated a rule and used specific examples to confirm the rule instead of providing a general argument.

Justifications receiving teacher feedback. Although 30% of the 59 proving events did not involve feedback from the classroom community, 70% (42) of the 59 proving events received teacher feedback (see Table 5), which took one of three forms: (a) sanctioning of justification (positive appraisal); (b) questioning related to justification; or (c) requesting of feedback from other students. Regardless of the kind of justification students produced or the kind of feedback given, the classroom

Table 5

Teachers' Moves During Proving Events for Which Justification Received Teacher Feedback by Justification Type

Teacher move	Type of justification	No. of instances
Sanctioning (positive appraisal)	General arguments	9
	Nonproof arguments	9
Questioning	General arguments	2
	Nonproof arguments	10
Requesting feedback from other students	General arguments	8
	Nonproof arguments	4
Total		42

discussions of the justifications were short-lived; the number of turns in proving events for which the teacher provided feedback did not exceed five turns. Thus, even in the cases in which teachers made instructional moves to address students' justifications, there likely was not enough discussion about the validity of students' arguments to change students' conceptions.

Teachers' feedback differed based on the kind of justification that was provided. Table 5 shows how the teachers' moves varied depending upon the kind of justification. The excerpts that follow exemplify some of those moves.

As shown in Table 5, teachers positively appraised 9 justifications out of a total of 19 proving events (out of 42) that utilized a general argument, sanctioning the use of the justification. An example of this type of feedback is given in the following excerpt from Ms. Xavier's seventh-grade class, as they discuss how the volume of a prism changes as the number of sides increase (e.g., from a triangular prism to a rectangular prism):

Ms. Xavier: Which one do you think holds more, a cylinder or rectangular prism?

Joy: The cylinder. The closer to a circle it is, the greater the area.

Ms. Xavier: There you go. The closer we get to a circle, the larger the volume is going to be.

In this episode, Joy provides a general argument "the closer to a circle it is, the greater the area" to justify her conjecture of "cylinder." Ms. Xavier apparently approved of Joy's conjecture and justification, using the statement "There you go." Ms. Xavier, perhaps accustomed to the sparse explanations that middle school students tend to provide, seems to have concluded that Joy understood that given a constant height, as the number of sides of the polygonal base of the prism increases, and its shape approaches that of a circle, the area of the base approaches its maximum and therefore maximizes the volume. However, Joy's classmates may have found Joy's justification unclear or even tangential to the question posed by Ms. Xavier. Considering A. J. Stylianides' (2007) definition of proof, and its emphasis on the importance of the understanding of the classroom community, the discourse around Joy's conjecture was insufficient to determine how other students evaluated Joy's reasoning. Ms. Xavier's move sidestepped input from other students in the classroom community, and missed a key opportunity to reinforce the importance of clarity and completeness when producing justifications.

In the 23 out of 42 proving events for which students provided nonproof arguments, teachers preferred to sanction the justification (9 events) or to question the justification (10 events) rather than to ask other students for feedback (4 events). The next two excerpts, taken from one observation in Ms. Kelly's sixth-grade class, illustrate how her use of different kinds of instructional moves depended upon the type of justification students provided during the implementation of a single problem. In the investigation, students explored definitions of odd and even numbers and generated conjectures about the parity of the result when adding or multiplying odd and even numbers. The textbook presented tile models to represent the differences between odd and even numbers, showing models that were two tiles high to represent a factor of two in each number. When constructed, rectangles formed by joining columns two tiles high represent even numbers, and rectangles formed by joining columns two tiles high with one extra square attached to one side (see Figure 3) represent odd numbers.

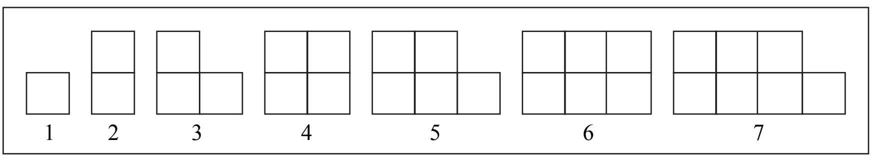


Figure 3. Tile models of numbers 1 through 7.

The implementation of the investigation took two full days. In this excerpt, Ms. Kelly had asked for volunteers to share their conjectures and justifications for a problem in which students are asked to make a conjecture about the parity of the result of the sum of two odd numbers.

Jason: All odd numbers added together make an even. The odd numbers always [have] a square hanging off and if you have another odd number with a square number hanging out they fit together.
Ms. Kelly: Any comments? Anyone have any additions to it?
Tyler: I just said that . . . I can't think of it any other way . . . adding two together you take the two extra pieces off and then add them together.
Ms. Kelly: Okay. Who wants to do problem c? Adding odd and even?

The excerpt illustrates, especially in lines 4 and 7, how Ms. Kelly's perceived indifference to the students' justifications provided in lines 1–3 and 5–6 served as an implicit positive appraisal of the students' arguments. Tyler's argument was simply a restatement of Jason's justification, and Ms. Kelly was satisfied enough with both justifications that she moved discussion to the next problem. Although both Jason and Tyler referred to the tile models in their arguments, they did not make an explicit connection between the tile models and definitions of odd and even numbers. Ms. Kelly missed a key opportunity to help students understand, concretely, why the sum of two odd numbers makes an even number.

The next excerpt, from the same class and the same discussion, features an examples-based justification produced in response to the conjecture that the product of two odd numbers is odd:

Ms. Kelly: Anyone want to do problem e? Product of two odds? Okay, I've got lots of nonvolunteers . . . Sophie?
Sophie: I think it [the product of two odd numbers] is odd because . . . 5 * 5 is 25 and they are both odd numbers.
Ms. Kelly: What did you notice about your [tile] models?
Sophie: It's like a puzzle. When you add the extra pieces, there is one out.
Ms. Kelly: Any questions?

In this excerpt, Sophie initially presented an examples-based argument to support her conjecture that the product of two odd numbers is odd (lines 3 to 4). Ms. Kelly responded by asking Sophie to think about the tile models, and Sophie offered a justification (line 6) more general in nature than her initial justification (lines 3 to 4). Ms. Kelly seemed satisfied with Sophie's revised justification, as she then opened up the discussion to other students (line 7). Both of these excerpts from Ms. Kelly's class exemplify typical discourse sequences in which students offered justifications concurrently with conjectures, but they also highlight an important facet of instructional practice when teachers evaluate students' justifications. In the excerpts already presented, Ms. Kelly was satisfied when students provided justifications that involved using tile models. In fact, an examples-based justification was presented that utilized the tile models and Ms. Kelly did not probe for further explanation. In a sense, Ms. Kelly was keeping her "eyes on the prize" during class discussion, in which the prize was a justification based on the tile models.

Evidence from both the curriculum and Ms. Kelly's responses to preobservation writing prompts confirmed this goal. The suggested response provided in the CMP teacher support materials to the problem of the sum of an even number and an odd number is: "The sum of two odd numbers is even. The tile models for the odd numbers each have an extra square. If we combine the models, we can pair the extra squares to form a rectangle" (Lappan et al., 2002, p. 35e). In her response to a preobservation writing prompt about the responses she expected from students, Ms. Kelly stated: "An ideal response to each would include an explanation using the model or something equally substantive that the students have created. The explanations should include ideas about extra blocks sticking off the end." Ms.

Kelly's focus on a particular set of accepted statements about tile models, namely "ideas about extra blocks sticking off the end," instead of attention to a mode of argumentation that treated a general case resulted in an overemphasis of the importance of using accepted statements without consideration as to the completeness, form, or generality of the argument.

Justifications receiving student feedback. Teachers asked students to provide comments or questions related to other students' justifications for roughly 20% of the proving events (12 proving events out of 59) and were more likely to ask for students' feedback to justifications based on general arguments than nonproof arguments (see Table 5). In only one event did another student respond to the teacher's request with a follow-up question, thus continuing the discussion of the justification. Although this is only one case out of 12 proving events in which students took up the opportunity to critique others' justifications, there is significance in this episode in that it was the only event in which the student's justification closely approximated proof. In other proving events, students primarily produced confirming examples to justify conjectures or produced general statements that were either incorrect mathematically or were not based on a deductive argument (such as using the statement "Two odd numbers sum to make an even numbers because numbers go even, odd, even, odd and after two odds you will get an even"). However, the proving event featured in the following excerpt closely approximates proof because the justification provided uses statements referring to a generic, visual example in a deductive way. In the episode, the teacher played a minor role in the discourse but the other students provided the critical community necessary for a mathematically sophisticated proof to evolve.

The excerpt comes from an observation of Ms. Conroy's eighth-grade class during an investigation that engaged students in producing a visual proof of the Pythagorean theorem via the context of constructing puzzles from two puzzle frames (Investigation 3.2, *Looking for Pythagoras*, Lappan et al., 1998). The pieces of the puzzle were three squares, one square with side length a, one square with side length b, and one square with side length c, where $a^2 + b^2 = c^2$. Additionally, students were given eight right triangles, each with side lengths a, b, and c, where $a^2 + b^2 = c^2$. The two puzzle frames were identical squares of side length $a + b$. To begin the activity, students discussed how the side lengths of the squares compare to the side lengths of the triangle (part A). This helped students establish that the squares represent the squares of the side lengths of the triangles, thus serving to link the "conjecture" of the Pythagorean theorem to the puzzle pieces. After identifying the relationships between the pieces, students arranged the pieces to fit exactly into the two puzzle frames (part B). In parts C and D, students were asked to think about the relationships between the areas of the three square puzzle pieces as well as the relationships between the side lengths of the triangle.

After Ms. Conroy introduced the investigation, students worked to cut out the puzzle pieces from a worksheet and answer parts B through D of the problem. Ms. Conroy canvassed the room as students worked, making sure students were on task. To launch the discussion, Ms. Conroy called one student, Becky, up to the overhead projector to demonstrate how to fill in the puzzle frames with plastic pieces. In one frame, Becky placed two triangles together to make a rectangle and placed the rectangular piece in the upper left corner of the frame, and then duplicated the process, situating the second rectangle below the first with its long sides perpendicular to the long sides of the first. She inserted the square with side length *b* in the upper right corner of the puzzle frame with the square with side length *a* directly beneath. In the second frame, another student, named Jessica, fitted a triangle into each of the four corners and then filled in the middle of the frame with the square of side length *c* (see Figure 4).

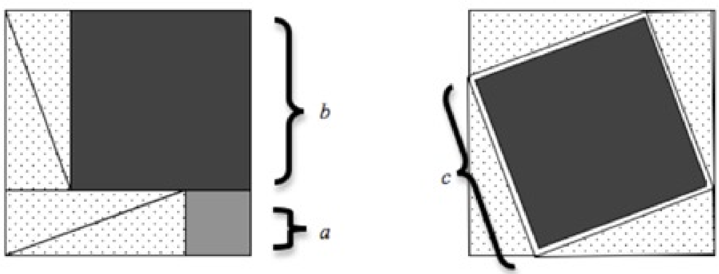


Figure 4. Visual proof of the Pythagorean theorem.

Kyle: Okay, I don't get what this tells us.

Becky: It means that with 4 triangles both the large box and the small and medium box fit.

Ms. Conroy: Go up and point out what you mean.

Becky: [at overhead projector] We all know that this [pointing to one of the two squares of side length b] is the same size as these [pointing to the longest leg of a triangle]. These 4 triangles [pointing to triangles in one puzzle frame] are the same size as these 4 [triangles in other puzzle frame]. They fit perfectly in there just like these so the two smaller squares must be equal to the larger square.

Kyle: How does that prove the theory?

Ms. Conroy: Our job today is to try to prove the theorem, not just accept it.

Jerome: Yeah, how does that prove that the two squares are equal to the larger square?

Kyle: Are you trying to say that the two squares are equal to the larger square?

Becky: I just proved that the sides are the same length because you can put them into a square with each other.

This excerpt shows a potential key to the development of students' abilities to prove, namely, the emergence of a critical community in which students' justifications are vetted before being accepted. It also raises the issue of the teacher's and students' roles in establishing proof. Because the title of the investigation was "Puzzling Through a Proof," students in the excerpt seem aware of the need for proof, especially in lines 11, 13, and 14. Although the mode of argument representation was given to the students in the activity, the goal, as achieved by Becky in lines 5–10, was to understand how the visual argument established the truth of the Pythagorean theorem. The most significant part of the excerpt from an analytic point of view was Kyle's response in line 11. His questioning of how Becky's argument proved the theorem was unique among the hundreds of questions, both by students and teachers, posed during proving events observed. But, his question is also an absolutely essential move to promote the development of an appropriate proof, namely, a move that requires students to explain why their arguments are sufficient to justify their conjectures.

When asked in a post observation interview whether Jessica and Becky had produced a proof, Ms. Conroy claimed: "If your definition of proof is just like getting why it works, then I would say yes. It was proof for them." Ms. Conroy's answer suggests that she recognizes a difference between the students' arguments and what the discipline of mathematics considers as proof, but her actions in the classroom reveal that she is unsure of how to impart her disciplinary knowledge. She is not alone; no other teacher was observed asking how students' justifications suffice as proof. Given that this is the only investigation in the entire CMP series in which proof is the main activity of the investigation, it is no surprise that the question of the sufficiency of an argument as proof arose in discussion. Given that there were 42 other investigations observed during the study (see Table 1), the excerpt also suggests that proof remains elusive in middle grades mathematics classrooms.

DISCUSSION AND IMPLICATIONS

This research investigated teachers' and students' involvement in discourse about proof-related problems from the CMP curriculum when experienced teachers well trained in implementing CMP tasks enacted such problems in middle school classrooms. The results show that students regularly responded to proof-related tasks by offering conjectures that created opportunities to prove, and students provided justifications for nearly half of the conjectures, resulting in proving events. Although students were regularly involved in responding to proof-related tasks, teachers in the classrooms observed did not provide sufficient feedback to sustain discussions about students' conjectures and/or justifications. Findings of this research also indicate that when a teacher provided feedback to students' justifications, it was not sufficient to establish standards for proof in a mathematics classroom. For instance, teachers were just as likely to sanction a justification with a positive appraisal if it was a justification based on nonproof arguments as a justification based on general arguments.

The lack of general arguments, and the predominance of examples-based justifications during the enactment of proof-rich tasks are not surprising considering the literature on students' tendencies to generate and prefer empirically based justifications as proof. Generating mathematical proof can be considered a demanding, "high-level cognitive activity" (Henningsen & Stein, 1997, p. 529). Henningsen and Stein identified several factors that contribute to the decline of high-level cognitive activity as tasks are enacted, including task challenges becoming nonproblems, time constraints, and lack of accountability. Findings of this study suggest that for proof-related problems in particular, time constraints and, more important, a lack of sufficient discussion about students' justifications to establish standards for proof-related activity, contribute to the reduction of cognitive demand for these problems. Given that the results were based on data from observations of classrooms and teachers that were, in many ways, well equipped for meaningful implementation of such proof-related tasks, one might expect even more superficial experiences in reasoning and proving in classrooms using curricula not developed to meet the NCTM Standards or teachers who are inexperienced or unfamiliar with the curricular materials.

Lack of Time

Time constraints are a given in school settings, and the brutal honesty of the clock forces teachers to prioritize the aspects of a lesson they believe to be most critical. Considering the results of Knuth's (2002b) research on secondary school teachers' beliefs about the value of proving in school mathematics, it is likely that, in response to time constraints, teachers would tend to minimize time spent on proving activities. Findings from the data collected in observations and postobservation interviews indicate that the reasons for proof-related problems not to be implemented during lesson enactment: Either the teacher perceived a lack of time for students to work on the problems or for the class to discuss their work before problems were assigned (6 out of 21 instances, or 29% of such occurrences), or the teacher chose not to discuss students' work after students' small-group work (15 out of 21 instances, or 71% of such occurrences).

The effects of time restrictions wreak havoc on the process of proving, potentially more so than on other kinds of mathematical activity. For example, if teachers do not have time to present a particular solution method or discuss a difficult problem, the result is a missed instructional opportunity that may or may not be remedied by adjusting plans for subsequent lessons. In terms of learning outcomes, students may not have adequate time to discuss a topic with the teacher and their peers, and they may only develop a surface-level understanding of the content. However, as problematic as this situation is, students are not taught erroneous concepts. When a teacher does not have enough time to discuss why an examples-based justification is not sufficient as proof, the result is not only a reduction in the cognitive demands of the proof-related task, but it may also be an implicit sanctioning of examples-based justifications as viable means of establishing knowledge in mathematics.

Setting Standards for Proof-Related Instruction

Existing research indicates there is a symbiotic relationship between teachers' and students' actions during the process of generating proof in whole-class discussions (Martin, McCrone, Bower, & Dindyal, 2005). Drawing from the theory of the didactical contract (Brousseau, 1984, as cited in Herbst, 2002a), Herbst asserts that the teacher's responsibility in classroom episodes in which students are involved in generating proof is to help students learn to produce arguments that meet standards of appropriate justifications and proofs in mathematics. In the middle school classrooms observed for this research, the sparse nature of teachers' involvement in providing feedback and instructional support during proving events suggests that the assumptions underpinning the didactical contract between teachers and students during proof-related activity were not sufficient to support meaningful learning about the standards for justification and proof. Throughout the observations, teachers' and students' feedback to justifications during proving events consisted of fewer than five turns between teachers and/or students. No evidence surfaced in the observations that students were expected to hold their classmates accountable for ensuring that their arguments provided a sufficient justification for their conjectures. Of the 59 proving events, 30% never received any feedback whatsoever. Of the 70% that received feedback, there was only one event in which a student revised an examples-based justification with an argument treating a general case. Further, the only time a member of the classroom community raised a question such as "How does this prove the theory?" occurred after a student had provided a justification nearly sufficient for proof, and it was a question posed by a student rather than the teacher.

Teachers' lack of feedback may be a pedagogical move to shift authority to students to determine what counts as acceptable justification. Reform in mathematics education over several decades has emphasized more constructivist, "student-centered" methods of teaching, methods that have been interpreted as downplaying the role of "telling," in which the teacher disseminates information to students (Smith, 1996). However, Lobato, Clarke, and Ellis (2005) note several reasons why telling is needed in mathematics classrooms and, in the case of proof, students' naïve conceptions from prior experience in providing proof in nonmathematical contexts may be incommensurable with the concept of proof in the discipline of mathematics (Fischbein, 1982). For discipline-specific conventions and norms, students may not be able to discover this knowledge without guidance from a teacher. Chazan and Ball (1999) suggest several ways to provide feedback that can support students in learning the norms of mathematics without telling.

What is absent in existing research, and within teacher support materials for middle school curricula such as CMP (Stylianides, 2009), are indications of appropriate standards for proving at the middle grade levels that could assist teachers in understanding how to provide feedback to students' justifications and proofs. Although teachers participating in this study received extensive support in how to teach with the CMP materials and had several years of experience in implementing

CMP lessons, the teachers identified proof-related problems as important in only 6 of the 39 observations. Further, their statements about their primary goals for the lesson implementations indicate skepticism of students' abilities to generate justifications and proofs:

Sixth-grade teacher:	Some of them will not understand because developmentally they aren't there yet. Some will catch on a bit and others will understand to justify their answer.
Seventh-grade teacher:	It may come up from the kids why this is, and then we will go with it . . .
Eighth-grade teacher:	Part C & D [explaining the proof of the Pythagorean theorem] are important. If they can really put their ideas into words, it is especially impressive.

The language that the teachers used to describe their expectations reveals some uncertainty about whether students are capable of understanding conceptually the mathematics in the investigation, let alone justifying their thinking. Justification is seen as something "especially impressive" or something for students who are developmentally ready. Teachers' beliefs about the role of justification and their mathematical understanding of what constitutes an appropriate justification are possible explanations for why so few of the students' justifications received feedback. In terms of the didactical contract, if the majority of the students are not expected to produce justifications of their mathematical thinking, and the instructional materials fail to make explicit what is expected in students' responses to proof-related tasks, it is not within the teacher's responsibility to ensure that students learn what constitutes acceptable mathematical justification and proof.

In summary, this study shows that even in the best of circumstances—teachers with ample experience and professional development using a curriculum that provides multiple opportunities for justification and proof—instruction that supports students' understanding of justification and proof at the middle school level is quite superficial. This suggests that greater emphasis is needed for middle school teacher preparation, professional development, and curricular support to make justifying and proving a routine part of middle school students' opportunities to learn.

REFERENCES

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*, 359–387.

Chazan, D., & Ball, D. (1999). Beyond being told not to tell. *For the Learning of Mathematics, 19*(2), 2–10.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational evaluation and policy analysis, 25*, 119–142.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9–18, 24.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Hawthorne, NY: Aldine.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *CBMS issues in mathematics education: Vol. 7.*

Research in collegiate mathematics education III (pp. 234–283). Washington, DC: American Mathematical Society.

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, 28*, 524–549.

Herbst, P. G. (2002a). Engaging students in proving: A double bind on the teacher. *Journal for Research in Mathematics Education, 33*, 176–203.

Herbst, P. G. (2002b). Establishing a custom of proving in American school geometry: Evolution of the two-column proof in the early twentieth century. *Educational Studies in Mathematics, 49*, 283–312.

Herbst, P. G. (2006). Teaching geometry with problems: Negotiating instructional situations and mathematical tasks. *Journal for Research in Mathematics Education, 37*, 313–347.

Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics, 17*(1), 7–16.

Hoyles, C. & Küchemann, D. (2002). Students' understandings of logical implication. *Educational Studies in Mathematics, 51*, 193–223.

Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33*, 379–405.

Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education, 5*, 61–88.

Knuth, E. J., Choppin, J. M., & Bieda, K. N. (2009). Middle school students' production of mathematical justifications. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). Mahwah, NJ: Erlbaum.

Knuth, E. J., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Weigel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the twenty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Vol. 4* (pp. 1693–1700). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (1998/2002/2005). *Connected mathematics project*. Upper Saddle River, NJ: Prentice Hall.

Lobato, J., Clarke, D., & Ellis, A. B. (2005). Initiating and eliciting in teaching: A reformulation of telling. *Journal for Research in Mathematics Education, 36*, 101–136.

Mariotti, M. A. (2006). Proof and proving in mathematics education. In Á. Gutiérrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 173–204). Rotterdam, the Netherlands: Sense Publishers.

Martin, T. S., McCrone, S. M. S., Bower, M. L. W., & Dindyal, J. (2005). The interplay of teacher and student actions in the teaching and learning of geometric proof. *Educational Studies in Mathematics, 60*, 95–124.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2009). *Focus in high school mathematics: Reasoning and sense making*. Reston, VA: Author.

Smith, J. P., III. (1996). Efficacy and teaching mathematics by telling: A challenge for reform. *Journal for Research in Mathematics Education, 27*, 387–402.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*, 289–321.

Stylianides, G. J. (2007). Investigating the guidance offered to teachers in curriculum materials: The case of proof in mathematics. *International Journal of Science and Mathematics Education, 5*, 191–215. DOI: 10.1007/s10763-007-9074-y

Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*, 258–288.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics, 48*, 101–119.

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to* Principles and standards for school mathematics (pp. 227–236). Reston, VA: National Council of Teachers of Mathematics.

Yin, R. K. (2006). Case study methods. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 111–122). Mahwah, NJ: Erlbaum.

Author

Kristen N. Bieda, Michigan State University, Department of Teacher Education, 312 Erickson Hall, East Lansing, MI 48824; kbieda@msu.edu

Accepted February 8, 2010