A Longitudinal/Cross-Sectional Study of the Impact of *Mathematics in Context* on Student Mathematical Performance

Classroom Observation Scale

(Working Paper #6)

Mary C. Shafer, Lesley R. Wagner, and Jon Davis

Wisconsin Center for Education Research
University of Wisconsin–Madison

Shafer, M. C., Wagner, L. R., & Davis, J. (1997). *Classroom observation scale* (*Mathematics in Context* Longitudinal/Cross-Sectional Study Working Paper No. 6). Madison, WI: University of Wisconsin.

The development of this instrument was supported in part by the National Science Foundation #REC-9553889.

Description of the Observation Scale

The observation instrument for the longitudinal/cross-sectional study was designed to measure one independent variable (curricular content and materials—the actual curriculum) and the three intervening variables: pedagogical decisions, classroom events, and student pursuits. The observation instrument is composed of seven sections. In the first section, the observer recorded pertinent information related to the teacher and students: the name of the teacher, the school, and the grade level of the students in the class. The observer also recorded information pertinent to the particular lesson: the date of the observation, times the lesson began and ended, text used, unit/chapter taught, and the page numbers taught during the lesson. In the second section of the observation instrument, the observer conducted and recorded notes from a brief preobservation interview of the teacher during which the teacher was asked to identify the mathematical content to be explored or conveyed in the lesson and the location of the lesson with respect to the development of concepts in the instructional unit/chapter. In the third section, the observer recorded the flow of the lesson, which was a list of lesson activities along with the time allotted to each.

The next two sections of the observation instrument were collectively composed of 12 indices for various dimensions of instruction, which addressed the three intervening variables in the research model for the study. Nine of these indices focused on classroom events; the remaining three indices focused on student pursuits. Pedagogical decisions, although not presented in a separate section of the observation instrument, were central to both classroom events and student pursuits.

The indices used to characterize each dimension were based on levels of authentic instruction, tasks, and assessment (Newmann, Secada, & Wehlage, 1995), Cognitively Guided Instruction (Fennema, Carpenter, Franke, Levi, Jacobs, & Empson, 1996), instruction that included teachers' understanding and beliefs about constructivist epistemology (Schifter & Fosnot, 1993), and utilization of particular instructional innovations (Hall, Loucks, Rutherford, Newlove, 1975, quoted in Schifter & Fosnot, 1993). Several levels for each index were preliminarily defined by describing each aspect of instruction and identifying differences between conventional approaches to teaching learning mathematics and approaches that were aligned with the NCTM *Standards* (1989, 1991, 1995), authentic instruction (Newmann, Secada, & Wehlage, 1995), and teaching mathematics for understanding (Carpenter & Lehrer, 1999). Further distinctions in the levels were identified through a review of literature that was specific to each dimension. The indices were refined as a result of classroom observations of nonstudy teachers who used MiC or conventional curricula during the year prior to the study. Three to four levels were identified for each dimension in order to identify differences in these variables between conventional and reform-based approaches to teaching and learning mathematics. The levels in each index are positioned along a continuum from the least appearance of a given characteristic to the most sophisticated implementation of the dimension being scaled. For example, levels of lessons that fostered conceptual understanding range from no attention to conceptual understanding during instruction to lessons in which the continual focus was on building connections among mathematical ideas.

The observation instrument was pilot-tested by project staff numerous times in both MiC and conventional classrooms in order to define and clarify descriptors for each item and to determine ways to achieve interrater reliability. Before the observation instrument was used in study classrooms, administrators, on-site observers, and curriculum specialists from anticipated research sites used and commented on a draft of the observation instrument in classrooms implementing MiC. As a result, project staff refined descriptions and clarified ratings for the final version of the index for each dimension. In order to maintain interrater reliability between the observers in each district and consistency in rating across all three years of the longitudinal study, these indices were not changed after review of data from study teachers.

In the sixth section of the observation instruction, the observer conducted and recorded notes from a brief postobservation interview of the teacher during which the teacher was asked to rate and comment on the degree to which the teacher felt the lesson achieved the purpose noted in the preobservation interview. The teacher was also asked whether any incidents occurred during the lesson which revealed student misunderstanding or provided opportunities to facilitated student understanding in any way. In this way, teachers had an opportunity to describe and explain modifications made during the lesson. In the final section of the observation instrument, the observer recorded any additional comments about the lesson.

Indices

Classroom Events

The lesson provided opportunities for students to make conjectures about mathematical ideas. In the conceptualization of conjectures in the longitudinal study, three types of student conjectures are described and sought in classroom interaction. First, students can make conjectures that are realizations of the connections between existing knowledge and the application of these concepts in new contexts. That is, students might see a connection between a new problems and problems they have already solved. Second, students may investigate the truthfulness of particular statements. Third, conjecturing may permeate a lesson. Given a pattern, for example, students are asked to devise a formula that captures the essence of the pattern in a concise form, which in turn leads to generalizations. Each type of conjecture is given a specific rating in the index, with an observation of the third type given the highest rating.

The lesson fostered the development of conceptual understanding. Instruction that fosters the development of conceptual understanding engages students in creating meaning for the symbols and procedures they use. Problems or questions posed by the teacher or in text materials may direct students' attention to linking procedural and conceptual knowledge. Lower ratings in this category describe classrooms in which teaching for conceptual understanding occurs, but is often overshadowed by an emphasis on procedural knowledge. The highest rating describes a lesson in which links between conceptual and procedural understanding are the main emphasis of the instruction.

Connections within mathematics were explored. In this index, mathematical topics can be thought of in two different ways. First, topics can be broad areas of mathematics such as probability, area, and ratios which connections can be made between factoring, completing the square, or using the quadratic formula. Even though these problems connect mathematical topics, instruction may not focus on discussing or developing these connections. The rating is meant to reflect both the problems and instruction.

Connections between mathematics and students' daily lives were apparent in the lesson. This index measures whether connections between mathematics and students' daily lives were apparent in text problems or problems presented in class or were discussed by the teacher or students.

Students explained their responses or solution strategies. This index is intended to measure the extent to which students elaborate on their solutions orally or in written form by justifying their approach to a problem, explaining their thinking, or supporting their results, rather than simply stating answers.

Multiple strategies were encouraged and valued. This index measures the extent to which students were asked to consider different perspectives in approaching the solution to a problem. Higher ratings on this index refer to lessons in which discussion of alternative strategies is a frequent, important element of classroom instruction.

The teacher valued students' statements about mathematics and used them to build discussion or work toward shared understanding for the class. This index is intended to measure the ways in which the teacher uses student responses during instruction. The highest rating is reserved for lessons in which the teacher not only probed individual students' thinking but also encouraged other students to comment on the solution strategies or used students' thinking processes to open discussions that encourage deeper understanding of mathematics.

The teacher used student inquiries as a guide for instructional mathematics investigation or as a guide to shape the mathematical content of the lesson. Occasionally a student's inquiry can be used to introduce the topic of the lesson, supplement a lesson, or connect the lesson to students' lives. In other cases, a student's question or response may provide a starting point for a rich mathematical journey. This index measures the teacher's responsiveness to student inquiries and the teacher's flexibility in using these inquiries in ways that enhance the lesson.

The teacher encouraged students to reflect on the reasonableness of their responses. This index is intended to measure whether the teacher encouraged students to reflect on the reasonableness of their answers and whether the discussion involved emphasis on conceptual understanding.

Student Pursuits

Student exchanges with peers reflected substantive conversation of mathematical ideas. Substantive conversation by students is characterized by interaction that is reciprocal, which involves listening carefully to others' ideas in order to understand them, building conversation on others' ideas, or extending an idea to a new level. Substantive conversation also promotes shared understanding of mathematical ideas and the use of higher order thinking, such as applying ideas, making comparisons, or raising questions. (Newmann, Secada, and Wehlage, 1995). While other items in this observation scale refer to the role of the teacher in mediating discourse, this item measures student discourse between peers in either large-group or small-group settings.

Interactions among students reflected collaborative working relationships. A low rating is given when students are physically sitting in groups but rarely working together. In contrast, the highest rating denotes a lesson in which students are actively involved in solving problems with their classmates and in which students made sure that all students in the group understood one problem before moving on to the next. N/A is reserved for lessons in which the goal is for students to work on problems independently.

The overall level of student engagement throughout the lesson was serious. This index measures the extent to which students remained on task during the lesson. Engagement is exemplified by behaviors in which students are attentive, complete assigned work, participate by raising questions, contribute to both large-group and small-group discussions, and help their peers (Secada and Byrd, 1993).

Observations

The observers (one each from Districts 1 and 2) were retired teachers with many years of experience teaching mathematics and were selected with district input. Throughout the class period, the observer continually judged the levels of each dimension of classroom events and student pursuits. During each observation the observer took field notes that pertained to the 12 indices. Immediately after observing a lesson, the observer rated each item and recorded evidence from the lesson (consisting of dialogue or an artifact) to support the given rating. In general, a rating of 1 on a particular item indicated that the dimension was rarely or never seen in the lesson; the highest rating indicated that the dimension received major emphasis in the classroom. In practice, high ratings were rarely attained on every item during one observation. Ratings also varied in different observations of the same teacher.

The number of observations per teacher varied in each district (see Table 1). Most teachers in District 1 were observed once a month for a total of nine observations per teacher. During the first year of data collection, one teacher in District 1 accepted an administrative position in December; consequently, she was observed three times, and the newly assigned teacher was observed five times. During the second and third years of data collection, one eighth-grade control class had three teachers, and two seventh-grade experimental classes had two teachers over the course of the school year. As a result, each teacher was observed only a few times. Teachers in District 2 were observed a total of two to nine times each. Fewer observations were conducted in District 2 due to differences in school schedules, procedures for assigning students to classes, and preparation for district and state standardized testing. In addition, four teachers from one school in District 2 withdrew from participation in the study during the spring semester of the first year of data collection; consequently, they were observed only three times. During the third year of data collection, two seventh-grade experimental classes were observed twice because the teacher had been on parental leave.

Table 1
Number of Observations Conducted, by Grade and Year

Grade *(No. of Teachers*)*	Number of Observations Per Teacher	Percent of Teachers Observed			
		1-3 Times	4-6 Times	7-8 Times	9 Times
1997-1998					
5 *(13)*	5-9	0	38	0	62
6 *(12)*	3-9	25	33	8	33
7 *(10)*	3-9	20	40	0	40
1998-1999					
6 *(12)*	5-9	0	33	25	42
7 *(12)*	5-9	0	42	25	33
8 *(10)*	2-9	20	40	30	10
1999-2000					
7 *(9)*	2-9	22	11	11	56
8 *(9)*	8-9	0	0	44	56

*Includes teachers who taught portions of the school year

Interrater Reliability

In the August prior to the study, each observer viewed two videotaped lessons with a graduate project assistant who developed the observation instrument and rated the lessons using the instrument. During these meetings discussions centered on consistency of ratings and descriptions of the types of conjectures observed, the nature of student–student conversation, and instances in which teachers used student inquiries to shape the lesson.

In the fall of 1997, each observer and a project assistant visited five classes in District 1 and nine classes in District 2. During the first few observations at each site, the project assistant's and observer's ratings of several items differed by one point. By the last observation, however, this disagreement had subsided considerably. The first dimension, student conjectures, initially caused difficulty for both observers. For example, observers initially categorized the repetitious practice of problems using a single prescribed algorithm as a first level conjecture, when the first level of conjecture is meant to describe the preponderance of students making conjectures that link concepts they have studied in the past with the same concept set within a new context. Another dimension, students' level of collaboration in the classroom, one observer tended to give the highest rating if students were physically sitting in groups. The project assistant emphasized the importance of circulating around the classroom to determine if students actually worked in groups to support each other's learning. After this training, the observers began observing each study teacher once a month and completed a report for each observation. Completed reports were sent electronically to the research center for analysis. Each observer was compensated an amount per observation as part of a subcontract between the observer and the University of Wisconsin. The amount of payment varied according to the length of the class period observed. In September 1998, both observers worked on interrater reliability with a project assistant during on-site classroom observations in District 1. Because of the lack of funds, on-site work for interrater reliability between observers and a project assistant were not conducted in the fall of 1999.

References

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Portsmouth, NH: Heinemann.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Classrooms that promote mathematical understanding* (pp. 19–32). Mahwah, NJ: Erlbaum.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education 27*(4), 403–434.

Hall, G. E., Loucks, S. F., Rutherford, W. L., & Newlove, B. W. (1975). *Levels of use of the innovation: A framework for analyzing innovation adoption*. Quoted in D. Schifter & C. T. Fosnot (1993), *Reconstructing mathematics education: Stories of teachers meeting the challenge of reform* (pp. 185–189). New York: Teachers College Press.

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

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

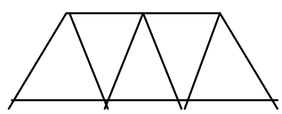
Newmann, F. M., Secada, W. G., & Wehlage, G. G. (1995). *A guide to authentic instruction and assessment: Vision, standards, and scoring*. Madison, WI: Wisconsin Center for Education Research.

Secada, W., & Byrd, L. (1993). *Classroom observation scales: School level reform in the teaching of mathematics* (pp. 14-15). Madison, WI: National Center for Research in Mathematical Sciences Education.

Observation Scale Descriptors

C.1. The lesson provided opportunities for students to make conjectures about mathematical ideas.

This scale measures the extent to which the lesson provided opportunities for students to make conjectures about mathematical ideas. There are three types of conjectures that students might make. One type of conjecture involves the student in making a guess about how to solve a particular problem based on experience solving problems with similar solution strategies. For example, students were solving problems in which they used properties of similar triangles. When asked to determine the height of a tree, students conjectured that an appropriate solution strategy would involve similar triangles. The students made a connection between the new problem and problems that they had previously solved. A second type of conjecture occurs when a student makes a guess about the truthfulness of a particular statement and subsequently plans and conducts an investigation to determine whether the statement is true or false. For example, a 12-year-old student disagreed with a statement that she was half as tall as she is now when she was 6-years old, and proceeded to support her argument by comparing her present height with heights of 6-year-old children. A third type of conjecture is a generalization. A generalization is created by reasoning from specific cases of a particular event, is tested in specific cases, and is logically reasoned to be acceptable for all cases of the event. For example, given that a beam is constructed of rods in the following configuration,



students are asked to describe the relation between the number of rods and the length of the beam[1] (Wijers, Roodhardt, van Reeuwijk, Burrill, Cole, & Pligge, 1998). Using a table to organize their reasoning, students described the pattern that emerged, explained how the pattern fit the given diagram, and generated formulas for the relationship. In this situation, students reasoned from specific cases, tested and supported their ideas with evidence from drawings and the table, and described the relation in a formula.

1. No conjectures of any type were observed in the lesson. Students were not encouraged to make connections between a new problem and problems previously seen, investigate the validity of their own guesses, look for patterns, or make generalizations.

2. Observed conjectures consisted mainly of making connections between a new problem and problems previously seen.

3. Observed conjectures consisted mainly of student investigations about the truthfulness of particular statements.

4. Students made generalizations about mathematical ideas.

1 The length of the beam is the number of rods on the bottom of the beam.

C.2. The lesson fostered the development of conceptual understanding.

Conceptual knowledge is described as the "facts and properties of mathematics that are recognized as being related in some way" (Hiebert & Wearne, 1986, p. 200), or as a network of relationships that link pieces of knowledge (Hiebert & Lefevre, 1986). In the primary grades, for example, students learn the labels for whole-number place-value positions. If this information is stored as isolated pieces of information, the knowledge is not conceptual. If this knowledge, however, is linked with other information about numbers, such as grouping objects into sets of ten or counting by tens or hundreds, then the information becomes conceptual knowledge. The network of relationships about place value grows as other pieces of knowledge related to place value, such as regrouping in subtraction, are recognized. Procedural knowledge, in contrast, is described as having two parts. One category comprises the written mathematical symbols, which are devoid of meaning and are acted upon through knowledge of the syntax of the system. A second category is composed of rules and algorithms for solving mathematics problems, step-by-step procedures that progress from problem statement to solution in a predetermined order. Procedural knowledge is rich in rules and strategies for solving problems, but it is not rich in relationships (Hiebert & Wearne, 1986).

Instruction that fosters the development of conceptual understanding engages students in creating meaning for the symbols and procedures they use. Problems or questions posed by the teacher or in text materials may direct students' attention to linking procedural and conceptual knowledge. In addition and subtraction of decimals, for example, lining up the decimal points should be linked with combining like quantities. Instruction might explicitly bring out the relationships between lining up the decimal point in addition and subtraction and lining up whole numbers on the right side for the same operations (Hiebert & Wearne, 1986).

1. The lesson as presented did not promote conceptual understanding.

2. The lesson asked few questions that fostered students' conceptual development of mathematical ideas, or conceptual understanding was a small part of lesson design.

3. Some lesson questions fostered students' conceptual development of mathematical ideas, or some aspects of the lesson focused on conceptual understanding, but the main focus of the lesson was on building students' procedural understanding without meaning.

4. The continual focus of the lesson was on building connections between disparate pieces of information or linking procedural knowledge with conceptual knowledge.

C.3. Connections within mathematics were explored in the lesson.

This scale measures the extent to which instruction addressed mathematical topics thoroughly enough to explore relationships and connections among them.[2] A low rating is given when the mathematical topic of the lesson was covered in ways that gave students only a surface treatment of its meaning, and instruction treated this topic in isolation of other mathematical topics. A high rating is given when the mathematical topic of the lesson was explored in enough detail for students to think about relationships and connections among mathematical topics. Rather than examining fragmented pieces of information, students looked for and discussed relationships among mathematical ideas, expressed understanding of mathematical topics, or provided explanations of their solution strategies for relatively complex problems in which two or more mathematical ideas were integrated.

Topics can be thought of in two different ways. First, topics can be broad areas of mathematics such as probability, area, and ratios, as in the following problem. Students are asked to determine the probability of a frog jumping from a cage and landing on white or black floor tiles and to express this probability as a fraction or percent (Jonker, van Galen, Boswinkel, Wijers, Simon, Burrill, & Middleton, 1998). In solving this problem, students use area, number, and probability concepts. Second, connections can be made among more narrowly defined areas such as a lesson involving the solution of quadratic equations. In this lesson, connections can be made between factoring, completing the square, or using the quadratic formula. Even though these problems connect mathematical topics, instruction may not focus on discussing or developing these connections. The rating should reflect both the problems and instruction.

1. The mathematical topic of the lesson was covered in ways that gave students only a surface treatment of its meaning. The mathematical topic was presented in isolation of other topics, and the teacher and students did not talk about connections between the topic of the lesson and other mathematical topics.

2. Connections among mathematical topics were present in the lesson. The teacher or students might have briefly mentioned that the topic was related to others, but these connections were not discussed in detail by the teacher or the students.

3. Connections among mathematical topics were discussed by teacher and students during the lesson, or connections were clearly explained by the teacher.

4. The mathematical topic of the lesson was explored in enough detail for students to think about relationships and connections among mathematical topics. During instruction, many students did at least one of the following: looked for and discussed relationships among mathematical ideas, expressed understanding of mathematical relationships, or provided explanations of their solution strategies for relatively complex problems in which two or more mathematical ideas were integrated.

---

2 Ideas were drawn from Newmann, Secada, & Wehlage (1995), Chapter 3, Authentic Instruction, Deep Knowledge (pp. 31-35).

C.4. Connections between mathematics and students' daily lives were apparent in the lesson.

This scale measures whether connections between mathematics and students' daily lives were apparent in text problems or discussed by the teacher or students. Examples of problems that foster such connections are estimating the sale price of an item or determining the amount of ingredients required to serve four people when a recipe serves seven. In contrast, word problems such as "Bart is two years older than Lisa. In five years Bart will be twice as old as Lisa. How old are they now?" are devoid of connections between mathematics and students' lives.

1. Connections between mathematics and students' daily lives were not apparent in the lesson.

2. Connections between mathematics and students' daily lives were not apparent to the students, but would be reasonably clear if explained by the teacher.

3. Connections between mathematics and students' daily lives were clearly apparent in the lesson.

C.5. Students explained their responses or solution strategies.

This scale is intended to measure the extent to which students elaborate on their solutions orally or in written form by justifying their approach to a problem, explaining their thinking, or supporting their results, rather than simply stating answers.

1. Students simply stated answers to problems. They did not explain their responses or solution strategies orally or in written form.

2. Students explained how they arrived at an answer, but these explanations focused on the execution of procedures for solving problems rather than an elaboration on their thinking and solution path.

3. Students explained their responses or solution strategies. They elaborated on their solutions orally or in written form by justifying their approach to a problem, explaining their thinking, or supporting their results.

C.6. Multiple strategies were encouraged and valued.

This scale measures the extent to which students were asked to consider different perspectives in approaching the solution to a problem. In a classroom where multiple strategies are encouraged and valued, students spend much of their time discussing different strategies in a substantive manner, and this discourse is an important element within the classroom. Multiple strategies might be elicited by the teacher during whole-class or small-group discussion in which students explicitly share their strategies. The task itself might clearly involve students in solving the problem in different ways (e.g., find the discount in another way), or the task may require students to consider alternative approaches for successful completion (e.g., list as many ways as you can to calculate 15 x $1.98).

1. Multiple strategies were not elicited from students.

2. Different problem-solving strategies were rarely elicited from students or only briefly mentioned by the teacher.

3. Students were asked if alternate strategies were used in solving particular problems, but this was not a primary goal of instruction.

4. Discussion of alternative strategies was frequent, substantive in nature, and an important element of classroom instruction.

C.7. The teacher valued students' statements about mathematics and used them to build discussion or work toward shared understanding for the class.

This scale is intended to measure the ways in which the teacher uses student responses during instruction. Teachers can give credence to students' responses by inviting students to listen carefully to other students, to ask each other questions that clarify meaning, and to compare other students' strategies with their own. Teachers can also use student responses to pose questions that stimulate further discussion, to illustrate a point, or to relate them to other aspects of the lesson.

1. The teacher was interested only in correct answers. The majority of the teacher's remarks about student responses were neutral short comments such as "Okay," "All right," or "Fine." No attempt was made to use students' responses to further discussion.

2. The teacher established a dialogue with the student by asking probing questions in an attempt to elicit a student's thinking processes or solution strategies.

3. The teacher valued students' statements about mathematics by using them to foment discussion or to relate them to the lesson in some way. The teacher opened up discussion about the student response by asking other students questions such as: "Does everyone agree with this?" or "Would anyone like to comment on this response?"

C.8. The teacher used student inquiries as a guide for instructional mathematics investigations or as a guide to shape the mathematical content of the lesson.

Occasionally a student's inquiry can be used to introduce the topic of the lesson, supplement a lesson, or connect the lesson to students' lives. In other cases, a student's question or response may provide a starting point for a rich mathematical journey. A student's question about whether the sum of the angles of every triangle is always 180°, for example, might lead to a discussion of non-Euclidean geometry. This scale measures the teacher's responsiveness to student inquiries and the teacher's flexibility in using these inquiries in ways that enhance the lesson.

Circle Yes, if the teacher used students' inquiries as a guide for instructional mathematics investigations or as a guide to shape the mathematical content of the lesson.

Circle No, if a student's comment or question potentially could have led to such a discussion, but the teacher did not pursue it.

Circle N/A, if no such opportunities came about during the lesson.

C.9. The teacher encouraged students to reflect on the reasonableness of their responses.

An unreasonable response refers to a response that is mathematically distant from the correct answer and might even be distant from an answer that students recognize as reasonable in contexts outside the classroom. One explanation for unreasonable responses is that students do not check the reasonableness of their answers. Although this may be true in some cases, unreasonable responses may also be the result of the lack of connections between symbols and their meaning. Evaluating the reasonableness of a solution involves connections between conceptual and procedural knowledge. These connections are especially significant at the end of the problem-solving process. Lining up decimal points when adding or subtracting decimals, for example, without connecting the process to place value concepts, may lead to unreasonable responses. Students might rely on rules or procedures to obtain correct answers and not have the conceptual knowledge to help them evaluate reasonableness of the answer (Hiebert & Wearne, 1986). This scale is intended to measure whether the teacher encouraged students to reflect on the reasonableness of their answers and whether the discussion involved emphasis on conceptual understanding.

1. The teacher rarely asked students whether their answers were reasonable. If a student gave an incorrect response, another student provided or was asked to provide a correct answer.

2. The teacher asked students if they checked whether their answers were reasonable but did not promote discussion that emphasized conceptual understanding.

3. The teacher encouraged students to reflect on the reasonableness of their answers, and the discussion involved emphasis on conceptual understanding.

D.1. Student exchanges with peers reflected substantive conversation of mathematical ideas.

With this scale we are attempting to capture the quality of student communication. Substantive conversation by students is characterized by interaction that is reciprocal, involving listening carefully to others' ideas in order to understand them, building conversation on them, or extending the idea to a new level. Substantive conversation also promotes shared understanding of mathematical ideas and the use of higher order thinking, such as applying ideas, making comparisons, or raising questions.[3] In contrast, student exchanges with little or no substantive conversation involve reporting facts or procedures in ways that do not encourage further discussion of ideas.

1. There were no exchanges between peers in small groups or as a formal part of the general discourse within a large-group setting.

2. Student exchanges with peers reflected little or no substantive conversation of mathematical ideas.

3. Most students only asked one another for a clarification of directions given by the teacher or simply accepted someone's answer without an explanation of how it was found. Few students asked how a solution was found or asked for a clarification of another student's answer.

4. Most of the students asked their classmates for a description of how they solved a particular problem, discussed alternative strategies, and/or questioned how classmates arrived at a solution.

3 Ideas were drawn from Newmann, Secada, & Wehlage (1995), Chapter 3, Authentic Instruction, Substantive Conversation (pp. 35-40).

D.2. Interactions among students reflected collaborative working relationships.

The collaborative nature of the classroom can be thought of as students working together, exchanging ideas, and finding solutions to the same problem. This includes providing assistance to one another, making sure that everyone understands and is working on the same problem, exchanging ideas, and seeking help from each other when it is needed. Student collaboration can occur in a small-group or large-group setting. If the major focus of the lesson is on providing students with individual work, then N/A should be selected.

N/A. The main purpose of the lesson was to give students needed individual practice, or students spent nearly all of the class period involved in independent work.

1. None of the students were working together in small groups or in a large-group setting. If students were working in small groups, then one student typically gave answers to other members of group without explanation of why certain procedures were used.

2. Few students were sharing ideas or discussing how a problem should be solved in small groups or in a large-group setting. Although students physically sat together, there was little exchange of ideas or assistance. Many of the students in a group were working on different problems and at different paces.

3. Some students were exchanging ideas, or providing assistance to their classmates; however, a few students relied on other members of the group to solve problems. Contributions to solving problems were not made equally by all students.

4. Most students were involved with their classmates in solving problems and made sure that other group members were caught up and understood the problems before moving on to the next problem.

D.3. The overall level of student engagement throughout the lesson was serious.[4]

This scale measures the extent to which students remained on task during the lesson.

1. Disruptive disengagement. Students were frequently off task, as evidenced by gross inattention or serious disruptions by many. This was the central characteristic during much of the class.

2. Passive disengagement. Students appeared lethargic and were only occasionally on task carrying out assigned activities. For substantial portions of time, many students were either clearly off task or nominally on task but not trying very hard.

3. Sporadic or episodic engagement. Most students, some of the time, were engaged in class activities, but this engagement was inconsistent, mildly enthusiastic, or dependent on frequent prodding from the teacher.

4. Widespread engagement. Most students, most of the time, were on task pursuing the substance of the lesson. Most students seemed to take the work seriously and were trying hard.

4 Ideas were drawn from Secada & Byrd (pp. 14-15).

References

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. Mahwah, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1986). Procedures over concepts: The acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics*. Mahwah, NJ: Erlbaum.

Jonker, V., can Galen, F., Boswinkel, N., Wijers, M., Simon, A. N., Burrill, G., & Middleton, J. A. (1998). Take a chance. In National Center for Research in Mathematical Sciences Education & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago: Encyclopaedia Britannica Educational Corporation.

Newmann, F., Secada, W., & Wehlage, G. (1995). *A guide to authentic instruction and assessment: Vision, standards, and scoring*. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin–Madison.

Secada, W., & Byrd, L. (1993). *Classroom observation scales: School level reform in the teaching of mathematics* (pp. 14-15). Madison, WI: National Center for Research in Mathematical Sciences Education.

Wijers, M., Roodhardt, M., can Reeuwijk, M., Burrill, G., Cole, B. R., & Pligge, M. A. (1998). Building Formulas. In National Center for Research in Mathematical Sciences Education & Freudenthal Institute (Eds.), *Mathematics in context*. Chicago: Encyclopaedia Britannica Educational Corporation.

Observation Scale

Observer: ________________________________ Time Lesson Begins: ______________________
Teacher: ________________________________ Time Lesson Ends: ______________________
School: ________________________________ Textbook: ________________________________
Grade: ________ Chapter/Unit: ____________________________
Date of Observation: ______________________ Lesson (pages): _________________________

A. Pre-observation Interview With Teacher

1. What mathematical concept(s) or important ideas are being conveyed in this lesson?
   ____
   ____

2. Where is this activity generally situated in the development of a unit? (For example, day 1 (introduction) of 5 days needed to complete the unit)
   ____

B. Lesson Flow

Describe the main activities that occurred during the class period and the amount of time devoted to each activity. For example: warm-up—5 minutes, introduction of concept through context—7 minutes, large group discussion—10 minutes, group activity—25 minutes, summary by teacher—5 minutes.

For sections C and D please refer to the observation scale descriptors on the attached sheets. Please provide evidence supporting your rating.

C. Classroom Events

		Evidence
1. The lesson provided opportunities for students to make conjectures about mathematical ideas.	1 2 3 4	
2. The lesson fostered the development of conceptual understanding.	1 2 3 4	
3. Connections within mathematics were explored in the lesson.	1 2 3 4	
4. Connections between mathematics and students' daily lives were apparent in the lesson.	1 2 3	
5. Students explained their responses or solution strategies.	1 2 3	
6. Multiple strategies were encouraged and valued.	1 2 3 4	
7. The teacher valued students' statements about mathematics and used them to build discussion or work toward shared understanding for the class.	1 2 3	
8. The teacher used student inquiries as a guide for instructional mathematics investigations or as a guide to shape the mathematical content of the lesson.	Yes No N/A	
9. The teacher encouraged students to reflect on the reasonableness of their responses.	1 2 3	

D. Pupil Pursuits

Evidence

1. Student exchanges with peers reflected substantive conversation of mathematical ideas. 1 2 3 4

2. Interactions among students reflected collaborative working relationships. 1 2 3 4 N/A

3. The overall level of student engagement throughout the lesson was serious. 1 2 3 4

E. Postobservation Interview With Teacher

1. On a scale of 1 to 4, how did the teacher rate the lesson as achieving the intended purpose cited in the pre-observation interview?

   Goals Not Met All Goals Met
   1 2 3 4

   Comments: ____

2. Did the teacher feel that any incidents occurred which revealed student misunderstanding or provided opportunities to facilitate student understanding in any way? Yes No If yes, please describe and explain if the lesson was modified in any way.

   ____

F. Additional Information

1. Please feel free to add any comments or information (on the back) that you think would be of interest in describing the classroom that you observed.