Effects of a Cooperative Learning Program on the Elaborations of Students during Help Seeking and Help Giving

American Educational Research Journal
Spring 2005, Vol. 42, No. 1, pp. 115–151

Effects of a Cooperative Learning Program on the Elaborations of Students During Help Seeking and Help Giving

Simon Veenman, Eddie Denessen,
Anneriet van den Akker, and Janine van der Rijt
Radboud University Nijmegen

In this study, the effects of a teacher-training program on the elaborations and affective-motivational resources (i.e., intentions and attitudes toward help seeking, help giving, and achievement goals) of students working on a cooperative task were examined. Participants were teachers from seven primary schools and 24 dyads of sixth-grade students. In general, the program showed moderately positive effects on use of elaborations among the treatment dyads. Dyads with experience in cooperative learning achieved more than dyads without such experience. Mastery- and performance-oriented goals were negatively related to use of high-level elaborations and to student achievement, while use of high-level elaborations was positively related to student achievement.

KEYWORDS: cooperative learning, help giving, help seeking, student elaborations.

Over the past few decades, research has demonstrated the potential of cooperative learning (CL) to enhance students' academic achievement and social relations (Cohen, 1994; Johnson & Johnson, 1999; Slavin, 1995). In light of its demonstrated efficacy, CL is now regarded in countries such as the United States, Australia, and Israel as an essential part of primary school instruction to foster active student participation and learning (Gillies, 2003b; Shachar, 2003; Webb & Mastergeorge, 2003). Although research has shown cooperative or peer learning to be educationally significant, CL does not have the place it deserves in the Dutch school curriculum. In the majority of primary classrooms, students sit in groups but rarely interact and work as groups. Instead, students work individually or as a class; learning is considered primarily an individual enterprise (Veenman, Kenter, & Post, 2000).

The aforementioned emphasis on individual and whole-class instruction goes hand in hand with a lack of teacher training in CL methods. Dutch teachers are not trained to facilitate learning in small groups and are therefore not familiar with CL. For this reason, a school improvement program focusing on CL and a supplementary program designed to promote the helping behaviors of students working in CL groups were developed by the Department of Educational Sciences at the Radboud University Nijmegen in collaboration with the Christian Pedagogical Study Center, Marant Educational Services (Arnhem/Nijmegen), and the Educational Faculty at the Teacher Education College Arnhem and Nijmegen. The aim of the school improvement program was to enable teachers to implement CL methods in their classrooms and thereby stimulate their students to cooperate successfully with each other in small groups. The effects of this program have been described by Krol (2004). However, cooperation does not automatically lead to productive interactions among students. Cooperation may, rather, make students more aware of the needs of others in the group and thereby the need to provide help and elaborated explanations. The aim of the supplementary program was therefore to provide experience with activities involving helping skills, particularly in the area of elaborated explanations.

While the training of CL methods generally contributes to the learning that occurs in cooperative groups (Gillies, 2003b; Hoek, Terwel, & van den Eeden, 1997; Qin, Johnson, & Johnson, 1995; Webb & Farivar, 1994), a number of questions still remain about how the specific processes of help seeking, help giving, and provision of explanations can best be fostered by instruction and training and how these processes relate to learning outcomes. Thus, the general question to be answered in the present study was as follows: Do students who have been previously trained with respect to cooperative helping behaviors demonstrate more elaborations and achieve more in their cooperative groups than students who have not been previously trained? In addition, relations among provision of elaborated help, performance on a cooperative task, and affective-motivational resources of students were explored.

SIMON VEENMAN is an Associate Professor, Department of Educational Sciences, Spinoza Building, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, the Netherlands; e-mail *s.veenman@pwo.ru.nl.* His research interests include effective teaching, coaching of teachers, multigrade classes, and cooperative and collaborative learning.

EDDIE DENESSEN is an Assistant Professor, Department of Educational Sciences, Spinoza Building, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, the Netherlands; e-mail *e.denessen@pwo.ru.nl.* His research interests include research methodology, the cultural contexts of education, and cooperative learning.

ANNERIET VAN DEN AKKER recently received her master's degree from the Department of Educational Studies, Radboud University Nijmegen. She is now a graduate student at Fontys PABO Limburg, Roermond, the Netherlands; e-mail *anneriet.vandenakker@student.fontys.nl.*

JANINE VAN DER RIJT recently received her master's degree from the Department of Educational Studies, Radboud University Nijmegen. She is now a Research Associate in the Department of Educational Innovation and Information Technology, Faculty of Law, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, the Netherlands; e-mail *janine.vanderrijt@edit.unimaas.nl.* Her research interests include cooperative learning, problem-based learning, and assessment.

In an attempt to offer an understanding of the importance of providing explanations, we first discuss the research on this topic. After that, we discuss two approaches to peer interaction and the framework that these approaches provide for understanding various academic helping behaviors. The cited studies were all conducted in the United States; they cover a variety of topics mainly in the domain of mathematics (e.g., fractions, decimals, time, money), and they encompass various grade levels (from elementary to middle school). Most of the studies also involved heterogeneous (two- or four-person) groups. The overview presented subsequently is by no means comprehensive; such a review is beyond the scope of a single journal article. Excellent reviews on academic helping behaviors have been published by Webb (1989, 1991) and Webb and Mastergeorge (2003).

Student Elaborations

Much of what is learned appears to depend on just how students interact during cooperative work (Fuchs, Fuchs, Hamlett, & Karns, 1998; King, 1994, 1999; Webb & Farivar, 1994, 1999). Students who construct explanations that clarify processes and help classmates arrive at their own solutions have been found to learn more than students who simply tell classmates the solution. In other words, opportunities to construct explanations constitute a critical mediator of successful CL experiences.

Theoretical support for the role of explanations in cooperative learning is provided by sociocognitive theory, based on the ideas of Piaget (1926), and sociocultural theory, based on the ideas of Vygotsky (1978). Both theories have emphasized the role of the social context in construction of knowledge and have stressed that peer interactions provide a rich and necessary context for revision of cognitive systems and creation of new meanings (Cole & Wertsch, 1996; De Lisi & Goldbeck, 1999; Hogan & Tudge, 1999).

One line of research influenced by the work of Piaget and Vygotsky that attempts to disentangle the circumstances in which social interaction benefits learning at school is the cognitive elaboration approach, which emphasizes the cognitive processing performed by interacting students (Slavin, 1996). More specifically, interaction with others can lead to active processing of information and modification of individual cognitive structures. Elaboration refers to the detailed explanation of a topic that occurs when peers provide examples, explain a concept, or supply specific argumentation (Webb & Farivar, 1994, 1999). Verbalization can lead to elaborate cognitive processing and thereby reflection, awareness, (re)organization, differentiation, fine-tuning, and expansion of knowledge (Van Boxtel, 2000). The process of elaboration typically involves comparison of different perspectives or conceptions, development of shared meaning, and joint construction of new knowledge by collaboratively resolving conflicting points of view. Theoretical support for construction of elaborations also comes from a generative model of learning (Wittrock, 1991) positing that, if new information is to be retained and meaningfully related to previously acquired knowledge, students must elaborate or generate connections between that information and representations in memory. One strategy for encouraging elaboration is to have students explain things to others (Fuchs et al., 1998).

Peer Interaction

Two approaches to peer interaction in the classroom can be distinguished in the research literature: the developmental-prosocial tradition and the instructional tradition (Hertz-Lazarowitz, 1989). The *developmental-prosocial tradition* involves a social-developmental perspective on help seeking in the classroom. According to Nelson-Le Gall (1981, 1992), help seeking can be conceived as an activity that serves a constructive purpose and should therefore not be construed as synonymous with dependence. Help seeking is a positive, instrumental skill that encompasses students' attempts to obtain assistance or intervention from another person (e.g., teacher or peer). Nelson-Le Gall distinguished two types of help seeking in the classroom: instrumental help seeking and executive help seeking. Instrumental help seeking refers to instances in which the help requested is limited to the amount and type needed to allow a student to solve the problem or attain the desired goal. Executive help seeking refers to instances in which the student's intention is to have someone else solve a problem or attain a goal on his or her behalf. The first form of help seeking is mastery oriented, whereas the second is dependency oriented. Studies conducted by Nelson-Le Gall and her colleagues have shown that students in need of help often do not receive it and that the majority of students view help seeking as a means of coping with problems beyond their own capabilities (Nelson-Le Gall & Glor-Scheib, 1985, 1986).

Newman (1994) developed a general model of help seeking and expanded Nelson-Le Gall's (1985) view of instrumental help seeking by defining instrumental or adaptive help seeking in the classroom in terms of a sequence of decisions and actions. If help seeking is to be adaptive or instrumental, it must (a) reflect the student's awareness of a lack of understanding; (b) involve consideration of the necessity of the request, the content of the request, and the target of the request; (c) be expressed in a manner suitable to the particular circumstance; and (d) involve processing of the help in such a manner that the possibility of success in subsequent help-seeking attempts is optimized. Studies conducted by Newman (1994) and Newman and Schwager (1993) have shown that the decision to seek help is filtered through an affective-motivational system that includes help-seeking intentions, attitudes toward both the benefits and costs of help seeking, and achievement goals. Achievement goals concern the purpose and meaning that a student ascribes to achievement behavior (Ryan & Pintrich, 1997). Students with mastery-oriented as opposed to performance-oriented goals have been found to be more likely to request help and to achieve more (Newman & Schwager, 1995). More specifically, students with mastery-oriented goals tend to view help seeking as a natural part of problem solving and focus on mastery of a goal not previously met. Students with performance-oriented goals are more concerned with avoiding demonstration of low ability, and help seeking could imply that one is lacking in a particular ability; thus, one risks loss of self-esteem by seeking help (see also Ames, 1983). In addition, research has documented that students infrequently ask questions in class, and low-achieving students—those most in need of assistance—appear to ask increasingly fewer questions as they proceed in school (Nelson-Le Gall & Glor-Scheib, 1985; Newman, 1990; Newman & Schwager, 1993).

The instructional tradition focuses on the importance of peer interaction in the social and academic growth of students. Research on helping behavior during small group work shows students working together to potentially benefit from provision and receipt of explanations (Webb & Farivar, 1994, 1999). Provision of elaborated explanations is assumed to benefit learning because the explainer must often reorganize or clarify the material and thereby understand the material more adequately. The explainer may also discover gaps in his or her knowledge or notice that his or her knowledge does not always match that of others. If these discrepancies in terms of knowledge are recognized and an attempt is made to neutralize any differences, cognitive restructuring can take place, and the explainer can develop new perspectives and knowledge (Kourilsky & Wittrock, 1992; Webb & Farivar, 1999).

Empirical results with regard to receiving and providing elaborated help are mixed. Most studies have revealed no statistically significant relations between receipt of explanations and achievement (Webb, 1991). According to Webb and Farivar (1994), however, this may be due to a failure to meet one or more of several conditions; for example, explanations must be relevant, timely, and clear. In addition, students must actually attempt to use the explanations to solve a problem.

In contrast to the inconsistent effects found for receipt of elaborated help, both provision and receipt of nonelaborated help have been shown to be detrimental or simply not related to achievement (Webb, 1989, 1991). Provision and receipt of simply the correct answer to a problem without further details on how to solve the problem does not help students correct any of their misconceptions or their lack of understanding; thus, it may have negative motivational effects and may sometimes cause students to actually stop trying to understand (Webb & Farivar, 1994). Provision of the direct answer to a problem (i.e., nonelaborated help) presumably does not call for cognitive restructuring.

In studies conducted by Mercer (1995) and Wegerif, Mercer, and Dawes (1999), students were explicitly taught the ground rules for talking together more effectively in small groups. The focus in these studies was on the use of "exploratory talk," which is a type of talk in which joint reasoning is made explicit. It involves a style of interaction that combines explicit reasoning and talk including hypotheses, arguments, challenges, and justifications (Wegerif, 1996). Use of exploratory talk can be increased through teaching, and it has been shown that this method helps groups reason together and increases students' reasoning skills (Wegerif et al., 1999).

In summary, the results of studies conducted in both the developmental-prosocial and instructional traditions show that students do not naturally develop constructive interactional patterns without instruction. Without such instruction, students' explanations appear problematic and confused; students provide few opportunities for the recipients of explanations to engage in constructive application of these explanations; and low-ability students are frequently omitted from group dynamics (Fuchs et al., 1996, 1998). As demonstrated by Webb, Troper, and Fall (1995) and Webb and Farivar (1999), however, constructive application of explanations is a strong predictor of achievement among the recipients of explanations. In addition, explicit instruction has been found to enhance students' group interactions by teaching them how to ask elaborated questions and provide elaborated answers, arguments, and justifications (Fuchs et al., 1996; King, 1994; Webb & Farivar, 1994, 1999; Wegerif et al., 1999). Given that students are more likely to obtain elaborated help when they ask for such help, the teacher-training program tested in the present study was aimed at promoting elaborated responses by encouraging students to seek and provide more elaborated help.

The Present Study

The purposes of the present study were (a) to assess whether students of teachers who participated in the teacher-training program provided more elaborations during small group work and performed better on a cooperative task than the students of teachers who did not participate in the training program and (b) to examine the relationships among provision and receipt of elaborations, performance on a cooperative task, and the affective-motivational resources of students.

Method and Instrumentation

Design

The present study followed a nonequivalent pretest-posttest control group design involving three instructional conditions: (a) a treatment group with 12 sixth-grade dyads from four primary schools using CL instruction and practices based on a 2-year staff development CL program in combination with a supplementary teacher-training program focusing on effective helping behaviors, (b) a control group with 6 sixth-grade dyads from two primary schools using CL instruction and practices based on a 1-year staff development CL program without a supplementary teacher-training program addressing effective helping behaviors, and (c) a control group with 6 sixth-grade dyads from one primary school not using CL.

Participants

Schools

The participants in the present study were originally drawn from seven primary schools located in the east and south of the Netherlands. Participation was voluntary, and all of the schools indicated an interest in the implementation of CL in their classrooms. Prior to the present study, none of the teachers in the schools had received systematic exposure to or training in CL. The schools were comparable with respect to location, size, enrollment, and interest in CL.

Four of the seven schools (Schools 1 through 4) participated in a 2-year school improvement program during the school years 1999–2001. The CL school improvement program was based on two prominent approaches, namely Johnson and Johnson's (1999) "learning together" approach and Kagan's (1994) "structural" approach. During the first and second years of implementation (school years 1999–2001), the teachers received 10 half-day training sessions on the fundamentals of CL. The training sessions were distributed throughout the 2 years and addressed such topics as the nature of CL, the teacher's role in CL, the essential components for successful CL (as defined by Johnson & Johnson, 1999), research on the use of CL, effective interaction patterns for CL, strategies for reciprocal teaching, paired reading, and promotion of social and communication skills. The design of the workshops and CL training process was guided by the recommendations of Joyce and Showers (1995) for effective training: (a) presentation of theory, (b) modeling or demonstration, (c) practice, (d) structured feedback, and (e) coaching.

The theoretical and practical principles underlying CL were presented in a teacher's manual. Cooperative activities were demonstrated by experienced trainers and via presentation of case studies. Practice was achieved via role playing and having the teachers attempt the CL activities in their classrooms. During the first year of implementation, expert coaches provided both feedback and coaching. During the second year of implementation, peers provided coaching. Krol (2004) and Krol, Veenman, and Voeten (2002) have provided detailed information on the design and conduct of the original training program.

Two of the seven schools included in the present study (Schools 5 and 6) began their participation in the 2-year-long CL school improvement program during school year 2001–2002. Finally, one school (School 7) had not yet participated in the program at the initiation of the present study, and the teachers and students thus had no experience with CL.

Teachers

Fifteen sixth-grade teachers from the four schools who participated first in the CL school program (Schools 1 through 4) also volunteered to participate in a follow-up study on students' helping behaviors. As already mentioned, a supplementary teacher-training program was developed for the present study. The design of this program was based on the assumption that the amount of learning that takes place in a group depends on how the students in the group interact. While the focus of the school improvement program was on the teacher and whether or not the teachers who participated in the program were able to implement the various CL methods, the focus of the supplementary program was on the student and whether provision and receipt of elaborated help can be stimulated, on the one hand, and can improve performance, on the other hand.

Two sixth-grade teachers from the two schools that started the school improvement program in 2001 (Schools 5 and 6) did not participate in the follow-up study because the supplementary teacher-training program built on the former program, which therefore had to be completed first. At the initiation of the present study, these teachers had 1 year of experience with CL; at the end of the study, they had 2 years of experience. One sixth-grade teacher from the school (School 7) that had not yet participated in the 2-year school improvement program at the initiation of this study, and therefore practiced whole-class instruction, also did not participate in the follow-up study.

One experimental and two control conditions were created on the basis of experience with CL and participation in the supplementary teacher-training program. The experimental group consisted of the 15 sixth-grade teachers who had participated in both the 2-year CL improvement program and the supplementary teacher-training program. Control Group 1 consisted of the 2 sixth-grade teachers who had participated at the start of the present study in the first year of the 2-year CL school improvement program but had not participated in the supplementary teacher-training program. Control Group 2 consisted of the 1 sixth-grade teacher who had not participated in either the CL school improvement program or the supplementary teacher-training program.

No statistically significant differences among the three groups of teachers were found with respect to age or length of teaching experience ($p > .05$). The average age of the teachers was 38.7 years ($SD = 10.1$), and the average amount of teaching experience was 12.8 years ($SD = 10.7$).

Students and the Formation of Dyads

A total of 48 sixth-grade students (11 and 12 years of age) participated in the present study. Six sixth-grade students were randomly selected from each of Schools 1 through 4 (treatment group) and Schools 5 and 6 (Control group 1). Twelve sixth-grade students were randomly selected from School 7 (Control Group 2). The parents of all of these students provided consent for participation. Before random selection of students, classes were divided into three levels of mathematics ability (low, medium, and high) according to students' national mathematics achievement test scores from the spring of the previous year. Within each of the classrooms included in the present study, a low-ability student was next paired randomly with a medium-ability student, and a medium-ability student was paired randomly with a high-ability student.

The decision to use dyads as opposed to three- or four-person groups was based on the following considerations: (a) Students in pairs have been shown to interact more than students in three- or four-person groups (Webb & Palincsar, 1996); (b) three- or four-person groups sometimes allow students to shirk responsibility for helping others (Webb, Ender, & Lewis; 1986); and (c) the majority of the teachers participating in the present study already grouped students in pairs, which is also a popular group size in the CL literature (Webb, Baxter, & Thompson, 1997). The formation of two-tiered abil-

Table 1
Distribution of the 24 Dyads According to Ability Level

Variable	Treatment group (Schools 1–4)	Control Group 1 (Schools 5 and 6)	Control Group 2 (School 7)	Total
Ability level	2 years of CL experience	1 year of CL experience	No CL experience	
Level 1: high-medium dyad	6	3	3	12
Level 2: medium-low dyad	6	3	3	12
Total	12	6	6	24

Note. The treatment group and Control Group 1 (6 schools) were composed of 6 sixth-grade students per school divided into 18 dyads; Control Group 2 (1 school) was composed of 12 sixth-grade students divided in 6 dyads. The treatment group also followed the supplementary teacher-training program. CL = cooperative learning.

ity groups (i.e., a high/medium-ability group and a medium/low-ability group) was based on the assumption that the ability levels of the participating students should differ to a sufficient degree to generate helping behaviors but not so much as to allow the students to work within their zones of "proximal development" (Vygotsky, 1978). Research from a Vygotskian perspective indicates the importance of pairing students with slightly more expert peers (O'Malley, 1995). In addition, medium-ability students have been found to miss numerous opportunities to construct explanations when working with high-ability peers in four-tiered ability groups (i.e., groups with one high-ability student, two medium-ability students, and one low-ability student) (Webb & Palincsar, 1996).

Pairings were subsequently checked by the teachers to exclude dyads of students who were not able to get along with each other. The distribution of the 24 dyads according to group and ability level is shown in Table 1.

Supplementary Teacher-Training Program

On the basis of the work of Dawes, Mercer, and Wegerif (2000); Farivar and Webb (1991); King (1994, 1999); Nelson-Le Gall (1981, 1992); and Webb and Farivar (1994, 1999), we developed a supplemental training program to help teachers promote students' ability to work effectively in small groups. The program involved three sets of activities: (a) activities to develop and practice conversational ground rules, (b) activities to develop help-seeking behavior in the classroom, and (c) activities to develop help-giving behavior in the classroom. The activities were clustered in sequential sets in order to address only a few skills at a time, and the activities proceeded from basic to advanced to allow the students to build on previously learned skills.

The activities designed to develop and practice conversational ground rules addressed such skills as listening, sharing information, cooperating, making arguments for and against different points of view, critically discussing alternative ideas, giving and asking for reasons, reaching agreement before making a final decision, and seeing that all members of the group are invited to contribute. While a number of these skills were addressed in the CL school improvement program, students were invited to design their own ground rules for effective conversation.

The activities focusing on developing help-seeking behavior in the classroom were intended to help students learn how to ask for help when needed. These activities involved a number of help-seeking behaviors based on Nelson-Le Gall's (1981, 1985) model of the help-seeking process and the instructional program of Farivar and Webb (1998, p. 173): "(a) recognize that you need help; (b) decide to get help from another student; (c) choose someone to help you; (d) ask for help; (e) ask clear and precise questions; and (f) keep asking until you understand." The students were also presented with a set of thought-provoking stem questions and items (e.g., "What does . . . mean?" "Explain why. . . ." and "Explain how. . . ."). Such guided cooperative questioning, developed by King (1994, 1999), is intended to elicit elaborated explanations, arguments, inferences, and justifications.

The activities involving development of help-giving behavior in the classroom were intended to help students learn how to provide elaborated explanations. These activities focused on different kinds of help and emphasized provision of extended explanations as opposed to shorter, less informative responses or nonelaborated help. In guided discussions with the class, the teacher thus creates a chart of productive help-giving behaviors such as the following: "(a) when someone asks for help, help him or her; (b) be a good listener; (c) give explanations instead of the answer; (d) watch how your teammate solves the problem; (e) give specific feedback on how your teammate solved the problem; (f) check for understanding; and (g) praise your teammate for doing a good job" (Farivar & Webb, 1998, p. 174).

The activities covered in the training program were first presented in seven teacher newsletters containing three to four exercises for practice with the students during the 2 weeks following receipt of the newsletter. The purpose and rationale behind the exercises, the materials needed, the amount of time needed, and concrete suggestions for the conduct of the exercises were also included in the newsletters. At the beginning of each exercise, the teacher was expected to introduce the entire class to the problem for the day, remind the class of the agreed-upon "ground rules for talk," and tell the students that they should consult each other first before asking the teacher for help. The students worked in small groups, and the teacher distributed prompt cards with the various help-seeking and help-giving behaviors previously discussed. The teachers also circulated among the groups, observed the groups' interactions, and answered questions as necessary during the exercises. At the end of the class period, each group presented its conclusions to the class and discussed how well the group had worked together.

We present one example as an illustration of the nature of the exercises. In the exercise "Using your mobile phone" (Newsletter 6), the dyads were asked to solve such problems as the following. "Find the difference in costs of two telephone companies for a 27-minute telephone call to 3611431 at 14.06 hours and at 20.15 hours. Company Flex charges for fixed phones off-peak €0.07 and peak €0.10; for mobile phones off-peak €0.25 and peak €0.30 (basic rate of €0.15). Company Telbase charges for fixed phones off-peak €0.15 and peak €0.15; for mobile phones off-peak €0.15 and peak €0.15 (basic rate of €0.00)." Each student was also given a help-asking chart and a help-giving chart of behaviors and examples to help him or her solve the relevant problem (adapted from Webb, Farivar, & Mastergeorge, 2002).

After the conduct of the exercises presented in the seven newsletters, the teachers participated in two 3-hour workshops. In the first workshop, time was taken to discuss the experiences of the teachers in implementing the exercises presented in the newsletters. Then the theoretical background underlying the activities included in the newsletters was briefly reviewed. Next, the teachers were shown video fragments with examples of "effective" and "ineffective" helping and asked to code these fragments as a form of practice in distinguishing between effective and ineffective helping behaviors. Finally, the teachers were asked to video or audio record the cooperative interactions of a pair of students from their classrooms and discuss these interactions with their students in terms of effective and ineffective helping; the results of this assignment constituted the main topic for the next workshop.

In the second workshop, the teachers presented their videotapes or audiotapes of a student dyad solving a math problem to allow discussion of the demonstrated helping behaviors. The guiding question for the discussion was "What went well and what can be improved next time?" The teachers then broke into small groups to role-play and discuss the contrasting roles of teachers in monitoring students working in small groups. The role-playing exercises were based on the outcomes of a study conducted by Webb and Farivar (1994) showing a relation between teachers' instructional styles and students' verbal interactions. At the end of the second workshop, the teachers were also asked to complete a brief questionnaire to evaluate the newsletters and workshops.

Fidelity of Treatment

The fidelity of the implementation of the 2-year CL staff development program that constituted part of the experimental condition was checked by observing the teachers in their classrooms on three occasions at approximately 11-month intervals; questionnaires were also administered to the teachers and students (for a full description of the CL staff development program and implementation results, see Krol, 2004). The results of the implementation study showed that the teachers were able to implement those components required for successful CL (e.g., structure positive interdependence, foster individual accountability, and stimulate social skills) (see also Johnson & Johnson, 1999). The students worked in small groups for 1 hour a day, three times a week.

The accuracy with which the participating teachers implemented the exercises presented in the newsletters for the supplementary training program was assessed via direct observation and regular telephone interviews with the teachers. Specifically, two of the authors visited all of the classrooms from which the dyads were selected to observe one of the exercises from the newsletters, and the teachers were interviewed by telephone to ascertain their reactions to the exercises after the second, fourth, and sixth newsletters. In general, it was found that the teachers conducted the exercises as intended.

The fidelity of the implementation of the CL staff development program by Control Group 1 was checked by interviewing the participating teachers, who also had their students work in small groups for 1 hour a day, three times a week. The teacher in Control Group 2 received no special training. The students worked individually and were instructed in a whole-class format.

Procedure

Before and after participation of the treatment teachers in the supplementary training program, all of the student dyads were asked to cooperatively solve a math task. At pretest, Version A of the math task was used; at posttest, Version B was used. All of the sessions took place in the morning. The researcher brought two students from a classroom (i.e., one of the 24 dyads) into a room where the materials were already set up. According to a standardized procedure, a brief description of the task and how the answers should be recorded was then provided, and it was explicitly stated that the students should cooperate to solve the task within 30 minutes. All of the sessions were video and audio recorded and later transcribed.

After completion of the math task at pretest, all of the students were also administered a questionnaire addressing their help-seeking intentions, their attitudes toward help seeking and help giving in the classroom, and the nature of their achievement goals. The math task, coding of transcripts, and the student questionnaire are described in the sections to follow. In light of recent recommendations (e.g., Wilkinson & American Psychological Association Task Force on Statistical Inference, 1999), and to help better interpret the importance of the effects found in the present study, we report reliability coefficients for scores achieved on the instruments assessed.

CL Math Task

The math task, which required formal reasoning and discussion, was developed to be challenging for sixth graders and did not include topics discussed previously in the classroom. During the development phase, a sample of two dyads from schools not involved in the study provided feedback on the adequacy of the materials.

Students were presented a 15-page booklet with 15 balance-beam problems to solve cooperatively. That is, the students had to predict which side of the beam would go up or down when various configurations of weights and distances were set up. In contrast to the work of Siegler (1976), Phelps and Damon (1989), and Tudge (1992), our study did not involve a real balance beam with removable pegs on each side of the fulcrum or the capacity to actually tip left or right. Rather, we used a paper-and-pencil task with drawings of a balance beam involving different configurations of weights and distances from the fulcrum (see also Ros, 1994).

The first five problems pictured a scale with weights on it, and the students had to indicate whether the scale was balanced or which side would go up or down. These five problems were the simplest and involved basic weight and distance problems that have been found to be fairly easy for most sixth graders (Phelps & Damon, 1989). The booklet also provided feedback on the solutions to the first five problems, which allowed the students to compare their solutions with the ones in the booklet. These problems highlighted the importance of different weights and different distances from the fulcrum and thereby familiarized the students with how to work with a balance beam. After completing the problems, the students were asked to cooperatively solve 10 more problems with the weights and distances from the fulcrum varied in a more complicated manner. The solutions to these problems required more formal reasoning, in that the students had to consider weight and distance from the fulcrum at the same time (Phelps & Damon, 1989). Each dyad received one booklet and one worksheet. After discussion of a problem, one of the students wrote the solution down.

As a means of assessing the overall difficulty of the math task, a pilot version of the task was administered to three sixth-grade classes from three schools not involved in the study. The teachers indicated the ability levels of the students, and a total of 30 dyads were formed with a high-ability student and a medium-ability student or a medium-ability student and a low-ability student. Two versions of the math task were constructed: Pilot Version A and Pilot Version B. Half of the dyads were then asked to solve Version A followed by Version B, while the other half were asked to solve Version B followed by Version A. Results showed that overall item difficulties were .53 (SD = .25) for Version A and .57 (SD = .24) for Version B. Difficulty ratings per item ranged from .04 to 1.00 for Version A and from .07 to .96 for Version B. Cronbach reliability coefficients were .71 and .73 for Versions A and B, respectively. The parallel-form reliability estimate (i.e., the Pearson correlation between the two versions) was .75. Thus, the pilot data showed Versions A and B of the math task to be largely equivalent with respect to item difficulty and reliability.

The total performance score for the math task was the sum of the points awarded for the final 10 balance beam problems. A correct answer was assigned 1 point, meaning that a maximum score of 10 points was possible.

Verbal Interaction Categories

Coding Scheme

The coding scheme used to analyze the verbal interactions of the students during the math task involved six dimensions pertaining to help seeking, help giving, constructive activities, procedural statements, affective statements, and noncontent. These dimensions were based on the work of King (1999); Nelson-Le Gall (1981, 1985, 1992); Newman and Schwager (1995); Mercer, Wegerif, and Dawes (1999); Puustinen (1998); and Webb and Farivar (1994, 1999).

The help-seeking dimension refers to the manner in which a student approaches and processes the help-seeking process. The help-giving dimension refers to the manner in which students provide help, especially to the level of elaboration provided in the work group. Constructive activities concern the behaviors that students engage in after receiving help or their construction of self-explanations after raising a problem. The procedural dimension refers to statements intended to help regulate the conduct of a task. The affective dimension refers to any emotional remarks made concerning the collaboration and contributions of those involved. Given that the success of working together may depend on not only cognitive but also affective factors, the coding scheme covered affective elements as well. Noncontent refers to noninformational responses or off-task statements. The coding scheme, summarized in Table 2, contained a total of 19 verbal interaction categories. All of the utterances included in the transcript were coded according to one of these categories.

Coding of Transcripts

To code the verbal interaction of the dyads, the interaction was first divided into conversational turns, defined as a change of speaker, and then into utterances, distinguished by a "perceptible pause," comma, or period in the transcript and a singular communicative function (Van Boxtel, van der Linden, & Kanselaar, 2000). The utterance constituted the unit of analysis, and although the length of an utterance can vary from a single word (e.g., "No") to an extended monologue, each utterance was assigned to only one of the categories within the coding scheme. The scores for a given student were the numbers of utterances falling into the different categories within the coding scheme.

Before the coding of the transcripts, two researchers completed a training program of about 40 hours. The training program involved formulation of rules for coding, instruction in applying the Multiple Episode Protocol Analysis computer program (developed by Erkens, 2002) to code transcribed verbal interactions, and practice with the coding scheme. The transcripts for the 48 dyadic interactions from the pretests and posttests were randomly assigned to one of the researchers for coding. Only the interactions for the final 10 problems were transcribed; the first 5 problems served as practice.

Interrater agreement for the coding of the transcripts, based on nine transcripts randomly selected from the overall pool (i.e., 25% of all transcripts), was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100, as well as by using Cohen's (1960) kappa value. The interrater agreement across the 19 categories was 88% (median = 86%, range = 67%–96%); Cohen's kappa was .86 (median = .84, range = .66–.95). Item-specific kappa values are presented in Table 2. According to Fleiss's (1981) general benchmark, a kappa value between .40 and .75 can be interpreted as intermediate to good, and a value above .75 can be considered excellent.

Cooperative Learning in the Classroom Questionnaire

The Cooperative Learning in the Classroom Questionnaire (CLCQ) contains a total of 25 items divided across four sections addressing (a) student intentions to seek help in the classroom, (b) student attitudes toward help seeking in the classroom (i.e., benefits and costs of help seeking), (c) student attitudes toward help giving in the classroom, and (d) the nature of students' achievement goals. Items are rated on a 5-point scale ranging from *highly disagree* (1) to *highly agree* (5).

Intentions to Seek Help

The first section of the CLCQ contains four items designed to measure students' intentions to seek help in the classroom. The items in this section were adopted from Newman (1990) but adjusted to address the intention to seek help from a peer as opposed to the teacher when having difficulty (e.g., "When I don't know how to do a problem, I ask help from a classmate"). The alpha coefficient for this section was found to be .79.

Attitudes Toward Seeking Help

The second section of the CLCQ contains eight items designed to measure students' attitudes about seeking help in the classroom. These items were adopted from Newman (1990) and Newman and Schwager (1993). Two items concern the potential *benefits of help seeking* (e.g., "I think that asking questions helps me to learn better"), and five items concern the potential *costs of help seeking* (e.g., "I feel shy about asking questions"). Cronbach alpha coefficients were .66 for benefit scores and .76 for cost scores.

Attitudes Toward Giving Help

The third section of the CLCQ contains four items designed to measure students' attitudes toward giving help in the classroom. These items, based on the work of Webb and Farivar (1994, 1999), concern the willingness of a student to help or explain something to a fellow student (e.g., "When I explain something to another child, I understand the subject better myself"). The alpha coefficient was .45.

Achievement Goals

The fourth and final section of the CLCQ contains nine items designed to assess students' achievement goals for mathematics. The items were adopted from Newman (1998) and Newman and Schwager (1993). Three items tap *mastery goals* (e.g., "I like to learn as much math as possible" or "I like to

Table 2
Coding Scheme for Verbal Interactions

Dimension	Category	Description	K
1. Help seeking	1a. Instrumental (INS)	Requesting an explanation of process (how to obtain the solution), not an answer (asking for little hints, e.g., "Why did you multiply it by 4"?)	.78
	1b. Executive (EXE)	Requesting the answer only (asking for big hints, e.g., "What is the outcome for problem 10?")	.82
	1c. Confirmatory (CON)	Requesting confirmation of the suggested solution (e.g., "Is this correct?")	.89
2. Help giving	2a. Help offering (HOF)	Offering or probing for help (e.g., "Shall I explain it to you?")	.86
	2b. Labeled explanation (EX+)	Providing arguments or justifications via an extended explanation (e.g., "It will balance that way because, here, exactly in the middle, are 3 kilograms and there are 6.")	.78
	2c. Unlabeled explanation (EX–)	Providing the answer without further argument or justification (e.g., "Anyway, I think it is 3.")	.88
3. Constructive activities	3a. Challenging help: labeled explanation (CEX+)	Raisings doubts about the help received by specifying reasons ("To achieve balance? I don't know. That's 9 kilograms on this side, and this is more! How did you get 12 kilograms?")	.75
	3b. Challenging help: unlabeled explanation (CEX–)	Raising doubts about the help received without specifying reasons (e.g., "Well, mine doesn't have that, I have 9 kilograms.")	.79
	3c. Acknowledging help: labeled explanation (AEX+)	Acknowledging the help received by explaining how it was applicable (e.g., "Yes, I see now because it's at the end of the bar.")	.66
	3d. Acknowledging help: unlabeled explanation (AEX–)	Acknowledging the help received without explaining why it was applicable (e.g., "Yes, okay, I think so too.")	.95
	3e. Self-questioning: labeled explanation (SQU+)	Producing arguments out loud, thereby showing how to solve the problem (e.g., "That's a bit strange. Let's see. Should these weights go all the way to the back or not? It's 4, because 4 makes it equally long.")	.66

(continued)

Table 2 (Continued)

Dimension	Category	Description	*K*
	3f. Self-questioning: unlabeled explanation (SQU–)	Producing arguments out loud without the intention of sharing the information (e.g., "Uhm . . . I think the fulcrum must be here. . . . Or here? . . . No, here!")	.84
	3g. Metacognitive statements (META)	Monitoring and testing the problem-solving process (e.g., "Shall we skip this problem and go on to the next one, and then come back to this one later?")	.86
	3h. Evaluative statements (EVA)	Evaluating the problem-solving process or providing a short summary (e.g., "Now, we've finished it and I understand it completely.")	.67
	3i. References to earlier-discussed ideas (REF)	Raising previously mentioned solutions (e.g., "Like problem 12, it is the same, the same distance.")	.86
4. Procedural	4a. Task-process responses (PROC)	Regulating the conduct of the task (e.g., "Sorry, I was looking at the wrong problem.")	.92
5. Affective	5a. Positive statements (APOS)	Providing positive remarks about the process of cooperation (e.g., "Now, that's good!")	.92
	5b. Negative statements (ANEG)	Providing negative remarks about the process of cooperation (e.g., "Boy, you sure took your time!")	.86
6. Noncontent	6a. Noninformational (NON)	Making off-task statements (e.g., "I dropped my pencil.")	.67

Note. Category codes are in parentheses. *K* = Cohen's kappa.

learn new things in math"); three items tap *performance goals* (e.g., "I want other kids to think I am smart in math"); and three items tap *social goals* (e.g., "I like to talk with other kids about math" or "I like to help other kids with their math work"). Research on help seeking shows that achievement goals influence help seeking in the classroom (Nelson-Le Gall, 1981, 1992; Newman & Schwager, 1995). Alpha coefficients were .79 for scores on the mastery goals scale, .72 for scores on the performance goals scale, and .36 for scores on the social goals scale. Given the low reliability coefficients for the scores on the social goals scale and the scale concerning attitudes of students toward giving help, these scales were omitted from any further analyses.

Data Analysis

Within the help-seeking, help-giving, and constructive activity dimensions of the verbal coding scheme, two levels of elaboration were identified to characterize the contributions of students to verbal interactions: high-level elaboration and low-level elaboration. This classification was indirectly based on the work of Webb, Nemer, Chizhik, and Sugrue (1998). With regard to the help-seeking dimension, the instrumental help category was taken to represent high-level elaboration, because the intention is to elicit elaborated help in the form of small hints. The executive help and confirmatory help-seeking categories were taken to represent low-level elaboration, because the intention is to either obtain the answer directly or confirm the correctness of one's work.

With regard to the help-giving dimension, the help giving with labeled explanations and probing the other for help to elicit elaborated explanations categories were taken to represent high-level elaboration. The help giving without labeled explanations category was taken to represent low-level elaboration.

With respect to the constructive activity dimension, the following five categories were taken to represent high-level elaboration: (a) challenging help received with labeled explanations, (b) acknowledging help with labeled explanations, (c) self-questioning/answering with labeled explanations, (d) metacognitive statements, and (e) references to previously discussed material. Four categories were taken to represent low-level elaboration: (a) challenging help received without labeled explanations, (b) acknowledging help received without labeled explanations, (c) self-questioning/answering without labeled explanations, and (d) evaluative statements. The procedural, affective, and noncontent dimensions of the verbal coding scheme were not considered at this point in the analyses.

The unit of analysis for the students' verbal interactions was the dyad. This unit of analysis was adopted because the knowledge building that occurs during dyadic interactions can be viewed as largely interdependent; that is, the questions and responses of one partner are elicited and stimulated to a great extent by the questions, statements, and responses of the other partner (King, 1994). For each dyad, the frequencies of high- and low-level elaborations were calculated by summing the individual code frequencies. The percentage of high-level (vs. low-level) elaborations was then calculated for each dyad by dividing the number of high-level (vs. low-level) utterances by the total number of utterances. Percentages were used because the absolute number of utterances varied widely across the dyads, from 41 to 243 at pretest ($M = 82.92$, $SD = 45.01$) and 26 to 209 at posttest ($M = 74.79$, $SD = 42.08$). A per-dyad composite score for both high- and low-level elaborations was also computed by averaging the frequencies for the eight categories of high-level elaboration and the seven categories of low-level elaboration separately.

For the math test, the unit of analysis was again the dyad, in that the math problems were solved jointly with resolutions written down on a single worksheet. For the CLCQ data, the unit of analysis was the individual student, because each student individually completed the questionnaire.

Analyses of variance were conducted to test for initial differences between the treatment group and the two control groups. Analyses of covariance (ANCOVAs) were used to test for differences between the treatment and control groups at posttest, with pretest scores used as the covariates. ANCOVA was considered suitable for correcting the posttest mean scores for existing pretest differences as well as for reducing the amount of unexplained variance in the posttest scores, which, according to Maxwell and Delaney (1990) and Onwuegbuzie and Daniel (2003), leads to an increase in the power of statistical tests.

For each ANCOVA, the assumption of homogeneity of regression within groups was tested. Simple contrasts were used to test the specific differences between the treatment group and Control Group 1 and between the treatment group and Control Group 2. For each contrast, effect sizes were computed with Cohen's (1988) d statistic. Olejnik and Algina (2000) suggested that, when ANCOVA is used, the means in the contrast be estimated with adjusted cell means rather than posttest cell means. The effect sizes in the present study were therefore computed by dividing the difference between the adjusted cell means in the contrast by the pooled standard deviation for all of the cells in the between-subjects design (Olejnik & Algina, 2000).

Results

Analysis of Verbal Interactions

Table 3 presents an overview of the interaction variables for the treatment and control groups. The pretest scores show that the dyads in both the treatment and control groups provided and received very few elaborations during the processes of help seeking, help giving, and constructive activities.

As a result of the low statistical power arising from small sample sizes, not all of the differences in the verbal interaction scores between the treatment and control groups were tested for statistical significance. ANCOVAs were conducted only on the composite scores for high and low levels of elaboration.

The tests for homogeneity of regression (i.e., interaction effects between the covariate and the independent variable) were nonsignificant, $F(2, 18) = 1.14$, $p > .05$, for high-level elaborations and $F(2, 18) = 2.95$, $p > .05$, for low-level elaborations. These results show that the assumptions for the ANCOVAs were not violated.

Comparison of the treatment and control groups showed no statistically significant differences at pretest for the composite variables reflecting the total percentages of high- and low-level elaborations, $F(2, 21) = 1.14$, $p > .05$, and $F(2, 21) = 2.08$, $p > .05$, respectively. The lack of differences between the groups at pretest shows that participation in an earlier version of the CL training program (in the case of the treatment group and Control Group 1) had no great impact on the elaborations of the dyads under study. Although there were no statistically significant pretest differences between the treatment and two control groups in terms of the composite variables reflecting total percentages of high- and low-level elaborations, ANCOVAs were performed to reduce the posttest error variances and increase statistical power (Maxwell & Delaney, 1990).

The results displayed in Table 4 show a statistically significant treatment effect for the use of high-level elaborated utterances, with the dyads in the

Table 3
Descriptive Statistics for Interaction Variables (Percentages): Treatment and Control Groups

	Pretest						Posttest					
	Treatment group ($n = 12$)		Control Group 1 ($n = 6$)		Control Group 2 ($n = 6$)		Treatment group ($n = 12$)		Control Group 1 ($n = 6$)		Control Group 2 ($n = 6$)	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Help seeking												
High-level elaboration	0.57	0.80	2.57	0.92	2.18	1.65	0.57	1.20	2.65	2.46	1.15	1.26
Low-level elaboration	8.41	3.91	15.07	6.18	7.50	4.23	6.24	3.68	12.54	6.72	8.15	5.85
Help giving												
High-level elaboration	10.87	6.08	4.00	3.61	4.24	3.25	10.09	6.07	2.79	5.19	2.14	1.09
Low-level elaboration	19.73	8.53	22.20	7.14	24.72	8.35	23.43	6.06	28.62	8.32	23.57	8.86
Constructive activity												
High-level elaboration	9.81	6.29	9.23	3.93	8.09	4.35	6.56	4.78	1.08	2.63	4.01	2.48
Low-level elaboration	41.69	4.95	38.66	1.77	46.94	10.26	48.10	6.68	46.64	3.66	47.41	3.76
Other												
Noncognitive utterances	8.32	5.16	8.08	4.59	5.57	3.79	4.59	3.24	5.37	6.79	12.50	7.82
Non-task-related utterances	0.61	0.92	0.20	0.50	0.77	1.20	0.23	0.54	0.32	0.78	1.08	1.28
Composite scores												
Total high-level elaboration	21.26	11.74	15.80	7.50	14.51	7.81	17.21	7.47	6.52	8.14	7.30	3.73
Total low-level elaboration	69.82	11.43	75.92	5.02	79.15	8.94	77.77	6.15	87.79	7.63	79.13	7.61

Table 4

Contrast Tests for Differences Between Mean Percentages of High-Level and Low-Level Elaborations Among the Treatment and Control Groups Controlled for Pretest Scores

Contrast	Contrast value (based on adjusted means)	*SD* (pooled)	$t(21)$	*d*
Total high-level elaboration				
Treatment group vs. Control Group 1	8.40	6.95	2.88*	1.21
Treatment group vs. Control Group 2	7.08	6.95	2.39*	1.02
Total low-level elaboration				
Treatment group vs. Control Group 1	−9.19	6.89	−2.56*	−1.33
Treatment group vs. Control Group 2	−0.08	6.89	−0.02	−0.01

*$p < .05$.

treatment group producing more high-level elaborations than the dyads in Control Group 1 and Control Group 2 (ps < .05, $d = 1.21$ and $d = 1.02$, respectively). The differences among the three groups are displayed in Figure 1, which shows that the treatment effect was due to a smaller posttest decrease in the use of high-level elaborations by the treatment group than by the two control groups. At the same time, the dyads in the treatment group produced significantly fewer low-level elaborations than the dyads in Control Group 1

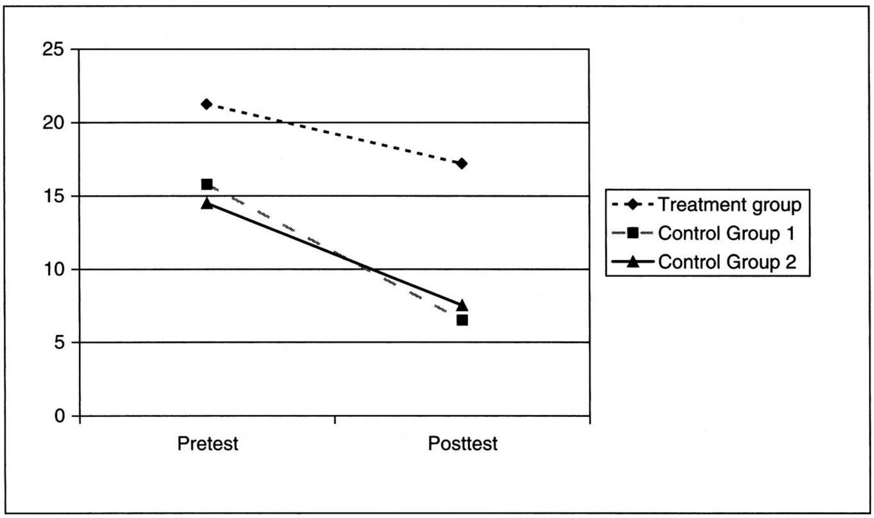

Figure 1. **Graphic display of differences between the treatment and control groups in provision and receipt of high-level elaborations.**

($p < .05$, $d = -1.33$), while the difference in the adjusted mean percentages of low-level elaborations for the dyads in the treatment group versus Control Group 2 was not statistically significant ($p > .05$, $d = -0.01$). In summary, with the exception of the differences in use of low-level elaborations between the treatment group and Control Group 2, the differences regarding total use of high- and low-level elaborations were in the expected directions and signify a treatment effect.

The two excerpts displayed in Figures 2 and 3 show the use of high-level elaborations by a treatment dyad (Figure 2) and a Control Group 1 dyad (Figure 3). These excerpts were selected because the overall use of high-level elaborations reflected the relevant group mean. The most striking difference between the two excerpts is the difference in the use of high-level elaborations

Line	Student and comment	Category code
1	Kim: [reads problem 7 aloud]	
2	Bart: Now, it is in the. . . . The beam is divided into three sections.	EX–
3	Kim:Yeah.	AEX–
4	Bart: So one such section is . . .	EX–
5	Kim: Umm. This is 4 and then there are two more sections, so this would be 8, right?	CON
	You then have twice as much. //	EX+
6	Bart: Umm. Yeah, that is possible. Yeah, that is right. 'Cause that should, yeah, that should have more weight on the long side, 'cause this side always has to have less weight.	AEX+
7	Kim: Yeah, 'cause that then . . . let's see . . . one-third is . . . and there's, there are two sections left and then it should be twice as much . . .	AEX–
8	Bart: Uh-huh, yeah.	AEX–
9	Kim: Do you understand?	HOF
10	Bart: Yeah, I get it but I don't know if 8 is the / or 8kilos . . . /	LOU
11	Kim: 'Cause there were just . . .	EX–
12	Bart: Maybe it's too much, but . . .	EX–
13	Kim: But there were three by . . . by the one just now . . .	REF
14	Bart: Oh, yeah, you're right. Look, there are two sections.	AEX–
15	Kim: Yeah.	AEX–
16	Bart: Yeah, it's right.	AEX–

Figure 2. **Use of high-level elaborations by a treatment group dyad. Bart is a high-ability male student; Kim is a medium-ability female student (to preserve the anonymity of the students, their names have been changed). The correct answer for this problem is 8 (Version B of the math task). / = pause shorter than 3 seconds; // = pause longer than 3 seconds. See Table 2 for descriptions of category codes.**

Line	Student and comment	Category code
1	Sam: [reading Problem 7 aloud] Uh . . . 8.	EX–
2	Lindy: 2.	EX–
3	Sam: 8, isn't it?	CON
	This should be twice as much?	CON
4	Lindy: Yeah, 8. //	AEX–

Figure 3. **Use of high-level elaborations by a Control Group 1 dyad. Lindy is a high-ability female student; Sam is a medium-ability male student (to preserve the anonymity of the students, their names have been changed). The correct answer for this problem is 8 (Version B of the math task). // = pause longer than 3 seconds. See Table 2 for descriptions of category codes.**

and the time needed by the dyads to solve the problem. The students in the treatment group produced more elaborations than the students in the control group. The students in the treatment group thought aloud, verbalized their thoughts, and provided each other with extended elaborations (e.g., lines 5, 6, and 7). An answer was provided only when both students understood the rule and were in agreement on its use (e.g., lines 9 and 14). Explanations of rules and justifications were exemplified by the use of such words as "because," "so," and "thus." The rules identified for previous solutions were also applied.

The students in the control group produced primarily nonelaborated explanations, with no attempt to understand each other. In addition, as can be seen from Figure 3, the control dyad needed only four utterances to solve the problem.

Performance Scores

Table 5 presents performance results for the math task (with a maximum possible score of 10). Comparison of the treatment group with the control groups prior to treatment showed no statistically significant differences in performance scores or time needed by the dyads to complete the math task. Tests for homogeneity of regression were nonsignificant in the case of both math performance, $F(2, 45) = 0.37$, $p > .05$, and number of minutes used to solve the task, $F(2, 45) = 2.59$, $p > .05$. Thus, the assumptions for ANCOVA were not violated. Simple contrast tests for differences at posttest showed no statistically significant differences between the treatment and control groups (see Table 6). The results nevertheless showed a moderate effect size in favor of the treatment group relative to Control Group 2 ($d = 0.76$).

Inspection of the data in Table 5 shows that dyads with CL experience scored higher at posttest than dyads without CL experience. No statistically significant differences were found for the number of minutes the dyads in

Table 5

Descriptive Statistics for Performance Scores and Number of Minutes Needed to Solve the Math Task

	Pretest						Posttest					
	Treatment group ($n = 12$)		Control Group 1 ($n = 6$)		Control Group 2 ($n = 6$)		Treatment group ($n = 12$)		Control Group 1 ($n = 6$)		Control Group 2 ($n = 6$)	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Performance scores	6.33	2.23	6.00	2.28	6.67	0.52	6.17	2.08	6.33	1.97	4.67	1.75
Number of minutes used	8.58	3.90	7.17	2.14	10.67	5.20	6.67	2.06	5.00	1.67	8.17	2.93

Table 6
Contrast Tests for Differences Between Mean Performance Scores and Minutes Needed to Solve the Math Task Among the Treatment and Control Groups Controlled for Pretest Scores

Contrast	Contrast value (based on adjusted means)	*SD* (pooled)	*t*(21)	*d*
Performance scores				
Treatment group vs. Control Group 1	−0.17	1.98	−0.17	−0.09
Treatment group vs. Control Group 2	1.50	1.98	1.52	0.76
Number of minutes needed to solve math task				
Treatment group vs. Control Group 1	1.67	2.22	1.50	0.75
Treatment group vs. Control Group 2	−1.50	2.22	−1.35	−0.68

the three groups needed to solve the mathematics task, although the results showed relatively large effect sizes indicating that (a) the dyads in the treatment group needed more time to solve the task than the dyads in Control Group 1 and (b) the dyads in Control Group 2 needed more time than the dyads in the treatment group. In general, the dyads required an average of 8.8 minutes ($SD = 3.93$) to complete the math task at pretest and an average of 6.6 minutes ($SD = 2.39$) to complete the task at posttest.

Variance in Levels of Elaboration

To determine the extent to which help-seeking intentions, estimates of the benefits and costs of help seeking, and the achievement goals of students (affective-motivational resources) contributed to the variance observed in levels of elaboration (indexed here as the composite scores for the total number of high- versus low-elaborated utterances), we conducted regression analyses in which all of the variables were entered simultaneously. Standardized regression weights (betas), R^2 values, and adjusted R^2 values are shown in Table 7. The analyses were supplemented with structure coefficients (r_s), which are not suppressed or inflated by collinearity (Courville & Thompson, 2001; Thompson & Borrello, 1985).

Table 7 also includes the descriptive statistics for the affective-motivational resources. The mean scores show that the dyads in both the treatment and control groups scored relatively high on mastery-oriented goals and help-seeking intentions and relatively low on performance-oriented goals. In addition, the results presented in Table 7 show that the different variables considered together accounted for a considerable amount of the variance in the total scores for high-level elaborations (47%; $p < .05$) and only a small amount of the variance in the total scores for low-level elaborations (17%; $p > .05$). The negative beta coefficients for mastery goals ($\beta = -.46$) and

Table 7

Descriptive Statistics for Affective-Motivational Resources, Standardized Regression Weights (β), and Structure Coefficients (r_s) for Predictors of Levels of Elaboration and Math Performance at Posttest (N = 48)

			Total high-level elaborations		Total low-level elaborations		Math performance	
Variable	*M*	*SD*	β	r_s	β	r_s	β	r_s
Help seeking								
Intentions to seek help	3.64	0.92	−.00	.02	−.16	−.38	−.19	−.15
Benefits of help seeking	3.57	1.12	.21	.10	−.04	−.11	.28	.21
Costs of help seeking	4.37	0.78	−.15	−.31	.08	.11	.08	.07
Achievement goals								
Mastery goals	4.26	0.76	−.46*	−.78	.19	.53	−.35*	−.57
Performance goals	2.47	0.99	−.42*	−.62	.30	.77	−.40*	−.74
R^2				.47*	.17		.34*	
Adjusted R^2				.41	.07		.27	

*p < .05.

performance goals (β = −.42) showed that students with more mastery- or performance-oriented goals produced fewer high-level elaborations. That is, a mastery or performance orientation negatively affected provision of high-level elaborations.

Variance in Math Performance

To determine the extent to which the aforementioned affective-motivational resources contributed to the variance observed in the math performance of the students, we again conducted regression analyses. Beta coefficients, R^2 values, and structure coefficients are shown in Table 7. As can be seen, 34% of the variance in math performance scores was explained by the different variables ($p < .05$). The statistically significant negative beta coefficients for mastery and performance goals (−.35 and −.40, respectively) show that students with higher mastery- and performance-oriented goals scored lower on the math task.

Levels of Elaboration and Math Performance

Finally, to determine the extent to which the use of high-level elaborations affected math performance scores, we once again conducted regression analyses. The use of high-level elaborations accounted for a statistically significant amount of the variance in math performance scores ($R^2 = .35$, $p < .05$, adjusted $R^2 = .33$). The positive beta coefficient (.59) showed the use of high-level elaborations to foster higher achievement.

Discussion

The present study was designed to examine the effects of a supplementary teacher-training program on the elaborations of students during help seeking and help giving while working in dyads, the effects of provision and receipt of elaborations on student performance, and the relationships among affective-motivational resources, student elaborations, and math performance. In general, positive effects on provision and receipt of elaborations by the treatment dyads were found. A statistically significant treatment effect was found for the composite variable concerning total percentage of high-level elaborations, with the treatment dyads providing more high-level elaborations than the control dyads (treatment group vs. Control Group 1, $d = 1.21$; treatment group vs. Control Group 2, $d = 1.02$).

To evaluate the practical significance of the effect sizes observed, it is not sufficient to simply apply Cohen's tentative benchmarks for what might be deemed small, medium, and large effects (Thompson, 2002). Thus, we compare our effects with those found in related domains of interest.

Hattie, Biggs, and Purdie (1996) outlined a measurement procedure designed to determine the typical effect of most education interventions. On the basis of a synthesis of 394 meta-analyses (including 40,567 studies), they ascertained that an effect size of 0.40 was a benchmark from which various educational interventions could be interpreted. That is, across the 394 meta-analyses, the typical effect size for educational interventions was 0.40 ($SD = 0.26$). In the case of interventions focusing on mathematics, the typical effect size was 0.32. A meta-analytic review of 90 group comparison design studies evaluating classroom-based peer-assisted learning interventions (Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003) revealed positive overall effect sizes indicating increases in achievement (weighted $d = 0.33$). For mathematics, the weighted d value was 0.22 (based on 33 peer-assisted learning studies).

These are, of course, global benchmarks, and more refined comparisons can be made with interventions regarding academic helping behaviors within the context of mathematics instruction, the topic of the present study. For this purpose, we computed effect sizes (as Cohen's d values) from the means and standard deviations reported in the studies of Gillies (2003a); Krol, Janssen, Veenman, and van der Linden (2004); Swing and Peterson (1982); and Webb and Farivar (1994). The intervention study of Webb and Farivar was conducted within the context of six seventh-grade general mathematics classes (i.e., a 4-week unit on operations with fractions) in the United States. Students' academic helping behaviors were compared in regard to two instructional conditions: CL with instruction and practice in both communication and helping skills and CL with instruction and practice in basic communication skills only. The d indexes for the behavioral categories of "receiving elaborations," "giving elaborations," and "receiving elaborations when requested" were 0.23, 0.26, and 0.69, respectively.

The study by Gillies (2003a) was conducted in Australia within the context of eighth-grade mathematics, science, and English classes (i.e., three school terms). Groups of students who received training in cooperative group behaviors, including helping behaviors, were compared with groups of students not receiving such training. The *d* index for the behavioral category "solicited explanations" was found to be 0.17.

The study by Swing and Peterson (1982) was conducted in the United States among fifth-grade students trained in small-group interaction and explaining skills and fifth-grade students not trained in these skills (the investigation involved a 4-week unit on long division). The *d* indexes for the behavioral categories of "explains conceptual/sequencing" and "receives conceptual/sequencing" were both found to be 0.74.

In a previous study of the effects of a 2-year CL school improvement among sixth-grade students in the Netherlands (Krol et al., 2004), a *d* index of –0.02 was found for provision of "high-level elaborations" during the solution of balance-beam problems. In comparison with the effect sizes found in this earlier Dutch study and the studies of Webb and Farivar (1994), Gillies (2003a), and Swing and Peterson (1982), the effect sizes revealed in the present study were larger. Although direct comparisons of the results of the present study and previous studies—including the global benchmark studies conducted by Hattie et al. (1996) and Rohrbeck et al. (2003)—should be made with caution owing to the differences in designs, reliability of scores, variables studied, treatments, participants, school subjects, and school cultures, we consider the effects shown in our study to be meaningful and promising.

However, as Figure 1 shows, the overall treatment effect on the use of high-level elaborations was not due to an *increase* in the use of high-level elaborations at posttest versus pretest but to a strong *decrease* in the use of high-level elaborations by the control dyads at posttest. This was an unexpected outcome, and it also does not correspond to the findings of Gillies (2003a), Swing and Peterson (1982), or Webb and Farivar (1994), which showed an increase from pretest to posttest.

To gain greater insight into the unexpected decrease in levels of elaboration we observed, we returned to the schools and interviewed the teachers and students with regard to the results of the present study. Results were presented in general terms to the teachers and students, who were then asked whether the outcomes corresponded to their own classroom experiences and what they had learned from the exercises presented in the newsletters. All of the interviews were audiotaped and summarized for further analysis. As shown subsequently, task characteristics, demand characteristics, and effort appeared to account for the decrease as opposed to increase in provision and receipt of high-level elaborations at posttest.

Task Characteristics

It has been argued by Cohen (1994), Gillies (2003a), and Webb and Farivar (1994) that the type of task assigned to a group will largely determine how the members of the group interact. When the task has a correct answer or the answer can typically be found by following a number of well-structured rules or procedures, for example, Cohen (1994) has argued that there is no need for the group to engage in extended discussions of how to solve the problem. Such well-structured tasks (Cohen, 1994) require only a limited exchange of information and thus a low level of cooperation. In contrast, when the task is more open and no one correct answer exists, high levels of cooperation may occur as the group shares ideas and information. Such ill-structured tasks (Cohen, 1994) more or less require interaction and therefore interdependence among the members of the group to reach agreement on a solution. Moreover, when a task has one correct answer or a well-defined set of procedures is used to solve the problem, and one group member knows the answer or has the sufficient expertise to solve the problem, it may actually be most efficient and effective for that group member to provide the answer or solve the problem without the participation of the other group members (Webb, 1995).

The math task used in the present study can be characterized as a "well-structured task." The problems had a single correct answer and could thus be solved by students who were knowledgeable of the rules involved in a balance-beam task. It was possible to solve the problems individually (although these were not the instructions provided), which means that the students were not really interdependent and that the task probably evoked less elaboration than more open or ill-structured tasks. While the post hoc interviews with the teachers and students showed the math task to have many of the characteristics of a well-structured cooperative task, the students in the treatment group were nevertheless found to produce and receive more high-level elaborations at posttest than the students in either of the control groups.

Demand Characteristics

From a motivational perspective, the reward or goal structures entertained by students may explain their achievement in CL situations in particular. Cooperative goals and incentive structures can create a situation in which the only way to attain one's personal goal is to help the group succeed, and the CL literature indeed shows that CL has the most pronounced effect on student learning when groups are recognized or rewarded on the basis of the individual learning of group members (Slavin, 1995). The combination of group goals and individual accountability appears to give students an incentive to help and encourage each other. In solving the math problems in the current study, the individual students were not held fully accountable for achievement of the goal. For instance, students were not informed prior to the cooperative solution of the math problems that their learning would be assessed individually or that the group product would be later rewarded. They were told only that they would be assigned a group score after completion of the math task. In other words, the motivation to simply tell each other the answers and not provide extended explanations was present. In addition, both "social loafing" (i.e., sitting back and letting others do the work) and "cognitive loafing" (i.e., lack of interest in the task or unsuccessful group interaction) were possible (Salomon & Globerson, 1989). Thus, future CL efforts should focus greater attention on the structuring of positive interdependence and the promotion of individual accountability.

Effort

One way to measure effort is to simply note how much time is spent on a task (Lambiotte et al., 1987). The dyads in both the treatment and control groups spent less time actively solving the mathematics task at posttest than at pretest. At pretest, the dyads took about 8.8 minutes to complete the task; at posttest, they took about 6.6 minutes. The performance scores of the dyads in Control Group 1 were also found to suffer from this decreased effort. The decreased time spent on the math task might be explained as follows. First, the posttest was administered 2 months after the national achievement tests used to determine the type of secondary school best suited for a given student. At posttest, most of the students were less motivated to solve the math problems, and the post hoc interviews with the students showed that they had very little motivation to do whatever the group or dyad needed to succeed. After completion of the national achievement tests (i.e., the most important tests in their school lives), most of the students appeared to think that the remaining months of school would not contribute much to their learning. In future research, posttesting should thus be undertaken before administration of national achievement tests.

Second, in the post hoc interviews, the students reported finding the math task at posttest to be very similar to the math task at pretest and therefore to require less effort (e.g., "The second time, the problems were about the same as the first time; you still know how to solve them, and if you both know the answer, there is no need for explanation"). Most of the students simply assumed that their partners understood the rules for solving the balance-beam task problems, and, as a consequence, the discussions engaged in by the dyads were rather superficial. Effective discussion was undermined by "polite vagueness" (Lambiotte et al., 1987) and a predisposition to assume that the solution being proposed by the other member was correct. While the overall difficulty levels of Versions A and B of the math task indicated that they were neither too easy nor too difficult, the data in Table 3 show that the students at posttest scored only between 4.5 and 6.5 (with a maximum of 10 points possible). That is, more effective discussion could have produced higher performance scores.

The results of the present study underline the need to structure learning in small groups and the fact that such discourse features as help seeking, help giving, provision of reasons, and exploratory talk must be practiced and reinforced (Dawes et al., 2000; Webb & Farivar, 1999). The opinions of the teachers in the post hoc interviews also underlined this position.

The performance results showed that the dyads in the treatment group achieved more than the dyads in Control Group 2. Inspection of the posttest data in Table 5 shows that the dyads with experience in CL achieved more than the dyads without such experience. Comparisons can be made with three intervention studies regarding academic helping behaviors and achievement. The overall *d* index for mathematics performance was 0.18 in the study of Webb and Farivar (1994; Grade 7, fractions); the *d* index was −0.05 in the study of Gillies (2003a, 2003b; Grade 8, constructing a two-dimensional pyramid); and the *d* index was 0.77 in the previous study conducted in the Netherlands (Krol et al., 2004; Grade 6, balance beam problems). The 0.76 effect size found in the present study is comparable to that of the earlier Dutch study and exceeds the effect sizes found in the studies by Webb and Farivar and by Gillies. However, no additional facilitation of math performance was found for the treatment group relative to Control Group 1, which shows that the supplementary teacher-training program had no added value above and beyond the value of the earlier 2-year CL school improvement program. At posttest, the schools in Control Group 1 had 2 years of experience with CL but did not differ with respect to achievement from the schools in the treatment group with 2 years of experience and supplementary teacher training.

The regression results showed mastery- and performance-oriented goals to be negatively related to both use of high-level elaborations and math performance. In addition, the use of high-level elaborations was found to be positively related to student achievement. Previous research has not shown that performance and mastery goals are negatively related to provision of elaborated explanations, but this result suggests that students who are performance or mastery oriented are less likely to offer elaborations. It may be that they are too involved with their own goals to recognize implicit or even explicit requests for help from peers, although this possibility needs further exploration in future research. The negative relations between mastery- and performance-oriented goals and student achievement suggest that students with relatively strong mastery- or performance-oriented goals are likely to avoid difficulties or cognitive conflicts and therefore do not bother to seek help, with the result being less learning. However, this explanation also needs further exploration.

The negative relations of mastery- and performance-oriented goals to student achievement may also be the result of the self-report scales used to assess these goals. Previous studies have produced low to modest alpha coefficients ranging from .53 to .78 for assessment of help-seeking and achievement goals. In some cases, this could be due to the low numbers of items constituting the scales (typically two to five) or conceptual problems calling the construct validity of certain items into question (Pajares, Cheong, & Oberman, 2004). Thus, future research should involve attempts to develop scales with stronger psychometric qualities than the scales used in the present study.

The finding that provision of elaborated explanations positively relates to achievement is in line with the results of King (1994), Webb (1989), Webb and Farivar (1994, 1999), and Webb and Palincsar (1996). Provision of elaborated help encourages the explainer to reorganize and clarify information in new ways, while provision of nonelaborated help does not promote such cognitive restructuring (Webb & Farivar, 1994, 1999).

In summary, the relations among help seeking/help giving, achievement goals, and student achievement found here and in the studies by Nelson-Le Gall and Jones (1990), Nelson-Le Gall (1985), Newman (1990), Newman and Schwager (1995), and Ryan and Pintrich (1997) underline the importance of motivational factors in the social-interactional processes that occur in the classroom. Similarly, social influences appear to play a critical role in learning within the context of the classroom.

Limitations

In interpreting the results of this study, some possible limitations should be kept in mind. First, the relatively small numbers of classrooms and dyads limited the statistical power of our analyses. Second, the data were collected on only one occasion at pretest and only one occasion at posttest, possibly limiting the representativeness of the interaction data. In future studies, data should be collected on multiple occasions to allow examination of the stability of student interactive behavior and facilitate assessment of the effects of treatment on student helping behaviors.

Third, the students were studied while working outside the classroom with a video camera running. The choice of videotaped sessions outside the classroom was a practical one: Bilateral interaction in dyads is easier to record and code than multilateral interaction in small groups. However, the outside classroom setting does not resemble the normal classroom situation, which means that the important roles of cooperative and prosocial norms within the classroom and the role of the teacher during CL activities (e.g., monitoring help-seeking and help-giving processes) are not taken into account. Therefore, we cannot rule out the possibility that, within more natural contexts, students in the treatment and control groups might have performed differently. Fourth, the findings regarding affective-motivational resources were based on students' intentions and attitudes. Inferences regarding behavior must be made with care. However, given the importance of attitudes and intentions in regard to human behavior (Ajzen, 1988; Fishbein & Ajzen, 1975; Newman, 1990), our results provide important information that can guide future research on the actual help-seeking and help-giving behaviors of students in the classroom.

Fifth, our study concentrated on the personal characteristics of the student and not the social-interactional or contextual characteristics of the classroom. Classroom norms for CL and prosocial behavior (i.e., it is "normal" to need and provide help) are factors that interact with and actually shape the attitudes, beliefs, and motivational characteristics of students over time, and they should be considered in future research as well (see also Newman, 1990). Additional research is needed to determine, for example, how classroom structures mediate the effects of CL training programs or the relations between different affective-motivational variables and actual help-seeking and help-giving behaviors.

Sixth, a disadvantage of our use of help-seeking and help-giving exercises provided via a newsletter is that the teachers worked largely in isolation. Collegiality and collaboration are important, however, for the successful implementation and sustained use of CL methods. Whole school participation and staff development are recommended (see also Johnson & Johnson, 1998).

Teachers also need support in adapting the generic skills presented during training programs to their unique classroom situations; research on teacher training has shown that transfer occurs only when in-class coaching is added to the initial training experience. Coaching can help teachers implement CL via the provision of technical support, personal assistance, and companionship; it can also promote professional development through feedback, reflection, and discussion (Joyce & Showers, 1995). Future studies should therefore examine the effects of our training program with team implementation and the addition of expert or peer coaching.

Finally, in the present study, students were not asked to occupy specific roles during the CL interactions. Nevertheless, structuring of interactions to increase participation rates merits serious attention in future research. One alternative is the use of a scripted form of cooperation for pairs of students working together. One partner may solve a problem aloud while the other partner asks for elaboration, notes omissions, and points out any errors, and the roles are then reversed for the next problem. Research on scripted cooperation shows such a strategy to be very effective in terms of learning because it appears to promote functional information processing (Lambiotte et al., 1987; O'Donnell & O'Kelly, 1994).

Summary

Despite the aforementioned limitations, it can be concluded that supplementary teacher-training programs aimed at improving the ability of students to provide elaborated explanations can positively influence students' help-seeking and help-giving behaviors. The participation rates of the different students in the present study are currently being examined, in addition to their provision of fully accurate elaborations. This analysis—in combination with the insights provided by the present study—should facilitate the development of an even more effective training program.

References

Ajzen, I. (1988). *Attitudes, personality, and behavior.* Milton Keynes, England: Open University Press.

Ames, R. (1983). Help-seeking and achievement orientation: Perspectives from attribution theory. In B. M. DePaulo, A. Nadler, & J. D. Fisher (Eds.), *New directions in helping* (Vol. 2, pp. 165–186). New York: Academic Press.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64,* 1–35.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cole, M., & Wertsch, J. (1996). Beyond the individual-social antimony in discussions of Piaget and Vygotsky. *Human Development, 39,* 250–256.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement, 61,* 229–248.

Dawes, L., Mercer, N., & Wegerif, R. (2000). *Thinking together: A programme of activities for developing thinking skills at KS2*. Birmingham, England: Questions Publishing.

De Lisi, R., & Goldbeck, S. L. (1999). Implications of Piagetian theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 3–37). Mahwah, NJ: Erlbaum.

Erkens, G. (2002). *Multiple Episode Protocol Analysis (MEPA)* (Version 4.8) [Computer software]. Utrecht, the Netherlands: Utrecht University.

Farivar, S., & Webb, N. (1991). *Helping behavior activities handbook*. Los Angeles: University of California.

Farivar, S., & Webb, N. (1998). Preparing teachers and students for cooperative work: Building communication and helping skills. In C. M. Brody & M. Davidson (Eds.), *Professional development for cooperative learning: Issues and approaches* (pp. 169–187). Albany: State University of New York Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior*. Reading, MA: Addison-Wesley.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Karns, K. (1998). High-achieving students' interactions and performance on complex mathematical tasks as a function of homogeneous and heterogeneous pairings. *American Educational Research Journal, 35,* 227–267.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., Dutka, S., & Katzaroff, M. (1996). The relation between student ability and the quality and effectiveness of explanations. *American Educational Research Journal, 33,* 631–664.

Gillies, R. M. (2003a). The behaviors, interactions, and perceptions of junior high school students during small-group learning. *Journal of Educational Psychology, 95,* 137–147.

Gillies, R. M. (2003b). Structuring cooperative group work in classrooms. *International Journal of Educational Research, 39,* 35–49.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66,* 99–136.

Hertz-Lazarowitz, R. (1989). Cooperation and helping in the classroom: A contextual approach. *International Journal of Educational Research, 13,* 113–119.

Hoek, D., Terwel, J., & van den Eeden, P. (1997). Effects of training in the use of social and cognitive strategies: An intervention study in secondary mathematics in co-operative groups. *Educational Research and Evaluation, 3,* 364–389.

Hogan, D. M., & Tudge, J. R. H. (1999). Implications of Vygotsky's theory for peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 39–65). Mahwah, NJ: Erlbaum.

Johnson, D., & Johnson, R. (1998). Effective staff development in cooperative learning: Training, transfer, and long-term use. In C. M. Brody & N. Davidson (Eds.), *Professional development for cooperative learning: Issues and approaches* (pp. 223–242). New York: State University of New York Press.

Johnson, D. W., & Johnson, R. T. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning* (5th ed.). Boston: Allyn & Bacon.

Joyce, B., & Showers, B. (1995). *Student achievement through staff development* (2nd ed.). New York: Longman.

Kagan, S. (1994). *Cooperative learning* (10th ed.). San Juan Capistrano, CA: Kagan Cooperative Learning.

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.

King, A. (1999). Discourse patterns for mediating peer learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 87–115). Mahwah, NJ: Erlbaum.

Kourilsky, M., & Wittrock, M. C. (1992). Generative teaching: An enhancement strategy for the learning of economics in cooperative groups. *American Educational Research Journal, 29,* 861–876.

Krol, K. (2004). *From independence to interdependence: Implementation of cooperative learning in primary schools.* Manuscript in preparation.

Krol, K., Janssen, J., Veenman, S., & van der Linden, J. (2004). Effects of a cooperative learning program on the elaborations of students working in dyads. *Educational Research and Evaluation, 10,* 205–237.

Krol, K., Veenman, S., & Voeten, M. (2002). Toward a more cooperative classroom: Observations of teachers' instructional behaviors. *Journal of Classroom Interaction, 37,* 37–46.

Lambiotte, J. G., Danserau, D. F., O'Donnell, A. M., Young, M. D., Skaggs, L. P., Hall, R. H., & Rocklin, T. R. (1987). Manipulating cooperative scripts for teaching and learning. *Journal of Educational Psychology, 79,* 424–430.

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective.* Belmont, CA: Wadsworth.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners.* Clevedon, England: Multilingual Matters.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal, 25,* 95–111.

Nelson-Le Gall, S. (1981). Help-seeking: An understudied problem-solving skill in children. *Developmental Review, 1,* 224–246.

Nelson-Le Gall, S. (1985). Help-seeking behavior in learning. In E. W. Gordon (Ed.), *Review of research in education* (Vol. 12, pp. 55–90). Washington, DC: American Educational Research Association.

Nelson-Le Gall, S. (1992). Children's instrumental help-seeking: Its role in the social acquisition and construction of knowledge. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 49–68). Cambridge, England: Cambridge University Press.

Nelson-Le Gall, S., & Glor-Scheib, S. (1985). Help seeking in elementary classrooms: An observational study. *Contemporary Educational Psychology, 10,* 58–71.

Nelson-Le Gall, S., & Glor-Scheib, S. (1986). Academic help-seeking and peer relations in school. *Contemporary Educational Psychology, 11,* 58–71.

Nelson-Le Gall, S., & Jones, E. (1990). Cognitive-motivational influences on the task-related help-seeking behavior of Black children. *Child Development, 61,* 581–589.

Newman, R. S. (1990). Children's help-seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology, 82,* 71–80.

Newman, R. S. (1994). Adaptive help seeking: A strategy of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and applications* (pp. 283–301). Hillsdale, NJ: Erlbaum.

Newman, R. S. (1998). Students' help seeking during problem solving: Influences of personal and contextual achievement goals. *Journal of Educational Psychology, 90,* 644–658.

Newman, R. S., & Schwager, M. T. (1993). Students' perceptions of the teacher and classmates in relation to reported help seeking in math class. *Elementary School Journal, 91,* 3–17.

Newman, R. S., & Schwager, M. T. (1995). Students' help seeking during problem solving: Effects of grade, goal, and prior achievement. *American Educational Research Journal, 32,* 352–376.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review, 6,* 321–349.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25,* 241–286.

O'Malley, C. (Ed.). (1995). *Computer supported collaborative learning.* Berlin: Springer-Verlag.

Onwuegbuzie, A. J., & Daniel, L. G. (2003). Typology of analytical and interpretational errors in quantitative and qualitative educational research. *Current Issues in Education, 6*(2). Retrieved February 19, 2004, from http://cie.ed.asu.edu/volume6/number2.

Pajares, F., Cheong, Y. F., & Oberman, P. (2004). Psychometric analyses of computer science help-seeking scales. *Educational and Psychological Measurement, 64,* 496–513.

Phelps, E., & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology, 81,* 639–646.

Piaget, J. (1926). *Judgment and reasoning in the child.* New York: Harcourt.

Puustinen, M. (1998). Help-seeking behavior in a problem-solving situation: Development of self-regulation. *European Journal of Psychology of Education, 8,* 271–282.

Qin, Z., Johnson, D. W., & Johnson, R. T. (1995). Cooperative versus competitive efforts and problem solving. *Review of Educational Research, 65,* 129–143.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95,* 240–257.

Ros, A. A. (1994). *Analyse en effecten van uitleg door leerlingen en leerkrachten* [Explanations by students and teachers: Analyses and effects]. Unpublished doctoral dissertation, University of Groningen, Groningen, the Netherlands.

Ryan, A. M., & Pintrich, P. R. (1997). "Should I ask for help?" The role of motivation and attitudes in adolescents' help seeking in math class. *Journal of Educational Psychology, 89,* 329–341.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Educational Research, 13,* 89–99.

Shachar, H. (2003). Who gains what from co-operative learning: An overview of eight studies. In R. M. Gillies & A. F. Ashman (Eds.), *Co-operative learning: The social and intellectual outcomes of learning in groups* (pp. 103–118). London: Routledge-Falmer.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481–520.

Slavin, R. E. (1995). *Cooperative learning: Theory, research and practice* (2nd ed.). Boston: Allyn & Bacon.

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

Swing, S. R., & Peterson, P. L. (1982). The relationship of student ability and small-group interaction to student achievement. *American Educational Research Journal, 19,* 259–274.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher, 31*(3), 25–32.

Thompson, B., & Borrello, G. M. (1985). The importance of structure coefficients in regression research. *Educational and Psychological Measurement, 45,* 203–209.

Tudge, J. R. H. (1992). Processes and consequences of peer collaboration: A Vygotskian analysis. *Child Development, 63,* 1364–1379.

Van Boxtel, C. (2000). *Collaborative concept learning: Collaborative learning tasks, student interaction, and the learning of physics concepts.* Unpublished doctoral dissertation, Utrecht University, Utrecht, the Netherlands.

Van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000). The use of textbooks as a tool during collaborative physics learning. *Journal of Experimental Education, 69,* 57–76.

Veenman, S., Kenter, B., & Post, K. (2000). Cooperative learning in Dutch primary classrooms. *Educational Studies, 26,* 281–302.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

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

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

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis, 17,* 239–261.

Webb, N. M., Baxter, G. P., & Thompson, L. (1997). Teachers' grouping practices in fifth-grade science classrooms. *Elementary School Journal, 98,* 91–113.

Webb, N. M., Ender, P., & Lewis, S. (1986). Problem-solving strategies and group processes in small groups learning computer programming. *American Educational Research Journal, 23,* 243–261.

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31,* 369–395.

Webb, N. M., & Farivar, S. (1999). Developing productive group interaction in middle school mathematics. In A. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 117–149). Mahwah, NJ: Erlbaum.

Webb, N. M., Farivar, S. H., & Mastergeorge, A. M. (2002). Productive helping in groups. *Theory into Practice, 41,* 13–20.

Webb, N. M., & Mastergeorge, A. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39,* 73–97.

Webb, N. M., Nemer, K. M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. *American Educational Research Journal, 35,* 607–651.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 841–873). New York: Simon & Schuster.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87,* 406–423.

Wegerif, R. (1996). Using computers to help coach exploratory talk across the curriculum. *Computers and Education, 26,* 51–60.

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible sociocultural model of cognitive development. *Learning and Instruction, 32,* 493–516.

Wilkinson, L., & American Psychological Association Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594–604.

Wittrock, M. C. (1991). Generative processes of comprehension. *Elementary School Journal, 92,* 169–184.

Manuscript received January 19, 2004
Revision received November 22, 2004
Accepted November 29, 2004