Enhancing Mathematical Reasoning in the Classroom: The Effects of Cooperative Learning and Metacognitive Training

Article *in* American Educational Research Journal · March 2003

DOI: 10.3102/00028312040001281

Bracha Kramarski
Bar Ilan University

Enhancing Mathematical Reasoning in the Classroom: The Effects of Cooperative Learning and Metacognitive Training

Bracha Kramarski and Zemira R. Mevarech
Am Educ Res J 2003 40: 281
DOI: 10.3102/00028312040001281

American Educational Research Journal
Spring 2003, Vol. 40, No. 1, pp. 281–310

Enhancing Mathematical Reasoning in the Classroom: The Effects of Cooperative Learning and Metacognitive Training

Bracha Kramarski and Zemira R. Mevarech
Bar-Ilan University

The purpose of this study was to investigate the effects of four instructional methods on students' mathematical reasoning and metacognitive knowledge. The participants were 384 eighth-grade students. The instructional methods were cooperative learning combined with metacognitive training (COOP+META), individualized learning combined with metacognitive training (IND+META), cooperative learning without metacognitive training (COOP), and individualized learning without metacognitive training (IND). Results showed that the COOP+META group significantly outperformed the IND+META group, which in turn significantly outperformed the COOP and IND groups on graph interpretation and various aspects of mathematical explanations. Furthermore, the metacognitive groups (COOP+META and IND+META) outperformed their counterparts (COOP and IND) on graph construction (transfer tasks) and metacognitive knowledge. This article presents theoretical and practical implications of the findings.

KEYWORDS: argumentation, graphs, mathematical reasoning, metacognition, transfer task.

Most research on mathematical education emphasizes the importance of mathematical reasoning as an integral part of doing mathematics (National Council of Teachers of Mathematics [NCTM], 1989, 2000; Cai, Lane,

---

BRACHA KRAMARSKI is a Lecturer and Head of the Teacher Training Department, School of Education, Bar-Ilan University, Ramat-Gan 52900, Israel; e-mail *kramab@mail.biu.ac.il*. Her areas of special interest are mathematical education, metacognitive instruction, advanced educational technologies, and teachers' professional development.

ZEMIRA R. MEVARECH is the Head of the School of Education, Bar-Ilan University, Ramat-Gan 52900, Israel; e-mail *mevarz@mail.biu.ac.il*. Her areas of special interest are mathematical education, metacognitive instruction, quantitative methods of education research, and advanced educational technologies.

& Jakabcsin, 1996). According to the NCTM standards, mathematical reasoning requires the attainment of abilities to construct mathematical conjectures, develop and evaluate mathematical arguments, and select and use various types of representations. To help students meet the standards, the NCTM emphasizes the importance of mathematical discourse in the classroom. Students not only should discuss their reasoning on a regular basis with the teacher and with one another but also should explain the basis for their mathematical reasoning, both in writing and in their mathematical discourse.

Research on the NCTM standards has moved beyond identifying their elements to focusing on the treatments, or conditions, under which they can be optimally enhanced in the classroom. In the literature two methods have been suggested: One focuses on cooperative learning (e.g., Slavin, 1996), the other on metacognitive training (e.g., Schoenfeld, 1985). The following paragraphs briefly describe the effects of each method on mathematical reasoning.

Cooperative Learning and Mathematical Reasoning

Research in cognitive psychology has shown that learning occurs when the learner is engaged in some sort of cognitive restructuring or elaboration (Wittrock, 1986). One of the most effective means of elaboration is explaining the material to someone else. Slavin proposes that "(a) through mutual feedback and debate, peers motivate one another to abandon misconceptions and search for better solutions; (b) the experience of peer interaction can help a child master social processes, such as participation and argumentation, and cognitive processes, such as verification and criticism; (c) collaboration between peers can provide a forum for discovery learning and can encourage creative thinking; and (d) peer interaction can introduce children to the process of generating ideas" (1996, pp. 49–50).

On the basis of this research, several cooperative learning programs have been designed to replace the traditional teaching of mathematics. The following are examples of programs for teaching mathematics in cooperative settings: the Interactive Mathematics Program, a 4-year program for learning algebra through precalculus (Fendel et al., 1996); the Core-Plus Mathematics Project (Hirsch et al., 1997) for high school students; Connect Mathematics for Grades 6–8 (Lappan et al., 1996); IMPROVE for Grades 7–9 (Mevarech & Kramarski, 1997a); and Number Power for Grades K–6 (Robertson et al., 1996). For an excellent review of cooperative programs for the teaching of mathematics, see Davidson, 1990.

The use of small groups, however, requires fundamental changes not only in the organization of the classroom but also in ways of learning. From a cognitive perspective, students who have poor communication skills are less likely to benefit from cooperative learning because they are not able to communicate their mathematical reasoning to others, nor do they know how to ask questions, reflect on their solution process to explain it to their peers, or offer constructive criticism (Webb, 1989, 1991; Webb & Farivar, 1994).

Cohen (1996), Webb and Farivar (1994), Meloth and Deering (1994), and Mevarech and Kramarski (1997a) suggest that developing mathematical reasoning in small-group activities must be structured to maximize the opportunities for each student to be engaged in questioning, elaboration, explanation, and other verbal communication through which students can express their ideas and group members can give and receive feedback. Researchers (Cohen; Cobb, Boufi, McClain, & Whitenack, 1997) have indicated that features of discourse are new behaviors that students can learn through practice and reinforcement. Using arguments for mathematical reasoning, for example, can become a norm for behavior that enhances mathematical reasoning. Participation in reflective discourse in which mathematical activity is objectified can be an explicit topic of conversation. Thus, students need to be exposed to explicit metacognitive training that focuses on skills for high-level discourse.

Metacognitive Training and Mathematical Reasoning

Several studies (e.g., Mevarech & Kramarski, 1997a; Schoenfeld, 1985; Lester, Garofalo, & Kroll, 1989) have examined the effects of metacognitive training on mathematics reasoning. In all of these studies, the metacognitive training was based on Polya's (1945) approach for solving mathematical problems. A major common element of these programs is training students who work in small groups to formulate and answer a series of self-addressed metacognitive questions that focus on (a) the nature of the problem or task, (b) the construction of relationships between previous and new knowledge, and (c) the use of strategies appropriate for solving the problem or task.

Mevarech and Kramarski (1997a) examined the effects of metacognitive training embedded in small groups. The program is called IMPROVE, the acronym of all the teaching stages: *I*ntroducing the new concepts, *M*etacognitive questioning, *P*racticing, *R*eviewing and reducing difficulties, *O*btaining mastery, *V*erification, and *E*nrichment. Mevarech and Kramarski reported that IMPROVE students who studied in heterogeneous classrooms without tracking or grouping outperformed their counterparts in nontreatment control groups who studied in small groups. In particular, observation showed that IMPROVE had positive effects on students' mathematical achievement and ability to explain their reasoning. These findings are in line with those of King (1989, 1991, 1994), who showed that students in various age groups (e.g., fourth-graders, sixth-graders, and college students) who were trained to formulate and answer metacognitive questions similar to the ones described earlier outperformed their counterparts in control groups on various measures of achievement and problem solving outside the area of mathematics.

Although these findings are promising, they all focus on combined cooperative learning and metacognitive training. Thus there remains the interesting question, What is the unique contribution by each component to students' mathematical reasoning? For example, one may convincingly argue that students exposed to metacognitive training in individualized settings perform similarly to or even better than students exposed to metacognitive training in cooperative learning settings because the former are trained to analyze the problem or task, to construct connections between new and previous knowledge, and to use strategies that are appropriate for solving the problems or completing the tasks. According to this view, metacognitive training is an effective method for enhancing mathematical reasoning regardless of classroom organization. By comparing the mathematical reasoning of students who learned in cooperative settings (with or without metacognitive training) and that of students who learned in individualized settings (with or without metacognitive training), better evidence can be provided regarding the conditions appropriate for attaining standards in mathematics (NCTM, 1989, 2000). To our knowledge, no previous study has addressed this issue.

It is particularly important to examine the differential effects of the four instructional methods on the completion of transfer tasks. Many studies have indicated that students' ability to transfer their knowledge to new situations is quite limited (e.g., Salomon & Perkins, 1987). One of the factors that contribute to that limitation is students' inability to identify the essential elements that are similar or different in the already-completed and targeted tasks. Cecil and Roazzi (1994) reported that training students to describe the similarities and differences between problems facilitated their ability to succeed at transfer tasks. On the basis of that study, we hypothesized that regardless of classroom organization, students who are trained to formulate and answer metacognitive questions such as those described earlier will be better able to transfer their knowledge to new situations than students who are not exposed to such metacognitive training.

Finally, there is a need to investigate the metacognitive knowledge that students acquire under the various conditions. Since the late 1970s, when Flavell (1979) first coined the term *metacognition,* much research has focused on the nature of metacognition (e.g., Schoenfeld, 1987), how it develops as a function of students' age (e.g., Schneider & Sodian, 1990), and how it can be used in the classroom (e.g., Garofalo & Lester, 1985). In particular, research has focused on various kinds of metacognition, including metamemory, metalanguage, metasolving, and metareading (Forrest-Pressley, Makinnon, & Gary Waller, 1985). These studies led us to suggest a distinction between general and domain-specific metacognitive knowledge, similar to the distinction made between cognitive processes (Salomon & Perkins, 1987). *General metacognitive knowledge* is knowing about and being able to control and regulate problem-solving processes regardless of the specific domain from which problems or tasks are drawn. *Domain-specific metacognitive knowledge* focuses on the unique features of each domain and therefore varies among domains. Schoenfeld indicated that in mathematics, metacognitive knowledge includes knowledge about one's own thought processes (e.g., accuracy in describing your own thinking), control or self-regulation (e.g., keeping track of what you are doing when solving problems or tasks), and beliefs and intuitions (e.g., ideas about mathematics that you bring to your work in mathematics). There is reason to suppose that different instructional methods may have different effects on students' general and specific metacognitive knowledge. Students who are exposed to metacognitive training are expected to be better at reflecting on solution processes (general and specific) than students who are not exposed to such training. Discussing metacognitive issues with others is expected to enhance metacognitive knowledge.

The present study addressed these issues. We compared students' mathematical reasoning, transfer ability, and metacognitive knowledge under four instructional methods: cooperative learning combined with metacognitive training (COOP+META), cooperative learning without metacognitive training (COOP), individualized learning combined with metacognitive training (IND+META), and individualized learning without metacognitive training (IND).

Method

Participants

The participants were 384 students (181 boys and 203 girls) who studied in 12 eighth-grade classrooms randomly selected from 4 junior high schools. The 4 schools were randomly selected from a pool of 15 schools located in one district, where mathematics was taught in heterogeneous classrooms without groupings or ability tracking. The schools were similar in size, average socioeconomic status as defined by the Israel Ministry of Education, students' mean age, and levels of mathematics achievement assessed before the beginning of the study. Within schools, classes were normally distributed in terms of students' ability and prior knowledge. The average students' age was 13.3 years.

Twelve teachers participated in the study, each teaching in one classroom. All teachers were female and had more than 5 years of experience in teaching mathematics, and all had taught in heterogeneous classrooms. All teachers had a degree in mathematics education. The teachers were exposed to a 2-day inservice training program (to be described later).

Conditions

All classes studied a linear graph unit. That unit was selected for three reasons: First, it introduced various kinds of mathematics representations. Second, comprehending graphs was one of the mathematical skills that were emphasized by the NCTM standards (1989, 2000). Finally, comprehending graphs is a central skill in the study of sciences and mathematics.

The main purpose of the unit was to develop students' understanding of linear graphs. In particular, all of the students studied (a) the concepts of slope, intersection point, and rate of change; (b) quantitative and qualitative methods of graph interpretation; and (c) transformation of algebraic expressions of the form $y = mx + b$ into graphic representations.

In all classrooms mathematics was taught five times a week, in accordance with the mathematics curriculum suggested by the Israel Ministry of Education. The linear graph unit was taught for 2 weeks. At the beginning of the unit the teacher introduced strategies of graph interpretation such as using tables, algebraic formulas, steps, and verbal explanations. All students in all conditions practiced those strategies with the same problems or tasks, and all used the same textbook. In addition, the teachers used the same problems or tasks when they introduced new concepts. In addition to the textbook, students under each condition used learning materials that covered the same kinds of exercises as those presented in the textbook but that emphasized the unique components of each condition. An example of an open-ended task and how it related to the instruction is provided in Appendix A.

The differences among the groups were in the instructional method. Because each condition was defined by the presence or absence of two components—metacognitive training and cooperative learning—we first describe the components separately and then describe how they were combined.

Metacognitive Training

The metacognitive training used in the present study was based on the techniques suggested by Schoenfeld (1985) and Mevarech and Kramarski (1997a). The metacognitive training used three sets of self-addressed metacognitive questions: *comprehension questions, strategic questions,* and *connection questions.* The comprehension questions were designed to prompt students to reflect on a problem before solving it. In addressing a comprehension question, students had to read the problem, describe the relevant concepts in their own words, and try to understand what the concepts meant. In graph interpretation, the comprehension questions guided students to interpret problems on both the local-to-global dimension and the quantitative-to-qualitative dimension (Leinhardt, Zaslavsky, & Stein, 1990). The comprehension questions included the following: What does the x-axis represent? What does the y-axis represent? What is the trend of the graph? What are the specific points on the graph? To assist students in remembering the comprehension questions, students used the acronym DATA: *D*escribe the x-axis and the y-axis; *A*ddress the units and the ranges of each axis; *T*ell the *T*rend(s) of the graph or parts of the graph; and *A*nalyze specific points.

The strategic questions were designed to prompt students to consider which strategies were appropriate for solving or completing a given problem or task and for what reasons. In addressing the strategic questions, students had to describe the “what” (e.g., What strategy, tactic, or principle can be used to solve the problem or complete the task?), the “why” (e.g., Why is this strategy, tactic, or principle most appropriate for this problem or task?); and the “how” (e.g., How can the suggested plan be carried out?). Possible strategies for interpreting graphs were adding steps to a graph to calculate the slope, using data tables, and referring to the algebraic representation of the graph.

The connection questions were designed to prompt students to focus on similarities and differences between the immediate problem or task and problems or tasks that they had already completed successfully. Connection questions regarding graphs guided students to find similarities and differences between the graph at hand and graphs they had already interpreted, or to compare different intervals on the same graph.

The metacognitive questions were printed in the students’ working sheets and in the teacher guide. Students used the metacognitive questions in their small groups or individualized activities and in writing when they used their booklets. In addition, the teachers modeled the use of the metacognitive questioning when they introduced new concepts to the whole class, reviewed the materials, and helped students in their small groups or individualized activities. Students were told that asking and answering the metacognitive questions would help them to understand and remember the material presented in mathematics classes.

Cooperative and Individualized Learning

Generally speaking, both cooperative and individualized learning sessions included three parts: teacher introduction to the whole class (about 10 minutes), cooperative or individualized seatwork (about 30 minutes), and teacher review with the whole class (about 5 minutes).

Each session started with a teacher's short presentation of the new materials to the whole class using the question-answering technique described by Marx and Walsh (1988). The teacher began by providing answers to two questions: What am I supposed to do in this task? (What's in the problem or task?) and, What do I already know about it? (What are the differences and similarities between . . . and . . . ?). Next, the teacher modeled strategies for completing the task and explained why the strategies were likely to succeed. Last, the teacher explained how to check the answer and what to do if the plan did not work.

After the introduction, students began work in small groups or individually, using the materials designed by us. The cooperative technique followed the method suggested by Brown and Palincsar (1989): Students learned in heterogeneous teams of four, each team including one high-achieving student, two middle-achieving students, and one low-achieving student. To ensure team heterogeneity, all students were tested on mathematics achievement before the beginning of the study and teams were constructed on the basis of students' performance on that test. The learning in teams was implemented as follows: Each student, in turn, read a problem or task aloud and tried to complete it. Whenever there was no consensus, the team discussed the issue until the disagreements were resolved. In talking about the problem or task, explaining it to each other, approaching it from different perspectives, balancing the perspectives against each other, and proceeding according to what seemed to be the best option at the time, students used the diversity in their prior knowledge to self-regulate their learning. When all team members agreed on a solution, they wrote it down on their answer sheets. When none of the team members knew how to complete a problem or task, they asked for teacher assistance. At the end of the period the teacher reviewed the main ideas of the lesson with the entire class. When common difficulties came to light, the teacher provided additional explanations to the whole class.

Components Combined to Form the Four Conditions

The metacognitive training and the cooperative or individualized components were combined as follows.

COOP+META Condition

The COOP+META condition resembled the IMPROVE procedure (Mevarech & Kramarski, 1997a). Students in that condition studied in small heterogeneous groups using the metacognitive questions described earlier: comprehension questions, strategic questions, and connection questions. The metacognitive questions were used by students individually when their turn arrived to solve a problem or complete a task aloud, by the group as a whole in mathematical discourse, and by the teacher when introducing the new concepts to the whole class, reviewing the lesson at the end of the class, and providing help in the small groups.

IND+META Condition

Under the IND+META condition, the metacognitive training was exactly the same as in the COOP+META condition, except that it was implemented in individualized rather than cooperative settings. As in the COOP+META condition, each period in the IND+META condition started with a 10-minute teacher introduction for the whole class. The teacher introduced the new concepts using the metacognitive question-answering technique and modeling the use of the questions. The students then began to work individually on problems or tasks by using the same metacognitive questions described earlier, and the metacognitive questions were printed in the students' booklets. Students had to answer the questions in writing. Finally, at the end of the period, the teacher reviewed the new concepts by using the question-answering technique and modeling the use of the metacognitive questions. When students worked on the problems individually, the teacher provided help to individual students as needed. When helping students, the teacher also modeled the use of the metacognitive questions.

COOP Condition

Under the COOP condition, students studied in small heterogeneous groups as in the COOP+META condition, but they were not exposed to metacognitive training. Each session started with a short presentation of the new materials to the whole class using the question-answering technique. The students then started work on the problems or tasks. The teachers encouraged students to discuss mathematical ideas in their small groups and to provide explanations to one another, but no explicit metacognitive guidance was provided. Each student, in turn, read a problem or task aloud and tried to solve or complete it. When students failed to solve a problem or did not agree on the solution, the team discussed the components of the problem until consensus was achieved. When all team members agreed on a solution, they recorded it in their notebooks. When no team members knew how to solve a problem, the team asked for teacher assistance. At the end of the lesson, the teacher reviewed the main ideas of the lesson with the entire class. When common difficulties were observed, the teacher provided additional explanations to the whole class. In each period the teacher worked with two to three groups.

IND Condition

Under the IND condition, students learned individually and without metacognitive training. Each class started with a 10-minute teacher introduction of the new concepts to the whole class. Students then worked on the problems or tasks individually without using the metacognitive questions. While the students worked, the teacher provided help to those who needed it. At the end of the class, the teacher reviewed the new concepts with the whole class. That group served as a control group.

Learning Materials and Teacher Training

We designed two sets of learning materials for the purposes of the present study. One set included the metacognitive questions and the other did not. Instead of metacognitive questions, the learning materials of the non-metacognitive groups (COOP and IND) included general instruction (e.g., Explain your answer in writing; or, Discuss your mathematical ideas with your classmates). Otherwise, the two sets of learning materials were identical.

Before the beginning of the study, all 12 teachers were exposed to a 2-day inservice training, which focused on pedagogical issues related to the teaching of the linear graph unit. Teachers were told that they were participating in an experiment in which new learning materials would be tried out.

Teachers in each learning condition were exposed separately to the theoretical background of their learning method and its practical implications. The teaching guides included explicit lesson plans, learning materials, and use of examples. One set of learning materials included the metacognitive questions; the other, general instructions (e.g., Discuss your mathematical ideas with your classmates; or, Explain your answers). The mathematical problems that were used as examples were identical in all conditions. The teachers who were assigned to the metacognitive groups (COOP+META and IND+META) were introduced to the rationale and techniques of the metacognitive method. The COOP and the COOP+META teachers were introduced to the rationale of cooperative learning and how to implement it. The IND and IND+META teachers were introduced to the rationale of individualized learning and how to implement it. Otherwise, the training of all teachers was identical.

Measurements

We used three measures to assess students' mathematical reasoning and metacognitive knowledge: a graph interpretation test, a graph construction test, and a metacognitive questionnaire.

Graph Interpretation Test

A 36-item test, adapted from the studies of Mevarech and Kramarski (1992, 1993), assessed students' ability to interpret graphs, particularly linear graphs. The test included items that required qualitative and quantitative graph interpretation skills. According to Leinhardt et al. (1990), qualitative interpretation of a graph requires "looking at the entire graph (or part of it) and gaining meaning about the relationship between the two variables and, in particular, their pattern of covariation" (p. 11). Quantitative interpretation is often associated with local and specific features, based on point identification and a lower level of graph processing (Wainer, 1992).

The test contained two kinds of items, presented in a mixed order. Twenty-seven were based on traditional evaluation procedures; they included multiple-choice and short, open-ended items regarding basic knowledge about the Cartesian system and linear-graph interpretation (e.g., In which months were the temperatures equal?). Nine items did not require (or invite) local interpretations or computations but rather were designed specifically to assess students' mathematical reasoning. These were open-ended items that asked students to give a final answer and explain their reasoning in writing. They required students to (a) draw conclusions and make algebraic generalizations on the basis of a given graph; (b) evaluate graphs representing the same story and decide which graph better represented the story; (c) resolve mathematical conflicts regarding linear graphs; (d) identify misconceptions regarding linear graphs (see, for example, Leinhardt et al., 1990; Mevarech & Kramarski, 1997b); (e) analyze graphs and decide whether certain mathematical expressions always, never, or sometimes represented the given graphs and vice versa; and (f) analyze the structure of graphs. An example of an open-ended task is presented later.

Students received a score of either 1 (correct answer) or zero (incorrect answer) for each item and a total score ranging from zero to 36. The Kuder Richardson reliability coefficient was .91. In addition, we analyzed the mathematical explanations that students provided in response to the nine open-ended items, as described in the next section.

Mathematical explanations could be correct or incorrect and could be expressed in formal or informal mathematical language. We considered an explanation to be correct if the argument fit the conventions, even if it was not expressed in a formal mathematical language.

Students could use one or more arguments to explain their reasoning. We analyzed their explanations by focusing on two dimensions: fluency and flexibility (California Learning Assessment System, 1993).

Fluency refers to the number of correct arguments provided by students. In addition, students also provided incorrect arguments. No significant differences were found between groups on the number of incorrect arguments before the beginning and at the end of the study, $F(3, 372) < 1$, $p > .05$; therefore, we will not report further on fluency of incorrect arguments. Hereinafter, fluency refers only to the total number of correct arguments. *Flexibility* refers to a student's providing more than one kind of correct argument to justify his or her reasoning.

We classified the arguments into four categories: logical-formal, numerical-computational, visual, and drawing. Definitions and examples are provided below. All examples refer to the following task, in which students were directed to examine the graph shown in Figure 1:

> The graph below represents the income of two companies between the years 1990 and 2000. Until the year 1994, was the change rate in the income of Company B greater than, smaller than, or equal to the change rate in the income of Company A? After 1994, did the change rates in the companies' incomes become different? Please explain your reasoning.

Student responses were classified as *logical-formal* (based on logical-mathematical arguments), *numerical-computational* (based on numerical computations or algebraic formulas), *visual* (based on intuitive, visual analysis of the graph), or *drawing* (based on drawings that students added to the graph). Here are examples of correct and incorrect responses of each type:

> Logical-formal
>
> > *Correct:* "The change rate of line A is greater because its slope is steeper than that of line B"; "The change rate of line A is greater because the angle it creates with the *x*-axis is bigger than the angle that line B creates with the *x*-axis."

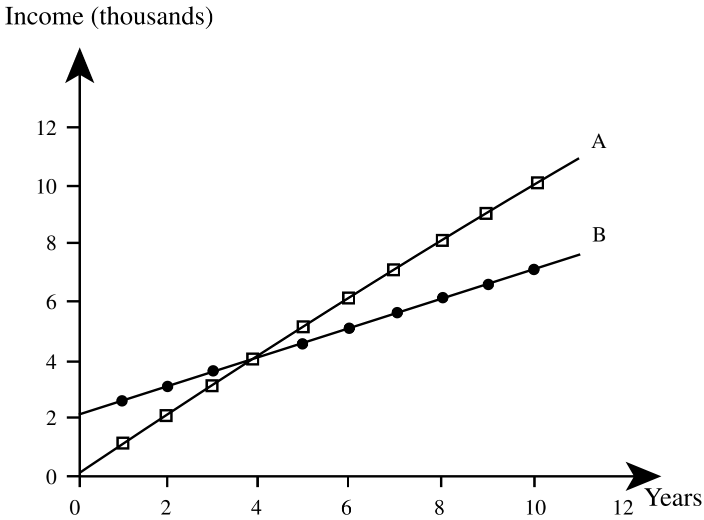


Figure 1. **The graph represents the income of two companies between the years 1990 and 2000.**

Incorrect: "The change rate of line A is smaller because it starts from the origin (0, 0)."

Numerical-computational

Correct: "The change rate of line A is 3 times greater than the change rate of line B."

Incorrect: "The change rate of line A is 5 times greater than the change rate of line B."

Visual

Correct: "Line A is steeper, line A is more diagonal."

Incorrect: "At M the line breaks"; "Line A is above line B."

Drawing

Correct: Adding one-unit steps to the graph and calculating the change rate by using the steps.

Incorrect: Adding lines that do not assist in finding the change rate.

Because students could use more than one kind of argument in explaining their reasoning, one explanation could be classified in several categories. Figure 2 presents a diagram of the categories used in the present study to classify students' mathematical explanations.

Two judges who were experts in mathematics education analyzed the students' explanations. The interjudge reliability coefficient was .88.

Graph Construction Test (Transfer Tasks)

The graph construction test, adapted from the study by Mevarech and Kramarski (1997b), assessed students' ability to construct graphs. That test was regarded as assessing transfer knowledge because in none of the classrooms did the students study graph construction. According to Leinhardt et al., "construction is quite different from interpretation. Whereas interpretation relies on and requires reaction to a given piece of data (e.g., a graph, an equation, or a data set) construction requires generating new parts that are not given" (1990, p. 12). Graph construction requires interpretation skills and involves more difficulties than does graph interpretation (Mevarech & Kramarski).

The graph construction test contained 7 items, each presenting a verbal description of a situation. The situations represented increasing, decreasing, constant, and curvilinear functions. Students were asked to transform the verbal descriptions into graphic representations. They could choose any kind of representation and were allowed to construct the graphs freehand, without using a ruler. Appendix B provides several examples.

For each item, students received a score of either 1 (correct answer) or zero (incorrect answer), and a total score ranging from zero to 7. A graph was considered correct if it followed the conventions of the Cartesian system and

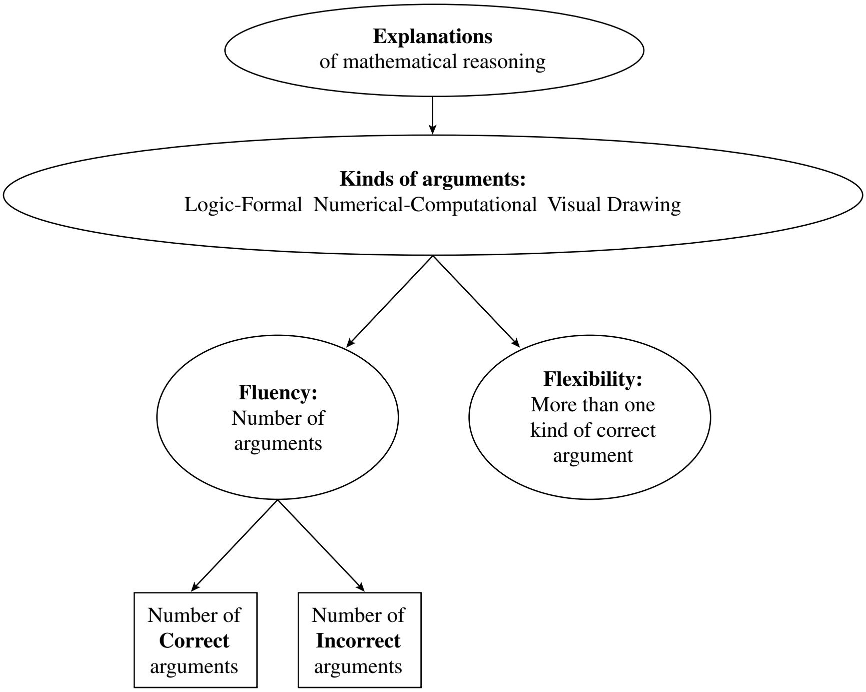


Figure 2. **Categories of mathematical explanations.**

represented correctly the situation described, regardless of the kind of graph used (e.g., histogram, bar graph, or line graph). Two judges who are experts in mathematics education scored students' responses. The interjudge reliability coefficient was .92.

Metacognitive Questionnaire

The metacognitive questionnaire, adapted from a study by Montague and Bos (1990), assessed students' general and specific metacognitive knowledge regarding graph comprehension. The questionnaire included 20 items: 7 items referred to general strategies (e.g., "When I solve a math problem I read the problem several times before attempting a solution"); 4 items referred to strategies for transformation from one kind of representation to another (e.g., "When I see a graph, I try to find out the algebraic expression it presents"); and 9 items referred to specific quantitative and qualitative graph interpretation strategies (e.g., "When I see a graph, I first look at specific points" [quantitative interpretation]; "When I see a graph, I first try to find out the general trend of the graph" [qualitative interpretation].

Each item was constructed on a 5-point Likert-type scale, ranging from 1 (never) to 5 (always). Thus students received scores on each item, ranging from 1 to 5, and a total score from 20 to 100. The alpha Cronbach coefficient was .86.

Procedure

As indicated, 12 classrooms from four schools participated in the study. Schools were randomly assigned to conditions, and in each school 3 classrooms (out of 6 parallel eighth-grade classrooms) were randomly selected to participate in the present study. We did not assign classrooms to different conditions within one school because our experience shows that teachers in the same school tend to share materials and talk to each other about their teaching activities.

About a month after the beginning of the school year, all students were administered the examinations described earlier: Graph Interpretation Test, Graph Construction Test, and Metacognitive Questionnaire. Then, each teacher began teaching the unit according to the instructional method to which she was assigned, using the materials specially designed for that condition. To ensure that the instruction was properly implemented as designed, all classrooms were observed twice a week by one of the authors of this article. At the end of the study, the same battery of tests was readministered.

Results

Mathematical Reasoning

The first purpose of the present study was to investigate the differential effects of COOP+META, IND+META, COOP, and IND on students' mathematical reasoning. Because a significant correlation was found between graph interpretation and graph construction scores ($r = .48$), a MANCOVA (Wilks's lambda test) was carried out on the posttreatment scores of those two variables simultaneously with classrooms nested in conditions and controlling for pretreatment on graph interpretation and graph construction. Before carrying out the MANCOVA, we checked the prerequisites for running it, $MS_e = 2.3$, $F(6, 743) < 1$, $p > .05$.

The results indicated significant differences between conditions on the posttest scores of graph interpretation and graph construction simultaneously, $MS_e = 10.7$, $F(6, 744) = 6.17$, $p < .001$.

Graph Interpretation

Given the MANCOVA findings, we carried out a one-way analysis of variance (ANOVA) on the pretest scores and a one-way analysis of covariance (ANCOVA) on the posttest scores with classrooms nested in conditions and with the corresponding pretest scores used as a covariant.

Table 1 presents the mean scores, adjusted mean scores, and standard deviations on graph interpretation by time and condition. The ANOVA with classrooms nested in conditions indicated no significant differences between conditions before the beginning of the study, $MS_e = 38.4$, $F(3, 372) = .71$,

Table 1
Scores on Graph Interpretation Test (by Time and Condition)

	COOP+META $n = 105$	IND+META $n = 95$	COOP $n = 91$	IND $n = 93$	F
Pretest M	15.5	14.4	14.2	16.0	< 1
SD	6.4	6.4	5.9	6.2	
Posttest M	24.4	20.9	19.2	19.8	3.98**
Adjusted M	24.0	21.4	19.8	19.1	
SD	7.2	6.9	6.4	6.6	

Note. Scores ranged from zero to 36.
**$p < .01$.

$p > .05$. Significant differences between treatment groups, however, were found at the end of the study after controlling for pretreatment differences, $MS_e = 44.6$, $F(3, 371) = 3.98$, $p < .05$.

Post hoc analyses of the adjusted mean scores based on the pairwise comparison t test technique indicated that the COOP+META group significantly outperformed the IND+META group, which in turn significantly outperformed the COOP and IND groups. No significant differences were found between the two groups that had not been exposed to the metacognitive training.

Students' Mathematical Explanations

In addition to analyzing the total scores on graph interpretation, we also analyzed students' explanations of their mathematical reasoning. As was indicated earlier, two dimensions of mathematical explanations were investigated: fluency (number of correct arguments) and flexibility (provision of more than one kind of correct argument).

Fluency

Table 2 presents the mean scores, adjusted mean scores, and standard deviations of students' fluency in providing correct arguments by time and condition. Although no significant differences were found between groups before the beginning of the study, $MS_e = 5.25$, $F(3, 372) = 1.38$, $p > .05$, significant differences were found at the end of the study with classrooms nested in conditions and with the corresponding pretest scores used as a covariant, $MS_e = 11.95$, $F(3, 371) = 7.53$, $p < .01$.

Post hoc analyses of the adjusted mean scores based on the pairwise comparison t test technique indicated that the COOP+META group outperformed the IND+META group, which, in turn, significantly outperformed the COOP and IND groups. However, no significant differences were found on that measure between the two groups that had not been exposed to metacognitive training.

Table 2
Scores for Fluency in Providing Correct Arguments (by Time and Condition)

	COOP+META $n = 105$	IND+META $n = 95$	COOP $n = 91$	IND $n = 93$	F
Pretest					
M	3.2	2.8	2.7	3.7	1.38
SD	2.4	2.2	2.1	2.6	
Posttest					
M	8.9	6.5	4.9	4.6	7.53*
Adjusted *M*	8.9	6.7	5.2	4.2	
SD	5.0	4.4	3.2	2.9	

Note. Fluency was scored by the number of correct arguments provided by a student. Scores ranged from zero to 21 for COOP+META, from zero to 18 for IND+META, from zero to 15 for COOP, and from zero to 13 for IND.
*$p < .05$.

Table 3 presents the number of students in each condition (percentages are in parentheses) who used each type of correct argument, by time and treatment. As can be seen from the table, under all conditions most students (60%) relied on numerical-computational arguments in justifying their reasoning. It is interesting to note that in the individualized groups (with or without metacognitive training), even more students (64%) did so than in the

Table 3
Frequency of Four Types of Correct Arguments (by Time and Condition)

	COOP+META $n = 105$	IND+META $n = 95$	COOP $n = 91$	IND $n = 93$
Logical-formal				
Pretest	19 (18.1)	16 (16.8)	6 (6.6)	18 (19.4)
Posttest	31 (29.5)	14 (15.3)	13 (14.3)	19 (20.4)
Numerical-computational				
Pretest	50 (47.6)	46 (48.4)	47 (51.6)	40 (43.0)
Posttest	54 (51.4)	61 (64.2)	55 (60.4)	60 (64.5)
Visual				
Pretest	6 (5.7)	3 (3.2)	4 (4.4)	5 (5.4)
Posttest	7 (6.6)	7 (7.4)	5 (5.5)	0
Drawing				
Pretest	0	0	0	1 (1.1)
Posttest	6 (5.7)	5 (5.3)	0	0

Note. Percentages (the figures in parentheses) for each answer category were calculated by dividing the number of students who gave answers in that category by the total number of students in the same condition.

cooperative groups, where 51% with metacognitive training used such arguments and 60% without such training did so. In addition to using numerical-computational arguments, students quite often used logical-formal arguments. The frequency of using logical-formal arguments, however, was significantly larger under the COOP+META condition (29.5%) than under all other conditions (15.3%, 14.3%, and 20.4% for the IND+META, COOP, and IND conditions, respectively). These differences were statistically significant ($\chi^2 = 27.0$, $p < .001$). Further analyses showed that under all conditions, students used the visual and drawing arguments quite infrequently (fewer than 8% of the students under each condition).

Flexibility

As indicated earlier, the term *flexibility* refers to a student's provision of more than one kind of correct argument to justify his or her reasoning. For example, referring to the task described earlier, a student responded as follows: "The rate of change of Company A is greater because its rate of change is 1 and that of Company B is 1/3. We can also see it by looking at the slopes. The slope of line A is steeper than the slope of line B."

Initial analysis of students' arguments indicated that students very rarely provided more than two kinds of argument to explain their ideas regarding the solution of a given task. Under each condition this kind of response occurred two or three times. Therefore, in further analysis we distinguished between students who provided one kind of correct argument and those who provided more than one kind of correct argument to justify their reasoning on each task. Table 4 presents that information. Although before the study there had been no significant differences ($\chi^2 = 9.3$, $p > .05$) between conditions on flexibility, at the end of the study the chi-square analyses showed significant differences on all tasks but two. According to Table 4, on average about 27%, 12%, 6%, and 4% of the students under the COOP+META, IND+META, COOP, and IND conditions, respectively, justified their mathematical ideas by basing their explanations on more than one kind of correct argument. Significant differences between conditions were also found among the students who provided one kind of correct argument to explain their reasoning, but the differences were not as large as those among students who provided more than one kind of argument to explain their reasoning. According to Table 4, on average about a quarter of the students under COOP+META (28%) and IND+META (26%) provided one kind of correct argument, as compared with 21% and 22% of the students under the COOP and IND conditions, respectively.

Graph Construction (Transfer Task)

The second purpose of the present study was to investigate the differential effects of COOP+META, IND+META, COOP, and IND on students' ability in graph construction, to which they were not exposed in the classroom. Table 5

Table 4

Percentages of Students Who Displayed Flexible Reasoning (by Task and Condition)

	COOP+META $n = 105$		IND+META $n = 95$		COOP $n = 91$		IND $n = 93$		
	One argument	More than one argument	One argument	More than one argument	One argument	More than one argument	One argument	More than one argument	χ^2
Task 1	22.3	35.9	26.8	16.4	30	14.5	24.7	7.5	57.9***
Task 2	32.1	39.8	37.6	20.3	37.8	21.1	26.6	21.5	22.1**
Task 3	15.5	24.3	18.3	11.8	12.2	4.4	19.4	3.2	22.0**
Task 4	30.6	15.6	29.0	12.4	26.7	5.6	31.2	3.2	23.0**
Task 5	25.2	31.1	23.7	16.2	16.7	5.6	18.3	1.1	46.2***
Task 6	35.9	28.2	38.7	14.0	21.1	3.3	24.7	2.2	30.0**
Task 7	38.8	20.4	31.2	14.0	17.8	1.1	32.3	2.2	40.3***
Task 8	15.8	1.3	14.0	0	11.6	0	11.1	0	0.7
Task 9	38.8	1.9	13.9	1.1	14.4	0	11.5	1.0	1.2

Note: A student was said to display flexibility when he or she provided more than one type of correct argument to justify a solution to a problem. The percentages in each argument category were calculated by dividing the number of students who gave arguments in that category by the total number of students in the same condition.

$^{**}p < .01$. $^{***}p < .001$.

Table 5
Scores on Graph Construction Test (by Time and Condition)

	COOP+META $n = 105$	IND+META $n = 95$	COOP $n = 91$	IND $n = 93$	F
Pretest M	3.4	3.3	3.5	3.5	< 1
SD	2.3	2.5	2.1	2.4	
Posttest M	5.3	5.0	4.5	4.6	7.19*
Adjusted M	5.4	5.1	4.4	4.5	
SD	2.3	2.2	2.0	2.1	

Note: Scores on the graph construction test ranged from zero to 7. This was a transfer task.
*$p < .05$.

presents the mean scores, adjusted mean scores, and standard deviations on graph construction by time and condition. An ANOVA of classrooms nested within conditions indicated no significant differences between conditions before the beginning of the study, $MS_e = 5.22$, $F(3, 372) = .2$, $p > .05$, but at the end of the study, significant differences were found between conditions controlling for pretreatment differences, $MS_e = 2.95$, $F(3, 371) = 7.19$, $p < .01$.

Post hoc analyses of the adjusted mean scores based on the pairwise comparison *t* test technique indicated significant differences between the metacognitive groups (COOP+META and IND+META) and the non-metacognitive groups (COOP and IND), but no significant differences were found between the two metacognitive groups or between the two non-metacognitive groups.

Metacognitive Knowledge

The third purpose of the present study was to investigate the differential effects of COOP+META, IND+META, COOP, and IND on students' metacognitive knowledge. A one-way MANCOVA (Wilks's lambda test) on the two criteria (general strategy and specific strategy), which were used as dependent variables with classrooms nested in conditions and with the pretest scores used as a covariant, indicated significant differences between conditions on both general and domain-specific metacognitive criteria simultaneously, $MS_e = .29$, $F(6, 744) = 2.97$, $p < .01$. Before carrying out the MANCOVA, we checked the prerequisites for running it, $MS_e = 2.8$, $F(6, 743) < 1$, $p > .05$.

Table 6 presents the mean scores, adjusted mean scores, and standard deviations by time and condition on general and domain-specific metacognitive knowledge. As one may see from Table 6, no significant differences were found between conditions on any aspect of the metacognitive measure before the beginning of the study (F values ranged from .65 to 0.94, all p values > .05). Yet at the end of the study, significant differences were found between conditions on domain-specific metacognitive knowledge with classrooms nested in conditions, $MS_e = 0.64$, $F(3, 371) = 7.75$, $p < .01$, controlling for pretreatment differences, but not on general metacognition, $F(3, 371) < 1$, $p > .05$.

Table 6
Scores on Metacognitive Knowledge of General and Domain-Specific Strategies (by Time and Condition)

	COOP+META $n = 105$	IND+META $n = 95$	COOP $n = 91$	IND $n = 93$	F
General strategies					
Pretest					< 1
M	3.0	3.0	2.9	3.0	
SD	.6	.6	.6	.6	
Posttest					< 1
M	3.1	3.1	3.0	3.0	
Adjusted *M*	3.1	3.1	.92	3.0	
SD	.6	.5	.5	.5	
Domain-specific strategies					
Pretest					< 1
M	3.2	3.3	3.3	3.3	
SD	.6	.6	.6	.6	
Posttest					7.75*
M	3.5	3.5	3.3	3.3	
Adjusted *M*	3.5	3.5	.23	3.3	
SD	.6	.5	.6	.5	

Note: Scores ranged from 1 to 5.
*$p < .05$.

Post hoc analyses of the adjusted mean scores based on the pairwise comparison *t* test technique indicated significant differences between the metacognitive groups (COOP+META and IND+META) and the non-metacognitive groups (COOP and IND), but no significant differences were found between the two metacognitive groups or between the two non-metacognitive groups.

In sum, students who were exposed to the metacognitive training in either cooperative or individualized settings significantly outperformed the other students on graph interpretation (total scores), fluency and flexibility of correct explanations, use of logical-formal arguments to justify their reasoning, and transfer tasks (graph construction). In addition, the metacognitive groups attained higher levels of domain-specific metacognitive knowledge than the non-metacognitive groups. Table 7 summarizes these findings.

Discussion

The present study compared the effects of the four instructional methods (COOP+META, IND+META, COOP, and IND) on mathematical reasoning, transfer of knowledge, and metacognitive knowledge. Although all four methods focused on promoting mathematical comprehension, each method was based on a different theoretical and operational approach. Therefore, the results should not be used to draw conclusions about which method is more

Table 7
Summary of Students' Performance (by Dependent Variable and Condition)

Dependent variable	Findings
1. Mathematical reasoning	
Graph interpretation (total score)	4 > 3 > 2 = 1
Mathematical explanations	
Fluency	4 > 3 > 2 = 1
Logical-formal	4 > 3 = 2 = 1
Numerical-computational	4 = 3 = 2 = 1
Visual	4 > 3 = 2 > 1
Drawing	4 = 3 > 2 = 1
Flexibility	4 > 3 > 2 = 1
2. Graph construction (transfer task)	4 = 3 > 2 = 1
3. Metacognitive knowledge	
General strategies	4 = 3 = 2 = 1
Specific strategies	4 = 3 > 2 = 1

Note: 1 = IND, 2 = COOP, 3 = IND+META, 4 = COOP+META. Use of the equals sign (=) indicates that there are no significant differences between two groups. The greater-than sign (>) indicates a significant difference in the mean scores of two groups.

"efficient." Rather, the purpose is to increase understanding of how each program operates relative to the others on these measures. It should also be emphasized that the present study focused on only one instructional unit: linear graphs. Although that unit is essential in the mathematics curriculum (NCTM, 2000), longitudinal investigations and large-scale studies that focus on other instructional units are needed to document the changes in other schooling outcomes. Moreover, methods for cooperative learning (e.g., Slavin, 1996) and metacognitive instruction (e.g., King, 1994; Schoenfeld, 1985) vary. Therefore, the generalizability of our findings may be limited to cooperative settings and metacognitive training that are similar to the ones we explored in this study.

Mathematical Reasoning

In the present study we analyzed mathematical reasoning in two complementary ways: One focused on the final answers, and the other on the written explanations provided by students to justify their mathematical ideas. The dual focus enabled us to develop a better understanding of how students learn to think flexibly about linear relationships and their representations in tables, graphs, and equations. Sometimes the final answers were correct but based on wrong lines of reasoning, as seen in the following example of a student's response to the question about the graph in Figure 1: "The change rate in the income of company A is greater [the conclusion is correct] because its line is above Company B [the explanation is incorrect]." In other cases the final answers were incorrect but the explanations were correct. That happened most frequently when students made mistakes in the calculations or in the solution of the equations. By formulating written explanations, students learned valuable lessons about the need for accuracy, precision, and completeness in their answers.

The data further showed that the average numbers of correct arguments were 4.2, 5.2, 6.7, and 8.9 for IND, COOP, IND+META, and COOP+META students, respectively. These data indicate that the number of arguments provided by IMPROVE students was almost double the number of arguments provided by IND students. Furthermore, IMPROVE students frequently used more logical-formal arguments, whereas other groups based their explanations mainly on numerical- computational arguments. The reason may be that, when studying the unit, IMPROVE students not only had to explain their strategies to their peers and the reasons for using those strategies but also had to analyze, compare, and contrast the meaningfulness of the tasks and their solutions.

The emphasis on fluency and flexibility, two important elements of mathematical discourse, has several implications. First, students tend to conceive of mathematics as a rigid subject based on arbitrary rules that allow only one correct answer and one correct argument to explain the answer (Schoenfeld, 1985). By contrast, it seems that IMPROVE teachers (and students) succeeded in building a mathematical community where students expressed their mathematical ideas fluently and flexibly, using various kinds of arguments to justify their ideas. Second, according to cognitive psychologists, elaboration is an important means for enhancing understanding (Wittrock, 1986). Our study confirmed that students who were exposed to metacognitive training did attain a higher level of mathematics achievement and were better able to explain their mathematical ideas in writing. In comparison with the written explanations of all other groups, those of IMPROVE students were longer, more accurate, and more frequently embedded with the new mathematical terms that had been introduced in the unit. Often their explanations included mathematical arguments and rationales, not just procedural descriptions or summaries (e.g., "The change rate of line A is greater because the angle it creates with the *x*-axis is larger then the angle that line B creates with the *x*-axis"). These findings are in line with the New Standards (NCTM, 2000), as well as with current studies showing that fluency and flexibility are essential components of mathematical thinking (Nohda, 2000; Brenner, Herman, Ho, & Zimmer, 1999; Dreyfus & Eisenberg, 1996).

Finally, beyond context and content, there is the role of the teacher. The study illustrates several important facets of teachers' roles, pertaining to metacognitive guidance, organization of the classroom, and the selection and use of worthwhile mathematical tasks that allow significant mathematical discourse to occur. Such tasks should include complex situations that present quantitative information in different contexts, allow multiple representations, or afford students opportunities to resolve mathematical conflicts (Mevarech & Kramarski, 1997a; NCTM, 2000). For example, the task described earlier, answering questions about the change rates of Companies A and B, was in many ways quite simple. It provided students with an opportunity to use their understanding of linear graphs, rate of change, slopes, and speed—important mathematical ideas in the middle grades. The task was simple enough that all students could perform it in a test situation, difficult enough to challenge students to think and reason about the properties of linear graphs and the applications of abstract mathematical ideas to everyday situations, and rich enough to allow students to engage in various levels of thinking—using quantitative or qualitative interpretations (Wainer, 1992) or computational or logical-formal arguments (Leinhardt et al., 1990).

Transfer Knowledge

The findings indicate that students who were exposed to metacognitive training were better able to transfer their knowledge from graph interpretation, which was taught in all classrooms, to graph construction, which was new to all students. Salomon, Globerson, and Gutterman (1989) reported similar findings, which showed that students who were exposed to metacognitive training in the area of reading comprehension improved their scores not only in reading comprehension but also in writing tasks (e.g., writing compositions). These findings also support earlier conclusions (Mevarech, 1999; Hoek, Eeden, & Terwel, 1999; Kramarski, Mevarech, & Liberman, 2001) that metacognitive training leads students to link new and existing knowledge.

It is interesting that on the transfer task we found no significant differences between students who were exposed to the metacognitive training in cooperative settings and those who were exposed to the training in individualized settings (COOP+META and IND+META). Two plausible explanations are that (a) a seven-item test and scoring procedure based on correct and incorrect responses is not sensitive enough to assess students' transfer ability; and (b) the metacogitive questions were internalized by both groups to such an extent that students' interactions could have only a small additional impact on transfer performance. Indirect support for the latter explanation comes from the analysis of students' metacognitive knowledge, which showed no significant differences between the COOP+META and the IND+META groups on several aspects of metacognitive knowledge.

Metacognitive Knowledge

The findings also showed that the two groups that were exposed to metacognitive training (COOP+META and IND+META) scored higher on the metacognitive questionnaire than the two groups that were not exposed to metacognitive training (COOP and IND).

It is interesting to note, however, that the differences between the metacognitive and non-metacognitive groups were observed only on domain-specific metacognitive knowledge but not on general metacognitive knowledge. That finding points to the importance of broadening the distinction between general and domain-specific knowledge (e.g., Salomon & Perkins,

1987) in the area of metacognition. Future research based on interviews and observations may explore the development of general and specific metacognitive knowledge under various conditions.

Comparisons Among the Conditions

Several issues relating to the comparison among the conditions need further consideration.

First, why did the COOP+META students outperform the IND+META students on graph interpretation (total scores), fluency, and flexibility in providing correct mathematical explanations? Two factors may explain these findings. One, it is possible that the type of metacognitive training used in the present study is more appropriate for cooperative than for individualized settings because learning in small groups provides a natural setting for students to formulate and discuss questions such as those used in the present study (Mevarech & Susak, 1993). Future research may focus on metacognitive training programs that are explicitly designed to be used in individualized settings. Two, cognitive psychologists (e.g., Sweller, Merrienboer, & Paas, 1998) indicate that the use of dual presentation techniques involving both auditory and visual-writing forms increases cognitive performance more than does the use of one form. Because students in the COOP+META condition more often used dual presentation forms, they were likely to outperform their counterparts in the IND+META condition who used mainly the visual-writing form.

Second, why did the COOP+META students outperform the COOP group on both aspects of mathematical reasoning? A partial answer relates to the quality of the mathematical discourse in the two cooperative conditions: Students who had been exposed to the metacognitive training in cooperative settings were better able to express their mathematical ideas in writing than were students who had studied in cooperative settings without metacognitive training. That finding was also consistent with studies by Yager, Johnson, Johnson, and Schneider (1986), and Webb (1989, 1991), who showed that the quality of discourse in the groups related to students' mathematical achievement. In particular, Webb indicated that giving and receiving elaborated help is more strongly related to mathematics achievement than is giving or receiving final answers without elaboration or asking for help and not receiving it. To gather further information on this question, students' interactions under both conditions should be videotaped and analyzed.

Third, why did the IND+META students outperform the COOP students? The cooperative learning approach is rooted in cognitive theories assuming that elaboration is an important means for enhancing understanding and that cooperative settings are appropriate contexts for encouraging students to elaborate information. The present study indicated that placing students in cooperative groups is not sufficient for enhancing mathematical reasoning. According to the present study, metacognitive training is an effective means of facilitating mathematical reasoning, even when the training is implemented in individualized settings.

Finally, why did the COOP students not outperform the IND students on mathematical reasoning? From a cognitive perspective, students who have poor communication skills are less likely to benefit from cooperative learning because they are not able to communicate their mathematical ideas and the strategies they have used, nor do they know how to ask questions, reflect on their reasoning to explain it to their peers, or specify what they do not understand. These findings support earlier studies arguing that there is a need to structure learning in small groups and that features of discourse such as "given reasons" must be practiced and reinforced (Webb, 1991; Webb & Farivar, 1994; Cohen, 1996; Mevarech & Kramarski, 1997a).

Practical Implications and Future Research

Assuming that these findings generalize to other settings, the study suggests several important practical implications. First, it appears that under certain conditions students in junior high schools can learn to provide mathematical arguments to justify their ideas. In particular, IMPROVE students under the COOP+META condition were more fluent and flexible than the other students in explaining their ideas in writing. These findings call for the design of additional learning environments based on similar components. Such environments are desirable at all grade levels and for all mathematical topics. There is a need, therefore, to adapt IMPROVE and other COOP+META programs to the needs of primary school children. The issue of how children's mathematical explanations develop under various conditions merits future research.

A second issue relates to the exposure of students to the COOP+META program in various classrooms, not only in mathematics classrooms. There is reason to suppose that if students study under the COOP+META condition in mathematics as well as in other subjects (e.g., languages), they will be better able to internalize and activate metacognitive processes. This issue is open for future research.

A third issue relates to the norms for evaluating mathematical reasoning in junior high school classrooms. Our study emphasizes the importance of changing classroom organization when providing metacognitive training. The study also describes the kinds of arguments that students use in their mathematical explanations. At present, many state proficiency tests and international examinations (e.g., TIMSS-1999 and PISA, administered by OECD [Organisation for Economic Co-operation and Development] countries) include tasks that require students to explain their reasoning in writing. To acquaint students with such tasks and the scoring procedure, teachers may prepare guidelines and ask students to score one another's explanations by using the guidelines and activating metacognitive processes. Thus students will learn how to construct explanations that are coherent, clear, and precise. Researchers in mathematics education may design "task banks" that challenge students to provide explanations. The task banks may be followed by analyses of students' correct and incorrect arguments.

References

Brenner, M. E., Herman, S., Ho, H. Z., & Zimmer, M. J. (1999). Cross-national comparison of representational competence. *Journal for Research in Mathematics Education, 30*(5), 541–557.

Brown, A., & Palincsar, A. (1989). Guided cooperative learning and individual knowledge acquisition. In L. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393–451). Hillsdale, NJ: Lawrence Erlbaum.

Cai, J., Lane, S., & Jakabcsin, M. S. (1996). The role of open-ended tasks and holistic scoring rubrics: Assessing students' mathematical reasoning and communication. In P. C. Elliott & M. J. Kenney (Eds.), *Communication in mathematics, K–12 and beyond* (pp. 137–145). Reston, VA: Academic Press.

California Learning Assessment System. (1993). *Portfolio assessment: Final report*. San Diego: CA: Center for Performance Assessment, Educational Testing Service.

Cecil, S. J., & Roazzi, A. (1994). The effects of context on cognition: Postcards from Brazil. In J. S. Sternberg & R. K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 74–100). Cambridge, UK: Academic Press.

Cobb, P., Boufi, A., McClain, K., & Whitenack, J. (1997). Reflective discourse and collective reflection. *Journal for Research in Mathematics Education, 28,* 28–47.

Cohen, E. G. (1996). *A sociologist looks at talking and working together in the mathematics classroom*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Davidson, N. (1990). *Cooperative learning in mathematics: A handbook for teachers*. Menlo Park, CA: Addison-Wesley.

Dreyfus, T., & Eisenberg, T. (1996). On different facets of mathematical thinking. In R. J. Stermberg & T. Ben-Zeev (Eds.), *The Nature of Mathematical Thinking* (pp. 253–284). Mahwah, NJ: Lawrence Erlbaum.

Fendel, C., Resek, D., Alper, L., & Fraser, S. (1996). *Interactive mathematics program (IMP)*. Berkley, CA: Key Curriculum Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychology, 34,* 906–911.

Forrest-Pressley, D. L., Makinnon, T., & Gary Waller, T. (1985). *Metacognition, cognition, and human performance*. Orlando, FL: Academic Press.

Garofalo, J., & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education, 16,* 163–176.

Hirsch, C., Coxford, A., Fey, J., & Schoen, H. (1997). *Contemporary mathematics in context: A unified approach*. (Core-Plus Mathematics Project). Chicago, IL: Janson.

Hoek, D., Van den Eeden, P., & Terwel, J. (1999). The effects of integrated social and cognitive strategy instruction on the mathematics achievement in secondary education. *Learning and Instruction, 9*(5), 427–448.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14,* 366–381.

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *Journal of Educational Psychology, 83,* 307–317.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31,* 338–368.

Kramarski, B., Mevarech, Z. R., & Liberman, A. (2001). The effects of multilevel—versus unilevel—metacognitive training on mathematical reasoning. *Journal for Educational Research, 94*(5), 292–300.

Lappan, G., Fey, J., Fitzgerald, W., Freil, S., & Phillips, E. (1996). *Connected mathematics (Grades 6–8)*. Palo Alto, CA: Dale Seymour.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60,* 1–64.

Lester, F. K., Garofalo, J., & Kroll, D. L. (1989). *The role of metacognition in mathematical problem solving: A study of two grade-seven classes,* pp. 1–126. (ERIC Document Reproduction Service No. ED314255)

Marx, R. W., & Walsh, J. (1988). Learning from academic tasks. *Elementary School Journal, 88,* 207–219.

Meloth, M. S., & Deering, P. D. (1994). Task talk and task awareness under different cooperative learning conditions. *American Educational Research Journal, 31,* 138–165.

Mevarech, Z. R. (1999). Effects of metacognitive training embedded in cooperative settings on mathematical problem solving. *Journal of Educational Research, 92*(4), 195–205.

Mevarech, Z. R. & Kramarski, B. (1992). How and how much can cooperative Logo environments enhance creativity and social relationships? *Learning and Instruction, 2,* 259–274.

Mevarech, Z. R. & Kramarski, B. (1993). Vygotsky and Papert: Social-cognitive interactions with Logo environments. *British Journal of Educational Psychology, 63,* 96–109.

Mevarech, Z. R., & Kramarski, B. (1997a). IMPROVE: A multidimensional method for teaching mathematics in heterogeneous classrooms. *American Educational Research Journal, 34,* 365–394.

Mevarech, Z. R., & Kramarski, B. (1997b). From verbal descriptions to graphic representations: Misconceptions regarding the construction of linear graphs. *Educational Studies in Mathematics, 32,* 22–263.

Mevarech, Z. R., & Susak, Z. (1993). Effects of learning with cooperative-mastery learning method on elementary students. *Journal of Educational Research, 86,* 197–205.

Montague, M., & Bos, C. S. (1990). Cognitive and metacognitive characteristics of eighth-grade students' mathematical problem solving. *Learning and Individual Differences, 2,* 371–388.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. New York: Routledge & Kegan Paul.

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

Nohda, N. (2000). Teaching by open-approach method in Japanese mathematics classroom. In T. Nakahara & M. Koyama (Eds.), *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education* (pp. 1–39). Hiroshima, Japan: Academic Press.

Organisation for Economic Co-operation and Development. (2000). Measuring students' knowledge and skills. *The PISA 2000 assessment of reading, mathematical and scientific literacy.*

Polya, G. (1945). *How to solve it?* Garden City, NY: Doubleday.

Robertson, L., Regan, S., Contestable, J., & Freeman, M. (1996). *Number power: A cooperative approach to mathematics and social development (Grades K–6).* Menlo Park, CA: Addison Wesley.

Salomon, G., Globerson, T., & Gutterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognition from a reading partner. *Journal of Educational Psychology, 81,* 620–627.

Salomon, G., & Perkins, D. N. (1987). Transfer of cognitive skills from programming: When and how? *Journal of Educational Computing Research, 3,* 149–169.

Schneider, W., & Sodian, B. (1990). Children's understanding of cognitive cueing: How to manipulate to fool a competitor. *Child Development, 1*(3), 697–704.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. San Diego, CA: Academic Press.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive Science and Mathematics Education* (pp. 189–215). Hillsdale, NJ: Erlbaum.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21,* 43–69.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.

Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher, 21*(1), 14–23.

Webb, N. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13,* 21–40.

Webb, N. (1991). Task-related verbal interaction in mathematics learning in small groups. *Journal for Research in Mathematics Education, 22,* 366–389.

Webb, N., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31,* 369–396.

Wittrock, M. C. (1986). Students' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315–327). New York: Macmillan.

Yager, S., Johnson, R. T., Johnson, D. W., & Schneider, B. (1986). The impact of group processing on achievement in cooperative learning. *Journal of Social Psychology, 126,* 389–397.

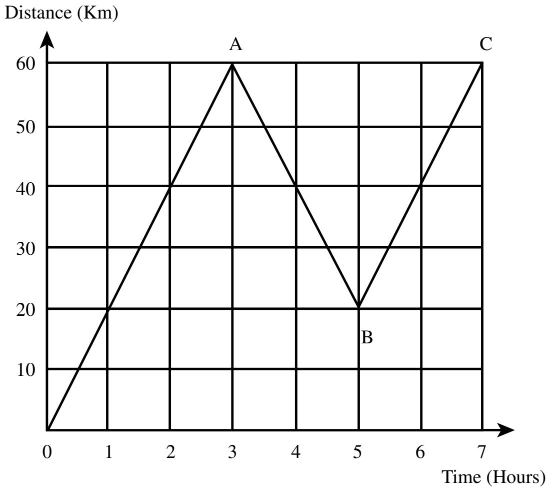
Manuscript received January 6, 1998
Revision received February 13, 2000
Accepted May 1, 2001

APPENDIX A

Presentation of an Open-Ended Graph Task in the Metacognitive and Non-Metacognitive Conditions

The task:

The following graph describes the distance that seventh-grade students traveled in 7 hours on an end-of-school-year trip. Was the speed during the trip constant? Explain your reasoning in writing.



META conditions	Non-META conditions
COOP+META: Discuss the task with your classmates.	COOP: Discuss the task with your classmates.
IND+META: What is the problem/task about? *Comprehension question:* Use the acronym *DATA* to describe the graph: *D*escribe the *x*-axis and the *y*-axis. *A*ddress the units and ranges of each axis. *T*ell the *T*rend(s) of the graph or parts of the graph. *A*nalyze specific points. *Strategic question:* Which strategy or principle is appropriate for solving or addressing the problem or task? *Connection question:* How is this problem or task different from what you have already solved?	IND: What are the students' speeds (a) During the first 3 hours? (b) Between the 3rd and 5th hours? (c) After the 5th hour?

APPENDIX B

A Sample Graph Construction Task: Increasing, Decreasing, Constant, and Curvilinear Graphs

The task:

Construct the following graphs:

(a) The more time a student prepares for tests, the better her grades on the tests are.
(b) The more time a student prepares for tests, the lower her grades on the tests are.
(c) No matter how much time a student spends preparing for tests, her grades on the tests are always the same.
(d) When a student prepares for tests for up to 3 hours, the more time she prepares for tests the better her grades on the tests are; but if she prepares for more than 3 hours, she becomes tired and her grades on the tests decline.