American Educational Research Journal
Summer 2004, Vol. 41, No. 2, pp. 365–399

Adapting Teacher Interventions to Student Needs During Cooperative Learning: How to Improve Student Problem Solving and Time On-Task

Ming Ming Chiu
Chinese University of Hong Kong

This study tested a model of teacher interventions (TIs) conducted during cooperative learning to examine how they affected students' subsequent time on-task (TOT) and problem solving. TIs involved groups of ninth-grade students working on an algebra problem; videotaped lessons were transcribed and analyzed. Results showed that teachers initiated most TIs and typically did so when students were off-task or showed little progress. After TIs, students' TOT and problem solving often improved. Teacher evaluations of student actions had the largest positive effects, serving as gatekeepers for other teacher actions. Higher levels of teacher help content tended to reduce post-TI TOT, while teacher commands reduced post-TI TOT only when a group grasped the problem situation. In summary, TIs can increase TOT and problem solving, especially if teachers evaluate students' work.

KEYWORDS: cooperative learning, problem solving, structural equation modeling, teacher evaluation.

Suppose you are a teacher and have several groups of students working on an algebra problem. Some groups are working productively, while others are not. When do you intervene? How should you intervene? What kinds of interventions are likely to improve students' problem solving?

In the present study, I addressed these questions by building on past research on cooperative learning (CL). A group of students engaged in CL works together to achieve shared goals. Ideally, students help one another learn

---

MING MING CHIU is an associate professor at the Chinese University of Hong Kong, Department of Educational Psychology, 314 Ho Tim Building, Shatin, NT, Hong Kong; e-mail *mingming@cuhk.edu.hk*. Using advanced statistical methods, he examines group processes during cooperative learning and compares student learning in large-scale international studies.

and perform better than they would if they were working alone (Johnson & Johnson, 1994). Teachers have used CL to improve their students' academic achievement, motivation, racial attitudes, and so on (see Good, Mulryan, & McCaslin, 1992, and Slavin, 1995, for reviews). However, simply putting students in a group does not ensure successful CL (e.g., Laughlin, VanderStoep, & Hollingshead, 1991; Newman & Thompson, 1987). Successful CL depends on teachers' choice of suitable group structures (e.g., group size, group composition, rewards, problem type, resources) and forms of student interaction (e.g., helping, communicating; for reviews of group structure and interactions, see Webb & Palincsar, 1996, and E. G. Cohen, 1994a). Even in well-planned CL lessons, students sometimes have difficulties and need help from the teacher (e.g., Webb & Farivar, 1999). If teachers could adapt their interventions to students' specific needs at specific times, students might work more productively afterward and thereby reap the potential benefits of CL. A working model of teacher interventions (TIs) can help teachers decide when to intervene and how to do so effectively.

In this article, I introduce and test a model of TIs implemented during CL. This type of TI refers to any sequence of teacher interactions with a group of students about their work. I analyzed 108 TIs involving 220 ninth-grade students working in 55 groups and examined their effects on students' subsequent time on-task (TOT) and problem solving. Specifically, I analyzed when TIs occurred, who initiated them, their relative success, and the effects of various patterns of teacher actions during these interventions.

Toward a TI Framework

This section begins with a discussion of past research on TI effectiveness. Next, I discuss factors that can affect TI effectiveness such as teacher actions during TIs. Then I extend this line of research by considering how the local student group context, initiation of the TI, and teacher adaptation to the group context might affect TIs and their outcomes. Finally, I introduce a TI model that combines all of these aspects.

TI Effectiveness

Past studies showed that TIs can improve student behaviors both during and after the intervention. For example, Harwood (1995) examined groups of primary school children discussing current world issues. Harwood examined groups' interactions (a) when the teacher facilitated their conversations and (b) when the teacher was elsewhere. In the teacher's presence, students stayed on topic more often (83% vs. 70%), made more correct inferences (53% vs. 14%), and made more justifications (58% vs. 32%) than in her absence. However, the students produced fewer new ideas (23% vs. 70%) in the teacher's presence than in her absence.

Meloth and Deering's (1999) reanalysis of Meloth and Barbe's (1992) study of 15 third-grade classrooms further showed that TIs can affect students' subsequent behaviors. After effective TIs, students cooperated more often, shared more information that was task relevant, and provided more explanations than before the TI. After ineffective TIs, however, students discussed the topic superficially or were off-task until the teacher returned. Although some TIs improved students' CL, others did not.

Teacher Actions During TIs

To explain why the effects of TIs can differ, researchers have explored teacher actions during TIs, such as help content and questions. Results have shown that teachers who provide low-help-content TIs using questions and those who provide high-help-content TIs without using questions can both succeed (E. G. Cohen, 1994b; Harwood, 1995; Meloth & Deering, 1999).

E. G. Cohen (1994b) argued that a teacher facing an off-task group should ask brief questions or make a few comments and then walk away, so that students have the opportunity to discuss the task on their own. Questions engage an audience more than statements do (Chiu, 2000, 2001; Tsui, 1994) and are less intrusive than commands, which involve demands on audience members to engage in specific actions (Brown & Levinson, 1987). By using questions, teachers can help students focus on the task while minimizing their intrusiveness. Harwood (1995) showed that when teachers asked students challenging questions with little problem content information, students often explained and justified their answers during the TI. Consider the following example in which a teacher intervened when a student erred.

[Teacher (Ms. T) approaches group]

Eva: Twenty plus five, twenty-five, and—
Ms. T: —Why did you do that?
Eva: Why? Cost is rate, rate times time. Twenty times five is a hundred.
Ms. T: Right. [walks away]

Rather than correcting and explaining the error, Ms. T challenged Eva to explain her answer ("Why did you do that?"). During her explanation, Eva realized that she should multiply rather than add the two numbers. The teacher's short response highlighted the students' responsibility for and control over their discussions. As students accept responsibility for their work, they tend to initiate more new ideas and more solution proposals than students who rely on the teacher (E. G. Cohen, 1994b). Increasing student responsibility likewise promotes group interdependence while reducing dependence on the teacher (E. G. Cohen, 1994b). In short, teachers who use questions to provide minimal problem content information can improve students' CL by bolstering their autonomy, initiative, and interdependence.

In contrast, Meloth and Deering (1999) showed that low-content TIs involving the use of questions were not necessarily helpful and that high-content TIs often improved student performance without reducing student autonomy. Sometimes teachers did not provide students enough information to discuss the topic meaningfully. As a result, these students did not make further progress until the teacher circled back to them for another TI. Meloth and Deering (1999) also showed that high-content TIs can be effective. They examined TIs in which teachers provided a group of students with five or more statements regarding task content or effective communication (*instructional exchanges*). Consider the following example:

> *Eva:* Twenty cents times five minutes is twenty-five.
> *Juan:* Um, how about three minutes—
> *Ms. T:* —If you're trying out different possibilities to find the critical number of minutes, try doing it in sequence beginning with simple numbers, like zero and one. So, try zero minutes, and then one minute, then two minutes, then three minutes. Write down the answers. See if you get a pattern. Then based on that pattern, estimate when you would reach the critical point and test numbers in that vicinity.
> *Eva:* OK. Twenty times zero is zero. . . .

After many of these high-content TIs, the students performed better than they had before the TI without any decline in their autonomy. They cooperated more, shared more problem-related information, and explained their ideas more fully than before the TI.

In instructional exchanges, teachers provide information, link it to the task, and ask only a few questions (Meloth & Deering, 1999). First, Ms. T provided students information that they probably lacked, such as conceptual explanations, solution tactics, and metacognitive actions. Metacognitive actions include evaluation of problem-solving progress and strategic planning (e.g., "Try doing it in sequence beginning with simple numbers, like zero and one"; for a discussion of metacognition, see Hacker, 1998). Second, Ms. T linked the information to the task ("If you're trying out different possibilities to find the critical number of minutes, try. . . ."). Third, Ms. T asked only a few questions and quickly moved to the main point (using the commands in the preceding example). With this new information and strategic plan, students could understand the problem and implement the teacher's plan to solve it.

Moreover, Meloth and Deering (1999) argued that student autonomy did not diminish during these high-content TIs. The teachers in 4 of the 15 classes used an increased number of instructional exchanges, but students did not show greater reliance on these teachers. For example, students did not ask for their immediate help more often than did those in the other classes.

TI Context and Initiation

Both low-content, question-based TIs and high-content TIs not involving questions can succeed, so each approach might fit different contexts. Few researchers have shown how differences in pre-TI group contexts affect intervention outcomes.

E. G. Cohen (1994b) argued that a teacher should intervene when groups are off-task, making very little progress toward a solution, or experiencing sharp interpersonal conflicts (e.g., see Salomon & Globerson, 1989). If students are on-task, teachers can monitor their work without intervening, as excessive TIs can disrupt student autonomy and interdependence (E. G. Cohen, 1994b). However, researchers have not shown that teachers usually intervene during CL under these conditions. For example, teachers might have other TI criteria, such as intervening mostly in groups whose members have shown low past achievement (as these groups might need more help than groups showing higher past achievement).

Furthermore, researchers have not shown how properties of the local group context affect TI effectiveness. Consider the following pre-TI contextual aspects: degree of off-task student behaviors, group's overall and recent problem-solving progress, and group's past achievement. One could argue that groups with more on-task members are more likely to stay on-task and to benefit from a TI than groups with fewer on-task members. Groups showing more overall problem-solving progress than other groups might also benefit more from a TI because they have more parts of the solution that they can use to understand and apply the teacher's help. Perhaps groups showing more recent problem-solving progress than other groups are more attentive and receptive to teacher input. Finally, groups showing higher past achievement than other groups might have more relevant knowledge or greater competence to capitalize on the TI.

The situation in which the TI is initiated might also affect its outcome. TIs in which a student asks for specific help might differ from those in which a student asks a general question or those initiated by a teacher. A student asking for specific help identifies the group's plight (e.g., "Does rate times minutes give us the total cost?") and often elicits helpful responses from other students (Peterson, Wilkinson, Spinelli, & Swing, 1984; Webb & Kenderski, 1984; Wilkinson, 1985). Hence, student requests for specific help might ease teacher evaluation of the group's work and elicit a helpful teacher response, which in turn might improve the TI outcome.

In contrast, students who cannot identify or recognize their plight rely on the teacher's diagnosis of the problem, so they might benefit less from a TI than those who recognize that they need help. Students who lack the metacognitive skills (Hacker, 1998; Nelson-Le Gall, 1992) to identify their problem can ask a teacher for *general help* (e.g., "What are we supposed to do next?"). A teacher can also begin a TI if students do not recognize their problem or are unwilling to seek help as a result of inadequate self-knowledge (Nelson-Le Gall, 1992), concern over loss of face (Weiner, 1980), or fruitless past experiences (Webb, Troper, & Fall, 1995).

Teacher Adaptation to TI Context

Because factors associated with local group contexts and TI initiation can differ, teachers' diagnoses of students' needs and teachers' subsequent choice of actions might affect intervention outcomes. By evaluating a group's past progress and current problem-solving approach, a teacher might pinpoint its needs and adapt TI actions to address them. In previous work (Chiu, 2001), I have argued that people's evaluations of an idea reveal whether they have listened to the person relating that idea, understood the idea, and thereby showed respect (Brown & Levinson, 1987). In this case teachers, in their evaluations, could advocate that groups either continue their problem-solving approach (agreement or a neutral approach) or change it (disagreement). Thus, these evaluations could guide students' problem solving. Consider a reanalysis of the first TI example in terms of teacher evaluations.

[Ms. T approaches the group]

Eva: Twenty plus five, twenty-five, and—
Ms. T: —Why did you do that?
Eva: Why? Cost is rate, rate times time. Twenty times five is a hundred.
Ms. T: Right. [walks away]

Ms. T, offering an implicit, critical evaluation, challenged Eva to explain her answer ("Why did you do that?"). While explaining her computation, Eva realized that she should multiply rather than add the two numbers. Then Ms. T confirmed that Eva's new computation was correct.

In contrast, a teacher who ignores students' ideas might impose preconceived directives rather than diagnosing the students' needs to adapt the TI accordingly. Consider the following example in which Eva was using a table to solve the problem.

Eva: Twenty cents times five minutes is twenty-five cents—
Ms. T: —Have you considered using algebraic equations?
Eva: Equations?

Ms. T could have helped Eva pinpoint an arithmetic error and then allowed her to pursue a promising solution path using tables. Instead, she introduced an unrelated solution method that Eva did not grasp, thereby interfering with Eva's progress.

TI Model and Research Questions

Combining the elements just described (local context, TI initiation, teacher actions, and TI effectiveness) yields the TI model shown in Table 1. The local context of the group might affect whether the teacher or a member of the group initiates a TI. Both the local group context and the TI initiation might affect the teacher's evaluation of the group's needs. The teacher's evaluation, in turn, might affect his or her choice of actions, including how much problem content information to provide and whether to use commands, questions, or statements. All of these elements might affect students' post-TI TOT and problem solving. TOT captures Harwood's (1995) "staying on topic" and Meloth and Deering's (1999) task-relevant behaviors, while problem solving captures both correct ideas (Harwood, 1995) and explanations (Harwood, 1995; Meloth & Deering, 1999). Past studies have shown that students are more likely to

Table 1

Teacher Intervention (TI) Model, Including Local Group Context, Initiation of the Intervention, Teacher Actions, and Effects on Student Behaviors

Pre-TI		TI initiated by:		Teacher action		Post-TI
Group time on-task Group overall progress Group recent problem solving Group past achievement	→	Student specific question? Student general question? Teacher? →	→	Evaluations Help content Questions Commands →	→	Group time on-task Group problem solving

solve a problem if they are on-task more often (e.g., Forman & Cazden, 1985) or generate more correct ideas (e.g., Chiu & Khoo, 2003).

Unlike past TI studies examining selected TIs (e.g., Meloth & Deering, 1999) or isolated (Harwood, 1995) aspects of TIs, this study systematically tested a TI model through the use of all TIs in each lesson. Key questions included the following. When do TIs occur? Who initiates them? Do TIs improve TOT and problem solving? Which patterns of teacher actions are effective? An ancillary analysis conducted at the group level showed the broader context of the TIs through the relationships among a group's past achievement, TOT, and solution score.

Method

Participants

Two female, non-Hispanic White teachers with 4 and 5 years of teaching experience each taught three classes. Both had bachelor's degrees in mathematics and teaching certification diplomas. They discussed lesson plans together and, based on Johnson and Johnson's (1994) "learning together" approach, used CL for a period of 2 years. (Johnson and Johnson [1994] did not advocate either a high or low teacher content approach.)

Two hundred twenty students from six 9th-grade classes formed 55 groups. (Groups of 3 students or fewer were videotaped, but their data were not used in this study.) They resided in mostly lower- to middle-class neighborhoods and attended a large, urban public school in which classes were not divided according to student ability (i.e., were not "tracked"). One hundred fourteen of the students were girls (52%), and 106 were boys (48%). Eight percent were Asian, 21% were Black, 38% were Latino, and 33% were White. According to the teachers, more than 80% of the school's students graduate, and about 40% of graduates pursue postsecondary studies.

The students taking part in this study had been working together in groups of four at least once per week for a month. For a period of 2 months before CL lessons, pairs of classmates discussed their work before sharing their ideas with the entire class ("think-pair-share"; Johnson & Johnson, 1994).

Procedure

In the unit assessed here, on single-variable equations, the teachers introduced a new problem involving representation of two phenomena.

> You want to get a mobile phone. Quickie Co. and Speedy Service both give away free phones but have different rates. Quickie charges $10 per month and 20 cents per minute. Speedy charges $25 per month and 5 cents per minute. How many minutes should you use each month so that Speedy costs less? Explain.

Both teachers typically introduced the problems with a 1–2-minute discussion focusing on context. In this case, the teachers asked students questions about mobile phones, such as how they were used. As with most of the CL problems assigned by these teachers, this one was difficult and involved multiple solution methods. Each teacher told the students to ask their group members for help first but otherwise freely responded to their questions. Both teachers displayed a poster that showed the following prompts: (a) understand the problem, (b) propose solution ideas, (c) listen to others, (d) give reasons for or against a proposal, and (e) ask others to clarify. The CL groups worked on the problem for 30 minutes with pens, paper, and calculators. Meanwhile, the teacher walked around the room, monitoring their progress. If students found a correct solution, the teacher asked them to check their answer by using another solution method.

After the CL activity, the teacher randomly selected a student to represent his or her group and to explain the group's solution(s) to the class, thereby giving students another incentive to cooperate during CL. After some of the CL activities, including this one, the teacher asked students to briefly write down (a) what they learned, (b) something they liked about the activity, and (c) something they did not like about the activity. The problem had algebraic, graphical, and tabular solutions. For example, students could write the inequality of the cost equations and solve for the number of minutes (m) as follows: $1{,}000 + 20m > 2{,}500 + 5m \rightarrow 20m > 1{,}500 + 5m \rightarrow 15m > 1{,}500 \rightarrow m > 100$ (answer: more than 100 minutes).

Data

Each group of students and each teacher were videotaped separately with 13 video cameras, resulting in 1,489 minutes of videotape. Research assistants (RAs) desensitized the students and the teacher to the video camera by videotaping the students for 1 week before collecting the data used in this study (although 2 days seemed sufficient). The RAs also audiotaped the classes to facilitate transcription. They collected and photocopied final solutions, notes, and other paperwork. To trace each group's problem-solving progress and to identify TIs, the RAs created summary logs by watching the videotapes. Then they transcribed the TI segments of the student audiotapes and added gestures from the videotapes (McNeil, 1992) and samples of students' writing from their papers. As mentioned, the final data set did not include groups with three or fewer students. It also did not include five TIs with groups of students that had already solved the problem correctly.

Variables

I used data from an earlier pilot study involving several groups of students to train two RAs in coding of the transcripts. Coders were unaware of the study's research questions. J. Cohen's (1960) kappa value was used in tests of intercoder reliability (see Tables 2 and 3). The RAs and the author settled coding disagreements by consensus.

TI

A TI was operationally defined as a sequence of speaker turns in which the teacher and students in a CL group addressed one another regarding the assigned problem.

Pre-TI and Post-TI TOT Rating

RAs coded each group of students' TOT 1 minute before the TI (pre-TI TOT) and 1 minute after the TI (post-TI TOT). The RAs coded each student's time as on-task or off-task (1 or 0, respectively). *On-task* was operationally defined as engaging in student behaviors conducive to problem solving, including suggesting or executing solution proposals, encouraging others to work on the problem, listening to solution proposals, and discussing solutions. Whenever students were not on-task (e.g., looking out of the window), they were coded as *off-task*. A student's TOT rating during a given 1-minute time period was computed as the following ratio: TOT seconds/60 seconds. Each group's TOT rating was the mean of all of its members' TOT ratings.

Pre-TI Progress

RAs scored each group (on a 0–3 scale) on the basis of its overall progress immediately before the TI. As noted earlier, the RAs created a log of each group's progress. "Speedy costs less if we use more than 100 minutes each month" was scored as a correct solution (i.e., a value of 3). Some groups used a correct method but did not solve the problem correctly (e.g., "1,000 + 20 cents per minute = 2,500 + 5 cents per minute, *25* cents per minute = 1,500, 60 minutes"); these groups were assigned a score of 2. Finally, other groups grasped the problem situation but did not use a correct method (e.g., "We have a base cost of 1,000 for Quickie and 2,500 for Speedy, then we need to add 20 cents per minute for Quickie and 5 cents per minute for Speedy"); these groups were assigned a score of 1. Some groups did not achieve any progress and

Table 2
Summary Table of Continuous Variables

Variable	*M*	*SD*	Min.	Max.	*N*	Cohen κ	*SE*
Highest mathematics grade in group	88.51	6.04	80	98	108	N/A	N/A
Overall mathematics grade in group	80.39	6.36	70	90.75	108	N/A	N/A
Post-TI TOT rating							
1 minute	0.44	0.73	0	1	108	.921	.000
2 minutes	0.42	0.56	0.04	0.98	108	.934	.000
3 minutes	0.40	0.32	0.06	0.95	108	.929	.000
4 minutes	0.40	0.19	0.07	0.93	108	.930	.000
5 minutes	0.39	0.16	0.06	0.90	108	.927	.000
Pre-TI TOT rating							
1 minute	0.25	0.80	0.00	1.00	108	.913	.000
2 minutes	0.27	0.55	0.06	0.95	108	.909	.000
3 minutes	0.28	0.33	0.05	0.92	108	.917	.000
4 minutes	0.30	0.22	0.08	0.91	108	.921	.000
5 minutes	0.31	0.19	0.04	0.91	108	.924	.000
Teacher agreement percentage	0.34	0.48	0	1	108	.903	.001
Teacher disagreement percentage	0.43	0.50	0	1	108	.903	.001
Teacher question percentage	0.38	0.36	0	1	108	.955	.002
Teacher command percentage	0.18	0.34	0	1	108	.955	.002
Teacher Command Percentage × Pre-TI Progress	0.11	0.29	0	1	108	.955	.002
Total words in TI	45.91	36.71	9	175	108	N/A	N/A
Percentage of teacher words in TI	0.44	0.25	0.05	0.96	108	N/A	N/A
Teacher interventions per class	18.00	2.10	14	20	6	N/A	N/A
TOT rating	0.58	0.15	0.27	0.98	55	.924	.000
Teacher interventions per group	1.85	0.76	1.00	3	55	N/A	N/A

Note. Min. = minimum; Max. = maximum; N/A = not applicable; TI = teacher intervention; TOT = time on-task.

received a score of 0. Pre-TI progress was a broad, rough measure of each group's overall progress before the initiation of a specific TI.

Pre-TI and Post-TI Problem Solving

Unlike pre-TI progress, which characterizes a group's entire problem solving before the TI, this measure captures only the group's work during a 1-minute

Table 3
Summary Table of Discrete Variables

	Scale rating (0–3 range) (percentage)						
Variable	0	1	2	3	*N*	Cohen κ	*SE*
Post-TI problem solving							
1 minute	30	26	44	1	108	.901	.005
2 minutes	23	17	38	22	108	.904	.003
3 minutes	22	10	30	38	108	.903	.006
4 minutes	19	9	23	49	108	.911	.002
5 minutes	14	6	20	60	108	.909	.002
Pre-TI overall progress	41	51	8	0	108	.856	.006
Pre-TI problem solving							
1 minute	81	13	5	1	108	.921	.000
2 minutes	67	14	11	8	108	.934	.000
3 minutes	54	15	16	16	108	.929	.000
4 minutes	44	13	14	29	108	.930	.000
5 minutes	32	10	17	41	108	.927	.000
Student specific initiation of TI	80	20	0	0	108	N/A	N/A
Student general initiation of TI	94	6	0	0	108	N/A	N/A
Teacher identification	49	51	0	0	108	N/A	N/A
Teacher help content level	48	22	29	1	108	.884	.001
Solution score	9	24	20	47	55	.803	.018

Note. TI = teacher intervention; N/A = not applicable.

interval. As with TOT, RAs coded each student group's problem solving 1 minute before the TI (pre-TI problem solving) and 1 minute after the TI (post-TI problem solving). RAs scored each group's work (on a 0–3 scale) during each 1-minute time period. Groups that offered a correct new idea (or ideas) with an explanation (e.g., "We convert 10 dollars to 1,000 because we need to use the same unit, cents") were assigned a score of 3. Groups that offered a correct new idea (or ideas) without an explanation (e.g., "Change 10 dollars to 1,000") were assigned a score of 2. Groups that could detect an incorrect idea ("10 dollars plus 20 minutes is 30—wait, that's not right") were assigned a score of 1. Finally, groups that expressed no new ideas or offered an (undetected) incorrect idea were assigned a score of 0. The summary log of problem-solving progress showed which ideas were new. If students had more than one idea in 1 minute, the highest scoring idea was used.

TI Initiation

Binary variables—students' specific start and students' general start—were used to indicate student responsibility for initiation of the TI. The former variable indicated whether the student(s) initiated the TI with a specific question such as "How do we find the cost of Speedy per month?" The latter variable indicated whether the student(s) began with anything other than a specific question, namely a general question (e.g., "What are we supposed to do?"). If the teacher initiated the TI, both variables were coded 0.

Teacher

A binary variable indicated which teacher taught the class.

Teacher Evaluations

In their evaluations, the teachers assessed the validity of the student's speaking turn during a TI (Chiu, 2000, 2001). Suppose that, before the teacher spoke, a student said "Twenty minus five is twenty-five" in the midst of a correct solution approach using algebraic equations. A teacher turn was operationally defined as "agreement" if the teacher *fully* agreed with the student turn. Likewise, a teacher turn was defined as "disagreement" if the teacher disagreed *at least in part* with the student turn (e.g., "That's the right idea, but check your subtraction"). A third possibility was that of the teacher ignoring the student and introducing a new topic (e.g., "Try making a table"). Percentage of teacher agreement during a given TI was computed as the ratio of total teacher agree turns to total teacher turns. Similarly, percentage of teacher disagreement during a particular TI was computed as the ratio of total teacher disagree turns to total teacher turns. Since there could be different causes and consequences of agreement and disagreement, these variables were coded separately in the analyses.

If a student question included a specific idea (e.g., "Do we multiply twenty times five?"), the teacher could agree, disagree, or ignore the idea. Consider the general question "What are we supposed to do next?" The teacher could agree to answer the question (e.g., "Try making a table"), decline to answer the question and ask the students to answer it instead (e.g., "What do you think the next step is?"), or ignore the question (e.g., "We only have 5 minutes left").

Teacher Help Content

RAs scored teacher help content on a 0–3 scale, as follows: none (0), drawing attention to a concept or aspect of the problem situation (1), explaining a concept or problem aspect (2), or giving a solution tactic (3). If a teacher provided more than one type of help during a particular turn, the type of help scored highest was used.

Teacher Questions and Commands

A teacher's *question percentage* during a TI was computed as total number of questions divided by total number of actions. Chiu (2000) defined an action as a sequence of words, motions, and/or drawings bracketed by pauses or falling intonations (in the case of words). Likewise, a teacher's *command percentage* during a TI was computed as total number of commands divided by total number of actions.

Solution Score

RAs scored each group's final answer on a 0–3 scale, as follows: correct solution (3), correct method (2), correct understanding of the problem situation (1), or none of the above (0). (See the Pre-TI Progress section for examples.)

Mean TOT Rating

A student's mean TOT rating was computed as total TOT divided by total amount of time. Total time referred to the interval between the beginning of the problem-solving session and either the end of the session or agreement on a final solution. Each group's TOT rating was the mean of all of its members' TOT ratings.

Past Mathematics Achievement

Students' midyear algebra grades were used to compute each group's highest grade and mean grade.

Robustness of Results

In the initial analysis, a 1-minute period was used to assess each of the following variables: pre-TI TOT, post-TI TOT, pre-TI problem solving, and post-TI problem solving. As a means of testing for robustness, analyses were repeated at 2, 3, 4, and 5 minutes. (Some TIs occurred within 6 minutes of one another.)

Analysis

I analyzed data both at the TI level (to test the hypotheses described earlier) and at the group solution level (to further illustrate the contexts of TIs). I used regression analyses, *t* tests, and Wilcoxon tests to examine when TIs occurred and whether they had positive effects on TOT and problem solving (Sheskin, 1997, pp. 291–302). I used regression analyses, path analyses, and structural equation modeling (Bollen, 1989) to test the TI model outlined here. The regression analyses estimated the effects of predictors on a given outcome variable. On the basis of the regression results, the path analyses estimated the effects of the significant predictors on one another *separately*. Finally, the structural equation model (SEM) *simultaneously* tested all of these significant effects to assess the overall fit of the model to the data.

First, when did TIs occur? They might have occurred primarily in groups that were off-task, had low past achievement, showed little overall progress, or showed little recent problem solving. As a means of testing these hypotheses, the number of TIs in each group was regressed against each of the following variables: group mean TOT, group mean mathematics grade, and highest mathematics grade in the group.

TIs could also have occurred when a group had a lower TOT than usual, less overall progress than usual, or less recent problem-solving success than usual. I analyzed the TOT data using a mean difference *t* test between pre-TI TOT and mean TOT (pre-TI TOT minus mean TOT). Because both overall progress and recent problem solving were ordered variables, rather than continuous variables, differences between pre-TI and median values were tested with a Wilcoxon signed rank test (Sheskin, 1997).

Next, did TIs improve TOT or problem solving? Mean difference *t* tests compared whether post-TI TOT differed significantly from pre-TI TOT (post-TI TOT minus pre-TI TOT). Similarly, a Wilcoxon matched pairs signed rank test (Sheskin, 1997, pp. 291–302) indicated whether pre-TI and post-TI problem solving differed significantly.

I tested predictors of post-TI TOT ratings and post-TI problem solving for CL groups using hierarchical set regressions, path analyses, and SEMs. Maximum-likelihood ordered Logit/Probit/Gompit regressions were used (Aitchison & Silvey, 1957) to examine all ordered outcome variables (e.g., post-TI problem solving and solution score). Ordered variables do not have fixed intervals between values, so using a least squares regression (which assumes fixed intervals) would lead to bias in estimation of standard errors (Finney, 1971). In contrast, Logit, Probit, and Gompit models estimate the likelihood that a variable value is of a higher value rather than a lower value (e.g., 1 rather than 0). Combining overlapping models, such as (0, 1) and (1, 2), yields an ordered model (0, 1, 2) that does not require fixed intervals (Aitchison & Silvey, 1957). The results also incorporated McFadden's (1974) R^2 value, an estimate of the coefficient of determination. The best fitting ordered Logit, Probit, or Gompit model had the lowest Akaike information criterion (AIC; Grasa, 1989), a measure of goodness of fit adjusted by a penalty that rises with the number of regressors included in the model. A Wald chi-square test was used to determine significant differences between coefficients (Davidson & MacKinnon, 1993, p. 278).

The sets of predictor variables were entered mostly in chronological order into a hierarchical regression (J. Cohen & Cohen, 1983) for the TI outcome variables, post-TI TOT ratings, and post-TI problem solving. Because the teacher variable could affect all subsequent predictors, it was entered first. Next, the pre-CL variables—highest math grade in group and mean math grade of group—were entered, followed by the pre-TI variables: pre-TI TOT rating, overall progress, and pre-TI problem solving. Subsequently, TI initiation variables were entered (i.e., students' specific start and students' general start). The TI variables followed. Because the teacher probably reacted to students' work before determining their help content (Chiu, 2000), teacher evaluations (percentages of agreement and disagreement) were entered first. Teacher help content was entered next, followed by percentages of teacher questions and teacher commands. Finally, because post-TI TOT could affect the post-TI problem-solving variable, it was entered in the last step of the regression analysis focusing on this variable.

Because effects could vary across subsamples, significant differences among groups were tested through the addition of an interaction term: Subsample Dummy Variable × Predictor. The different subsample groups included the following: (a) teachers (Teacher 1 vs. Teacher 2), (b) different classes, (c) groups with different levels of PS progress before the TI, and (d) groups categorized according to problem-solving success (i.e., groups that solved the problem correctly vs. those that did not). In the case of multiple subsamples, such as different classes, subsamples that did not show different effects were pooled together.

A nested hypothesis test (chi-square log-likelihood) was used to assess whether each added set of variables was significant (Judge, Griffiths, Hill, Lutkepohl, & Lee, 1985, pp. 184–187). I used Hochberg's (1988) variation on Holm's (1979) method to address the likelihood of spurious correlations caused by conducting a large number of tests on a single data set. Non-significant variables were removed from each set to reduce the risk of Type I error. The hierarchical regression results incorporated both unstandardized and standardized regression coefficients.

Path analysis results served as the initial candidate for an SEM that assessed goodness of fit using the LISREL program (Joreskog & Sorbom, 2001).[1] Because some variables were not continuous, polychoric and polyserial correlations, along with their asymptotic covariances, were used to create the SEM via weighted least squares. The SEM results included standardized coefficients to simplify comparisons of the effects of different predictors. The results reported also include reduced form squared multiple correlations designed to estimate explained variance.

A similar analysis was conducted at the group level to assess solution scores. Because the teacher variable could affect all subsequent predictors, it was entered first. Then the group's mean math grade and the highest math grade among its members were entered. Next, mean TOT ratings and TIs were entered together.

With the exception of the SEMs, the computations just described were performed with the E-views statistical software (Lilien, Startz, Ellsworth, Noh, & Engle, 1995). An alpha level of .05 was used in all of the statistical tests conducted. (Because there were too few classes [six] and too few interventions per group [less than two on average], multilevel analyses [Bryk & Raudenbush, 1992; Goldstein, 1995] were not used.)

Results

This section begins with three transcript segments illustrating the relationships among different variables during TIs. These three segments comprise two TIs initiated by a student, one asking for general help and one asking for specific help, and a TI initiated by the teacher. Quantitative analyses were conducted at the TI level and then at the group outcome level. Only the main results are included here (all robustness tests and ancillary results are available upon request from the author).

Transcript Examples of TIs

In the first segment, a group of students who had not worked on the problem began a TI with a general question. (All names used in transcripts are pseudonyms.)

Dan: What are we supposed to do?
Ada: Do the problem.

Dan: How?
Ada: [shrugs her shoulders]
Kay: Ms. T! We have a question! [Ms. T walks over] What are we supposed to do?
Ms. T: What does the problem ask you to do?
Kay: [looks at problem description, reads] How many minutes should you use each month so that Speedy costs less? Explain.
Ms. T: And how do you do that?
Kay: Figure out how much Speedy and Quickie cost?
Ms. T: Uh-huh. [Ms. T walks away]

The teacher, Ms. T, responded to Kay's general question ("What are we supposed to do?") by directing her to the problem ("What does the problem ask you to do?"). After Kay picked out the question ("How many minutes . . .?"), Ms. T asked for a plan to answer it ("How do you do that?"). When Kay proposed two subgoals ("how much Speedy and Quickie cost"), Ms. T affirmed Kay's idea and left, probably satisfied that this group could begin doing the problem on its own.

After Ms. T leaves, the group hesitantly works on the problem.

Ada: So how much does Speedy cost?
Kay: [reads] Speedy charges twenty-five dollars per month and five cents per minute.
Ada: How many minutes do we use?
Dan: Um, one?
Kay: So for one minute, Speedy. . . .

Ada picked one subgoal ("How much does Speedy cost?"), and Kay read the relevant facts. Ada noticed the unknown number of minutes and specified another subgoal ("How many minutes do we use?"). Dan proposed 1 minute, which Kay readily accepted. Then she began computing the cost of Speedy for 1 minute.

In the second segment, a student started the TI with a specific question. The teacher responded to the students' concerns and bolstered their autonomy through the use of low help content and questions (rather than commands).

Eva: Yeah, Speedy always costs less.
Xia: That seems too easy. Ms. T! Ms. T! [Ms. T walks over]
Eva: Ms. T, if this [Quickie] costs twenty cents a minute and this [Speedy] costs only five cents a minute, won't it [Speedy] always be cheaper?
Ms. T: No, not necessarily. What does this say? [points to Quickie sentence]
Eva: Quickie charges ten dollars per month and twenty cents per minute.
Bill: Oh, so it's ten-twenty [10.20] and twenty-five oh five [25.05].
Ms. T: OK. So, when is Speedy going to cost less?
Bill: Like a thousand minutes?
Ms. T: You can try that [leaves].
Eva: Do that on the calculator. A thousand times twenty plus ten.

Eva started the TI with a specific question about their answer. Ms. T answered their question ("no") but did not explain the cause of the error (low teacher help content). Instead, Ms. T pointed out a key fact (i.e., she pointed to the Quickie sentence) that the students had overlooked by asking the group to read a part of the problem statement ("What does this say?"). After the students recognized their omission, Ms. T asked them how to proceed rather than telling them to follow a specific solution path (low help content). When Bill suggested an answer, Ms. T let them try it. After Ms. T left, the students continued their problem solving.

Eva: Do that on the calculator. A thousand times twenty plus ten.
Bill: That's twenty thousand ten [20,010].
Sean: [taps calculator keys] Yep, twenty thousand ten.
Eva: And Speedy [looks at problem description], um, a thousand, a thousand times five plus twenty-five.
Bill: Five thousand twenty-five [5,025].
Sean: [taps calculator keys] Yep, five thousand twenty-five.
Xia: So, now Speedy's cheaper.
Bill: Is that it? Are we done yet?
Xia: I don't think so.
Eva: Let's do, like two hundred.

[After doing the computations for "200 minutes," they discussed whether they had finished the problem until Ms. T returned. Ms. T explained that they needed to find the exact number of minutes at which Speedy is less costly. This group found the correct answer by computing the costs for 50, 100, and 101 minutes.]

The students then computed the costs for Quickie and Speedy for 1,000 and 200 minutes. However, the group faced another difficulty, deciding when they had solved the problem. Despite discussing this question at length, they could not agree. Later, Ms. T explained it to them, and the group eventually solved the problem.

In another group, Ms. T did not pinpoint the students' difficulty, and her use of high help content and commands intruded on the students' autonomy.

Li: Twenty-five dollars and five cents.
Pia: Adding more minutes [to Speedy] costs more, but we need to make it less. [17 seconds of silence; Ms. T walks by]
Ms. T: OK, start working on the problem. [None of the students have written anything]
Pia: Uh, Ms. T, we're not sure how to make Speedy less.
Ms. T: You can start with how much each company costs if you use zero minutes.
Li: Ten dollars for Quickie and twenty-five for Speedy.
Ms. T: OK, now try different numbers of minutes, try one minute.
Pia: Um, ten dollars and twenty cents and twenty-five dollars and five cents.

Ms. T: OK. Keep trying different numbers of minutes [leaves].
Pia: OK. [Ms. T leaves; 6 seconds of silence]
Pia: Uh, it's still going to get bigger, right?
Li: Mmm.
Maya: Zoe [other group member], you free tonight?

[Off-task conversation between Maya and Zoe followed. Despite another TI, neither the teacher nor the students diagnosed the group's misinterpretation of the problem.]

Before Ms. T walked by their group, the students said that the total cost of using Speedy for extra minutes had to be less than using Speedy for no minutes (instead of less than the total cost of using Quickie). When Ms. T walked by, the students had not written down any of their work and showed no visible signs of having started the problem. Rather than using questions to accurately assess their progress, she commanded them to begin working ("Start working on the problem"). Pia referred to their plight, but her words resembled a restatement of the problem. As a result, Ms. T did not respond to Pia's specific concern. Instead, she told the students to compute costs for various numbers of minutes (high help content, command). After the students computed these costs, Ms. T left. However, the students' problem remained, and they did not make further progress.

Summary of TIs

The 108 TIs ranged from 14 to 20 for each class (see Tables 2 and 3 and the appendixes). Students rarely sought help from the teacher, initiating a TI with a specific question 22 times and with a general question 6 times. In contrast, the teachers were very active, initiating 80 TIs (74% of all TIs; 1.45 TIs per group). The teachers intervened in every group, suggesting that the teachers were being fair and maintained relationships to facilitate future student requests for help. In the following transcript, for example, the teacher first spoke with this group during the CL activity with only 5 minutes left.

Jay: Twenty times a hundred is two thousand. [Ms. T walks over and looks at their work while Jay is talking]
Nina: Plus one thousand is three thousand. Right, now they're the same.
Ms. T: You all seem to be working well together. Keep on going.
Jay: OK, Ms. T. And Speedy has to cost more.
Nina: So, we add one more. [Ms. T walks away]
Jay: Right, so it's a hundred and one minutes.

This group worked together toward a correct solution and did not seem to need any help. Still, the teacher stopped by, urged them to continue their good work, and quickly left. Hence, the teacher's motive for this interaction seemed more social than instructive.

TIs varied in length (total word range: 9 to 175), ranging from short commands, such as "get to work," to explanations of a solution tactic. The teachers used evaluations often (77% of the time: 34% agreement, 43% disagreement), asked many questions (38%), and gave few commands (18%). (Interrater reliability values [Cohen's kappas] ranged from .86 to .93 [see Tables 2 and 3]. Sixteen minutes could not be coded because of poor sound quality; however, none of these 16 minutes occurred during the TIs.)

When Do TIs Occur?

Groups' TOT and past work predicted TIs. TIs often occurred when a group was more off-task than usual (see Table 4), regardless of the time period measured. Furthermore, the strongest effect was at the 1-minute time period, suggesting that the TIs were responses to groups' recent low TOT more so than their earlier TOT. Although students' TOT immediately before a TI was very low on average, the teachers did not primarily target low TOT groups or groups with low past mathematics grades for TIs. Groups' mean TOT, groups' mean mathematics grade, and groups' highest mathematics grade all failed to predict number of TIs ($\beta = -0.19$, $p = .66$; $\beta = -0.01$, $p = .24$; and $\beta = -0.02$, $p = .15$, respectively).

TIs often occurred when a group had done little work. When a group received a TI, it often showed little overall progress or recent problem solving (see Table 3, pre-TI overall progress and pre-TI problem solving). Before 92% of all TIs, the group had not discovered a correct solution method (Wilcoxon signed rank value = 4.8, $p < .001$). Before 41% of TIs, the group did not show correct understanding of the problem situation. Groups often showed less-than-usual recent problem solving before the TI (Wilcoxon signed rank value = 9.6, $p < .001$). In the case of 94% of TIs, the group had offered no correct new ideas in the most recent minute.

TI Effectiveness

Post-TI TOT and problem-solving levels generally exceeded pre-TI levels, regardless of the time period measured (see Tables 5 and 6). However, these

Table 4
Differences Between Group TOT Immediately Before an Intervention and Overall Mean TOT (*N* = 108)

Measured time period	Difference[a]	*SE*	*t*	*p*
1 minute	–0.273	0.028	9.84	.000
2 minutes	–0.255	0.027	9.26	.000
3 minutes	–0.244	0.027	9.22	.000
4 minutes	–0.206	0.025	8.15	.000
5 minutes	–0.159	0.023	6.85	.000

Note. Differences were measured, via *t* tests, during the 1- to 5-minute time period. TI = teacher intervention; TOT = time on-task.
[a]Pre-TI TOT – mean TOT.

Table 5
Effects of TIs on Group TOT (*N* = 108)

Measured time period	Post-TI TOT – Pre-TI TOT			
	Difference	*SE*	*t*	*p*
1 minute	0.192	0.054	3.54	.001
2 minutes	0.149	0.042	3.53	.001
3 minutes	0.123	0.036	3.43	.001
4 minutes	0.099	0.032	3.09	.003
5 minutes	0.082	0.027	3.04	.003

Note. Effects were measured, via *t*-test differences, during the 1- to 5-minute time period. TI = teacher intervention; TOT = time on-task.

effects waned (see Table 2, mean post-TI TOT for 1 to 5 minutes, and Table 3, mean post-TI problem solving for 1 to 5 minutes).

Explanatory Models of TI Effectiveness

The effects of TIs on groups' subsequent TOT and problem solving varied. Some patterns of teacher behaviors within TIs were more successful than others.

Table 6
Effects of TIs on Group Problem Solving (*N* = 108)

Measured time period	Problem solving	Mean rank	Median	No. of observations greater than overall median	Wilcoxon statistic	*p*
1 minute	Pre-TI	78.4	0	20		
	Post-TI	138.6	1	76		
	All	108.5	0	96	7.09	.000
2 minutes	Pre-TI	82.6	0	21		
	Post-TI	134.4	2	65		
	All	108.5	1	86	6.09	.000
3 minutes	Pre-TI	87.1	0	34		
	Post-TI	129.9	2	73		
	All	108.5	1	107	5.03	.000
4 minutes	Pre-TI	91.5	1	31		
	Post-TI	125.5	2	53		
	All	108.5	2	84	4.00	.000
5 minutes	Pre-TI	94.9	2	0		
	Post-TI	122.1	3	0		
	All	108 .5	3	0	3.21	.001

Note. Effects were measured, via Wilcoxon tests, during the 1- to 5-minute time period. TI = teacher intervention.

Predicting Post-TI TOT

The hierarchical set regressions showed that post-TI TOT tended to increase if students initiated a TI with a specific question, teachers showed proportionately more agreement, or teachers showed proportionately more disagreement (see Table 7). The first result bolstered the claim that students' specific questions helped their teacher pinpoint and address their needs, thereby encouraging them to continue working on the problem. Effects of teacher agreement and disagreement showed that students were more likely to work on the problem after teachers had evaluated their work than in the absence of evaluation. The effects of teacher agreements and disagreements also differed significantly, Wald $\chi^2(1, N = 108) = 9.87$, $p = .002$, suggesting that students were more likely to work on the problem after a teacher agreed rather than disagreed with them.

Meanwhile, some teacher actions tended to reduce post-TI TOT. Greater problem-solving content reduced post-TI TOT, supporting E. G. Cohen's (1994b) claim that brief teacher responses involving small levels of content foster student autonomy and interaction. In groups showing correct under-

Table 7

Hierarchical Regressions Predicting Post-TI Time On-Task Ratings at 1 Minute (*N* = 108)

Step and predictor(s)	*B*	*SE B*	β
Step 1			
Student specific initiation of TI	0.329	0.072	0.407***
Step 2			
Student specific initiation of TI	0.306	0.053	0.379***
Teacher agreement percentage	0.627	0.057	0.666***
Teacher disagreement percentage	0.451	0.088	0.338***
Step 3			
Student specific initiation of TI	0.236	0.046	0.292***
Teacher agreement percentage	0.491	0.053	0.522***
Teacher disagreement percentage	0.330	0.077	0.247***
Teacher help content	−0.152	0.023	−0.377***
Step 4			
Student specific initiation of TI	0.219	0.039	0.271***
Teacher agreement percentage	0.425	0.047	0.452***
Teacher disagreement percentage	0.206	0.068	0.154**
Teacher help content	−0.157	0.020	−0.389***
Teacher command percentage	0.122	0.080	0.115
Teacher Command Percentage × Pre-TI Progress	−0.472	0.084	−0.387***

Note. Each regression included a constant term. For Step 1, $R^2 = .166$, Akaike information criterion (AIC) = 0.620. For Step 2, $\Delta R^2 = .465$, ΔAIC = 0.779. For Step 3, $\Delta R^2 = .109$, ΔAIC = −0.330. For Step 4, $\Delta R^2 = .081$, ΔAIC = −0.338. TI = teacher intervention.
$^{**}p < .01$. $^{***}p < .001$.

standing of the problem situation (i.e., pre-TI progress > 0), greater percentages of teacher commands also reduced post-TI TOT, implying that teacher commands interfered with these groups' autonomous problem solving. For example, consider the following segment.

Oona: Speedy still costs more. [Ms. T, the teacher, walks over]
Coco: Use one million.
Oona: Quickie's like twenty million cents and ten dollars and Speedy's like five million cents and twenty-five dollars. Quickie's more. How about one thousand? We get twenty thousand cents and ten and five thousand cents and twenty-five. Quickie still costs more, but not as much now.
Ms. T: Yes. Try writing down all the total costs for both companies. It'll help you compare them. [Oona writes "Quickie 200.00 + 25 = 225 Speedy 50.00 + 10 = 60; Ms. T leaves]
Oona: Two twenty-five minus sixty is one sixty-five.
Coco: What does that tell us?
Oona: I don't know.
Ty: This problem is just too hard.
Guy: What time is it?

Although the students were converging on the correct answer, the teacher interrupted them and asked them to write down the costs. The students stopped and compared the costs by subtracting one from the other. The resulting value, 165, puzzled the students. Exploiting the lull in the problem solving, Ty and Guy discussed the movie they had seen the previous night.

However, commands did not significantly reduce post-TI TOT in groups showing inadequate understanding of the problem situation. Consider the following segment.

Qi: Sixty minutes per hour [writes "20 × 60"]
Raj: Times twenty-four hours a day times, uh, how many days a month? [Qi writes "× 24 ×"]
Yen: Say, did you hear about Charlie? [Charlie is not in this class. Schoolmate?]
Han: No, what?

[Off-task conversation for 1 minute 16 seconds; Ms. T walks over]

Yen: She's got to—Hi, Ms. T.
Ms. T: What do you have so far? [Everyone looks at Qi's paper]
Raj: Ms. T, how many days in a month do we use?
Ms. T: Days in a month? Let me see your work. [Qi hands Ms. T her paper] I don't think this'll work. Let's go back to the problem. You're trying to choose a phone company that costs less, right?
Yen: Uh-huh.

[All students nod]

Ms. T: How much is Quickie per month if you don't make any calls?
Qi: That's what we're trying to figure out.
Ms. T: Read this [points to Quickie sentence].
Qi: Ten dollars plus twenty cents per minute.

Ms. T: Assume that you make no calls.
Qi: Oh, ten dollars. And Speedy costs twenty-five dollars.
Ms. T: Then, assume that you make some calls and use some minutes.
Raj: How about ten minutes?
Ms. T: OK. Try that.
Qi: Ten times twenty cents is two dollars plus ten is twelve dollars [writes "10 × 20¢ = $2 + $10 = $12]. Speedy [Ms. T walks away] is twenty-five plus five times ten, fifty [writes $25 + 5¢ × 10 = $25 + 50¢]. Speedy costs more.
Raj: Do fifty minutes.

Ms. T evaluated and rejected the students' approach. Using a combination of commands and questions, she redirected the students to a different understanding of the problem. Unlike the earlier segment, the teacher's commands helped the students understand the problem situation correctly.

No other predictors significantly affected post-TI TOT. A group's past mathematics grades did not predict post-TI TOT, implying that higher achieving groups of students were neither more nor less likely to benefit from a TI. Moreover, pre-TI TOT did not significantly predict post-TI TOT, suggesting the malleability of students' TOT to TIs. Percentage of teacher questions also had no significant effect.

The impact of TIs on TOT waned over time. The sizes of the regression coefficients decreased as the time period moved from 1 minute to 5 minutes (student's specific start: .24 → .11; percentage of teacher agreements: .42 → .14; percentage of teacher disagreements: .21 → .08; teacher help content: −.16 → −.09; Percentage of Teacher Commands × Pre-TI Progress: −.47 → −.21; all effects remained significant). Likewise, explained variance (R^2) dropped somewhat, from 82% (for 1 minute) to 72% (for 5 minutes).

Predicting Post-TI Problem Solving

The factors that increased post-TI TOT typically increased post-TI problem solving (see Table 8). When the time period used was 1 minute, student initiation of a TI with a specific question, greater percentages of teacher agreement and disagreement, and greater post-TI TOT all increased post-TI problem solving. Unlike their effects on post-TI TOT, the effects of teacher agreements and disagreements on post-TI problem solving did not differ significantly, Wald $\chi^2(1, N = 108) = 2.06$, $p = .15$. Similarly, higher levels of teacher help content and higher percentages of teacher commands reduced post-TI problem solving.

The post-TI problem-solving results differed from the post-TI TOT results in two major ways. First, pre-TI progress positively predicted post-TI problem solving. Yet, mathematics grade did not. Hence, in terms of post-TI problem solving, students' overall progress on this specific problem was more important than their general past achievement. Second, higher percentages of teacher commands reduced post-TI problem solving regardless of students' pre-TI progress. No other predictors were significant.

Table 8
Hierarchical Ordered Logit Regressions Predicting post-TI Problem Solving at 1 Minute (*N* = 108)

Step and predictor(s)	*B*	*SE B*	β
Step 1			
Pre-TI overall progress	0.796	0.307	0.573**
Step 2			
Pre-TI overall progress	0.618	0.308	0.445*
Student specific initiation of TI	1.203	0.466	0.611**
Step 3			
Pre-TI overall progress	1.639	0.542	1.180**
Student specific initiation of TI	0.933	0.689	0.474
Teacher agreement percentage	7.036	1.523	3.075***
Teacher disagreement percentage	8.438	1.877	2.595***
Step 4			
Pre-TI overall progress	0.969	0.599	0.698
Student specific initiation of TI	1.297	0.713	0.659
Teacher agreement percentage	7.193	1.551	3.144***
Teacher disagreement percentage	7.158	1.794	2.201***
Teacher help content	−0.817	0.314	−0.834**
Step 5			
Pre-TI overall progress	0.674	0.659	0.486
Student specific initiation of TI	1.390	0.771	0.707
Teacher agreement percentage	9.008	1.837	3.937***
Teacher disagreement percentage	7.504	1.996	2.308***
Teacher help content	−0.786	0.351	−0.802*
Teacher command percentage	−5.986	1.421	−2.333***
Step 6			
Teacher agreement percentage	12.027	3.067	5.257***
Teacher disagreement percentage	10.243	2.839	3.150***
Teacher command percentage	−6.216	2.567	−2.423*
Post-TI time-on-task rating (1 minute)	15.232	3.670	6.263***

Note. For Step 1, McFadden R^2 = .029, Akaike information criterion (AIC) = 2.241. For Step 2, ΔMcFadden R^2 = .030, ΔAIC = −0.048. For Step 3, ΔMcFadden R^2 = .370, ΔAIC = −0.790. For Step 4, ΔMcFadden R^2 = .030, ΔAIC = −0.048. For Step 5, ΔMcFadden R^2 = .096, ΔAIC = −0.196. For Step 6, ΔMcFadden R^2 = .108, ΔAIC = −0.277. Prediction-expectation accuracy was 92%. TI = teacher intervention.
*$p < .05$. **$p < .01$. ***$p < .001$.

The impact of TIs on problem solving also declined. The sizes of the regression coefficients plummeted after the first minute and mostly had dissipated 5 minutes later (regression coefficients after 1 minute, 2 minutes, and 5 minutes, respectively, were as follows: teacher agreement percentage: 12.0, 3.3, and 0; teacher disagreement percentage: 10.2, 3.6, and 0; teacher command percentage: −6.2, −1.8, and −1.6; post-TI TOT: 15.2, 0, 0; a value of 0 indicated that the effect was no longer significant). Likewise, explained variance plummeted after the first minute (66% to 17% for 2 minutes and 11%

for 5 minutes). Prediction accuracy also fell (92% to 51% and 48%). Post-TI problem-solving analyses involving ordered Probit and ordered Gompit produced results similar to the ordered Logit results described earlier, but they indicated poorer fits (higher AIC).

TI Model

The path analyses involving the use of the 1-minute time period provided an initial SEM candidate. Removing nonsignificant variables yielded the final model shown in Figure 1. The SEM effects matched the earlier-described regression results and showed a good fit.[2] This SEM also accounted for most of the variance in both outcome variables (reduced form squared multiple correlations were .52 for post-TI TOT and .68 for post-TI problem solving). SEMs for 2- to 5-minute periods showed similar results.

Teacher evaluations, help content, and commands all affected students' post-TI behavior. Teacher agreements and disagreements had the most pronounced effects on students' post-TI TOT (standardized total effects of 0.63 and 0.48, respectively). Meanwhile, post-TI TOT had the most pronounced effect on post-TI problem solving (0.73), with levels of teacher agreement and disagreement also showing large overall effects (0.72 and 0.56, respectively).

Teacher evaluation directly affected all other teacher actions, bolstering the view that it served as a gatekeeper for other teacher actions. When these teachers evaluated student work, they were less likely to provide high help content (teacher agreement and teacher disagreement effects of −0.43 and −0.39 on help content). These results support the claim that understanding students' progress allowed the teachers to use their ideas and, hence, reduced the need for new information, explanations, or solution tactics. Furthermore, if a teacher used an evaluation, she tended not to use a command (teacher agreement and teacher disagreement effects of −0.28 and −0.38 on commands). When groups showed inadequate understanding of problem situations, the teachers also tended to use commands in providing help (+0.18), suggesting that they perceived these students as needing more direction and more task content information than other students.

Predicting Solution Score

The second set of analyses examined predictors of group solution score. Most of the groups (64%) correctly solved the problem. Of the 1,489 minutes of videotape coded for TOT, students were on-task 58% of the time and off-task 42% of the time.

After control for groups' highest past grade, TOT predicted solution score (see Table 9), consistent with past research showing that TOT often correlates with group outcomes (Ames & Ames, 1984, 1985; Forman & Cazden, 1985; Stipek, 1988). Aside from groups' highest past grade and TOT, no other predictor was significant. The results yielded insufficient degrees of freedom to test the model fit with an SEM. Ordered Gompit showed a better fit to the data than ordered Probit and ordered Logit (higher AIC); the latter two analyses showed similar results.

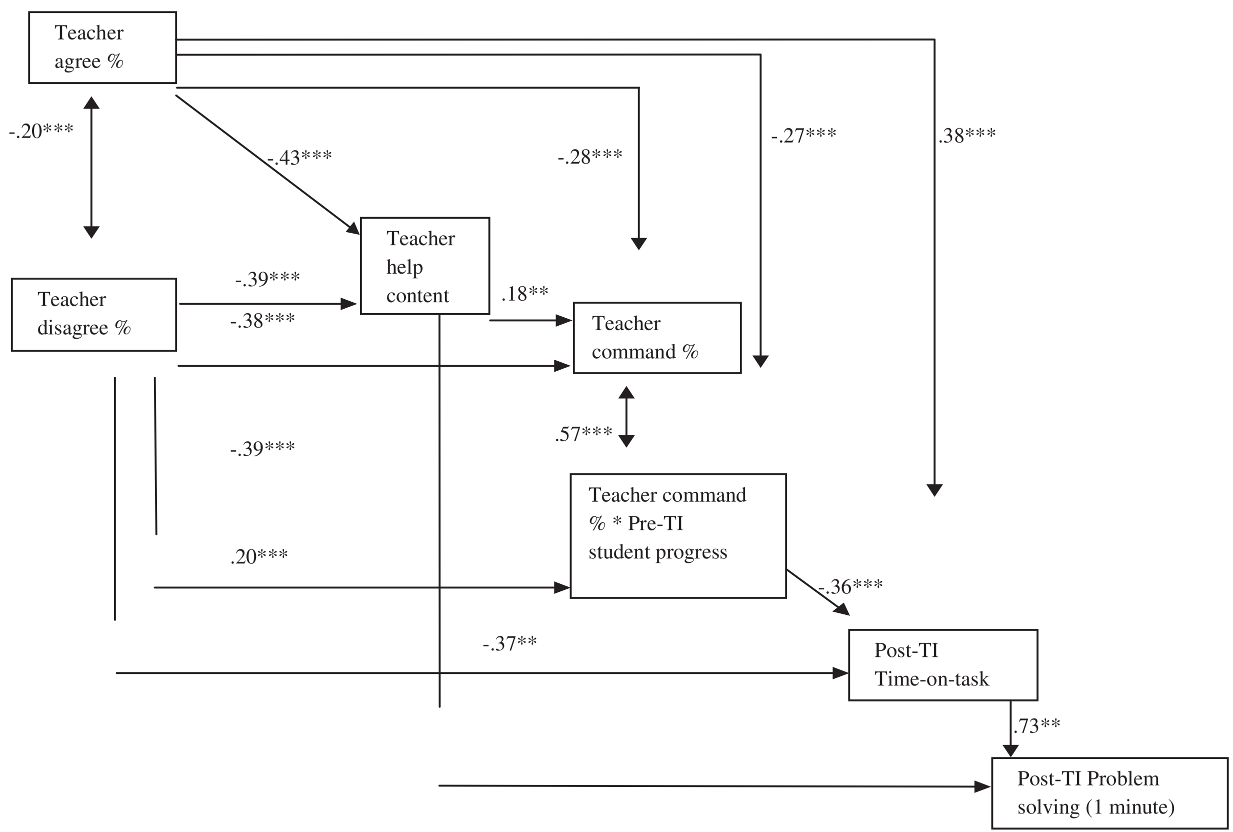

Figure 1. **Structural equation model showing standardized coefficients of effects of teacher interventions at 1 minute. *****p* **< .01, ******p* **< .001.**

Table 9
Hierarchical Ordered Gompit Regressions Predicting Solution Scores (*N* = 55)

Step and predictor(s)	*B*	*SE B*	β
Step 1			
Highest math score	0.181	0.036	1.141***
Step 2			
Highest math score	0.143	0.040	0.902***
Mean time-on-task rating	3.545	1.092	0.795**

Note. For Step 1, McFadden R^2 = .198, Akaike information criterion (AIC) = 2.127. For Step 2, ΔMcFadden R^2 = .088, ΔAIC = −0.181.
**$p < .01$. ***$p < .001$.

Discussion

In this study, I examined how TIs affect student behaviors during group problem solving. Past studies have shown that students can have difficulties even during well-planned CL lessons (e.g., Meloth & Deering, 1999; Webb & Farivar, 1999). If a teacher can identify and respond to students' needs, students might work together more productively and be more likely to solve the problem correctly than would otherwise be the case. Previous studies have explored the teacher's role during CL by examining a few interventions in detail (Meloth & Deering, 1999) or by analyzing the isolated properties of many interventions (e.g., Harwood, 1995). Extending this field of research, I created and systematically tested a model of TIs. Analyses of TIs with students during six classes revealed the powerful effects of teacher evaluations and the importance of attending to local group contexts.

TI Effects

The overall results showed that these teachers intervened effectively, consistent with results of other TI studies (E. G. Cohen, 1994b; Johnson & Johnson, 1994; Meloth & Deering, 1999). Students were more likely to be on-task after speaking with the teacher than before doing so. Furthermore, students were more likely to recognize their errors, develop new ideas, and explain ideas to one another. A second set of analyses showed that groups that were more often on-task usually provided better solutions than other groups. Together, these results show that teachers intervened effectively to improve students' short-term problem solving, which in turn tended to improve their group solutions.

The beneficial effects of TIs persisted for at least 5 minutes, but they faded over time. On-task behavior declined gradually, and problem-solving effectiveness plummeted after the first minute. The transcripts presented earlier illustrated students' subsequent behavior. Immediately after a teacher spoke with a group of students, the students usually talked about the problem and often generated new, correct ideas or at least recognized incorrect ones. As time passed, students' attention often gradually wandered to other topics, such as recent events in their personal lives. Likewise, students often capitalized on the teacher's comments to detect errors, correct them, or create new ideas. However, many groups soon returned to their earlier problem-solving routines. For example, the previously mentioned group that used a strategy of "do a computation and interpret" returned to that strategy after the teacher left. Although many of the interventions were successful, their declining effects over time indicate the importance of continuing to monitor students' progress.

TI Model

The results generally supported the relationships hypothesized in the TI model. Students' problem solving before the teacher's arrival at their table influenced whether the teacher intervened, how the intervention was initiated, the effects of teacher actions, and TI outcome. In theory, these teachers could have intervened at random times, or they could have targeted mostly groups with lower mathematics grades, but they did not do so. In practice, these teachers typically intervened when students were more off-task than usual, had made little overall progress, or had done little recent problem solving. These results suggest that the two participating teachers actively monitored students and based their intervention decisions on students' local behaviors (as advocated by E. G. Cohen, 1994a).

A few of the students initiated TIs. Some identified their plight and asked for specific help. The teacher usually answered these students' specific questions quickly, and they were more likely to continue working on the problem than other groups of students. In contrast, other students who did not identify their plight relied on their teacher to diagnose their needs. As shown earlier, the teacher's diagnosis was not always correct. These students were sometimes confused about how to proceed, and thus they were more likely to discuss off-task topics. This result suggests that when students want help, a teacher should encourage them to ask specific questions to express their needs.

Most students did not ask their teacher for help at all. Instead, they continued their personal off-task conversations or tried various computations to see whether the numbers would become meaningful. This failure to seek help is consistent with the results of past research on help seeking within groups (Nelson-Le Gall, 1992; Weiner, 1980). Because the teacher could not rely on students to ask for help, she had to determine when to intervene.

Effective Teacher Actions

Some teacher actions were more effective than others during TIs, and teacher evaluations explained most of the differences in groups' subsequent TOT and problem solving. During evaluations, the teachers tried to understand the students' work and, hence, diagnose their needs (Chiu, 2001). As a result, they made more informed decisions about the most suitable help needed by the students. In contrast, when a teacher did not evaluate students' work, she did not adapt to their specific needs. Instead, she might have relied on preconceived ideas about how students should proceed regardless of their progress.

Teacher evaluations during interventions served as a gatekeeper and directly affected subsequent teacher actions. In the present study, teachers, when evaluating student work, provided less help and issued fewer commands. Evaluations help teachers understand students' work; thus, teachers who make evaluations are more likely to help students use their ideas than to provide them with new information. When a teacher evaluated students' work, she tended to issue fewer commands. By working with students' ideas instead of simply telling students what to do, these teachers respect and validate students' ideas as worthy of consideration. In this study, such teacher actions might have bolstered students' autonomy and initiative (E. G. Cohen, 1994b) so that they were more likely to be on-task and to create new ideas.

Teacher actions do not necessarily have uniform effects. Rather, some teacher actions interact with local group contexts to yield different effects. When students already understand the problem situation, teacher commands tend to harm students' subsequent problem solving. In the present study, however, teacher commands had no effect on groups that did not grasp the problem situation. Perhaps teacher commands were perceived as less threatening to these students' initiative because they had no particular problem-solving approach. Thus, understanding the local group context could be vital to teachers in determining actions that will produce the desired effects.

Implications for Researchers

This study represented a step toward building a comprehensive understanding of TIs during CL. I sketched an initial model of TIs and illustrated analytical methods for these types of studies. The TI model outlined here identified important elements (local group context, TI initiation, and teacher actions) and relationships that researchers can consider when building a comprehensive theory of TIs. In particular, this study showed how teacher evaluations and local group context affected teacher actions and intervention outcomes.

This study also showed how to test a TI model systematically by using all interventions from each lesson. Using a single TI as the level of analysis, I conducted *t* tests and Wilcoxon tests on pre-TI and post-TI student behaviors to identify overall intervention effects. Then I used sequential set regressions to test the effects of each part of the model. Next, I ran a path analysis to create a candidate model for a final SEM analysis that tested all of the model's effects and interactions simultaneously.

Implications for Teachers

This study identified student actions and teacher actions that seemed to improve students' subsequent problem solving. Students did not always work on problems diligently and productively, and teachers monitored them to identify those who were off-task or showing little progress. After identifying such students, teachers can gain an understanding of their particular needs by encouraging them to ask specific questions.

In the present study, teacher evaluations of students' work, provision of low levels of help content, and use of a limited number of commands seemed to result in better student performance than other teacher behaviors. Evaluating students' work could help a teacher diagnose their needs, especially if students are unable to identify their concerns. Then a teacher could help students use their own ideas to address these concerns rather than simply giving them more problem-solving information. In this study, when students had a good grasp of the problem situation, the teacher did not need to explicitly tell them what to do.

Although many of the interventions observed here were successful, their effects faded over time. Thus, teachers should continue to monitor these students, especially if they have shown little overall problem solving progress prior to the intervention.

Limitations and Future Research

Because this study involved only two teachers and a single type of problem, the implications discussed here are subject to a replication of the results. Also, the two participating teachers had similar teaching experiences and discussed lesson plans together. Furthermore, the effects described earlier may occur only in specific contexts, such as among groups solving algebra problems.

Beyond replication with other participants, topics, and contexts, researchers can also consider how to help teachers evaluate students' immediate needs more accurately. This study showed the importance of teacher evaluations but did not consider their accuracy. Greater accuracy might help a teacher adapt an intervention more closely to students' needs and thereby help them improve their problem solving. Focusing on teachers' observation skills, knowledge of student errors, and questioning techniques might improve evaluation accuracy.

Conclusion

TIs initiated during CL helped improve subsequent TOT and problem solving among students who were off-task or making little progress. Students often did not ask for help when they needed it, so the teachers monitored their work and intervened as necessary. They typically intervened when groups were off-task or had shown little problem-solving progress. TIs were particularly successful when the teachers evaluated the students' work, provided lower levels of help content, and issued fewer commands. After interventions involving these teacher actions, students were more likely to be on-task and to show problem-solving progress. If teachers can accurately evaluate and adapt their interventions to students' needs, students might work together productively to realize the many potential benefits of CL.

Notes

This research was partially funded by a National Academy of Education/Spencer Foundation postdoctoral fellowship and a direct grant from the Chinese University of Hong Kong. I would like to thank Sung Wook Joh, Kit Tai Hau, Lawrence Khoo, Chang Lei, and Koanhoi Kim for their many helpful comments.

[1]Researchers have identified several robust goodness-of-fit measures using Monte Carlo simulation studies (Gerbing & Anderson, 1993; Hu & Bentler, 1995; Marsh, Balla, & McDonald, 1988; MacCallum & Hong, 1997). These measures include the comparative fit index (CFI), incremental fit index (IFI), Tucker–Lewis index (TLI; also known as the non-normed fit index; Tucker & Lewis, 1973), and root mean square error of approximation (RMSEA). Hu and Bentler (1999) showed that using a combination of one of these fit indexes and the standardized root mean residual (SRMR) tends to minimize Type I and Type II errors under many conditions. For SRMR, a value of .08 or less indicates a good fit; for RMSEA, .06 or less indicates a good fit; and, for TLI and IFI, .96 or higher indicates a good fit.

[2]See Note 1; SRMR = .03, RMSEA = .01, CFI = 1.00, IFI = 1.00, TLI = 1.00.

References

Aitchison, J., & Silvey, S. D. (1957). The generalization of Probit analysis to the case of multiple responses. *Biometrika, 44,* 131–140.

Ames, C., & Ames, R. E. (1984). *Research on motivation in education: Vol. 1. Student motivation.* San Diego, CA: Academic Press.

Ames, C., & Ames, R. E. (1985). *Research on motivation in education: Vol. 2. The classroom milieu.* San Diego, CA: Academic Press.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage.* New York: Cambridge University Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Chiu, M. M. (2000). Group PS processes: Social interactions and individual actions. *Journal for the Theory of Social Behavior, 30,* 27–50.

Chiu, M. M. (2001). Analyzing group work processes: Towards a conceptual framework and systematic statistical analyses. In F. Columbus (Ed.), *Advances in psychology research* (Vol. 4, pp. 193–222). Huntington, NY: Nova Science.

Chiu, M. M., & Khoo, L. (2003). Rudeness and status effects during group problem solving: Do they bias evaluations and reduce the likelihood of correct solutions? *Journal of Educational Psychology, 95,* 506–523.

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

Cohen, E. G. (1994b). *Designing groupwork: Strategies for the heterogeneous classroom* (2nd ed.). New York: Teachers College Press.

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

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics.* Oxford, England: Oxford University Press.

Finney, D. J. (1971). *Probit analysis* (3rd ed.). New York: Cambridge University Press.

Forman, E., & Cazden, C. B. (1985). Exploring Vygotskian perspectives in education: The cognitive values of peer interaction. In J. Wertsch (Ed.), *Culture, communication and cognition* (pp. 323–347). Cambridge, England: Cambridge University Press.

Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 40–65). Newbury Park, CA: Sage.

Goldstein, H. (1995). *Multilevel statistical models.* Sydney: Edward Arnold.

Good, T., Mulryan, C., & McCaslin, M. (1992). Grouping for instruction in mathematics: A call for programmatic research on small-group processes. In D. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 165–196). New York: Macmillan.

Grasa, A. A. (1989). *Econometric model selection: A new approach*. Dordrecht, the Netherlands: Kluwer.
Hacker, D. J. (1998). Definitions and empirical foundations. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 1–23). Mahwah, NJ: Erlbaum.
Harwood, D. (1995). The pedagogy of the World Studies 8–13 Project: The influence of the presence/absence of the teacher upon primary children's collaborative group work. *British Educational Research Journal, 21,* 587–611.
Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75,* 800–802.
Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandanavian Journal of Statistics, 6,* 65–70.
Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). London: Sage.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.
Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone: Cooperative, competitive, and individualistic learning* (4th ed.). Boston: Allyn & Bacon.
Joreskog, K., & Sorbom, D. (2001). *LISREL*. New York: Scientific Software International.
Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T. C. (1985). *The theory and practice of econometrics* (2nd ed.). New York: Wiley.
Laughlin, P. R., VanderStoep, S. W., & Hollingshead, A. B. (1991). Collective versus individual induction: Recognition of truth, rejection of error, and collective information processing. *Journal of Personality and Social Psychology, 61,* 50–67.
Lilien, D. M., Startz, R., Ellsworth, S., Noh, J., & Engle, R. (1995). *Eviews*. Irvine, CA: Quantitative Micro Software.
MacCallum, R. C., & Hong, S. (1997). Power analysis in covariance structure modeling using GFI and AGFI. *Multivariate Behavioral Research, 32,* 193–210.
Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness of fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391–410.
McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in economics* (pp. 105–142). New York: Academic Press.
McNeil, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
Meloth, M. S., & Barbe, J. (1992, April). *The relationship between instruction and cooperative peer-group discussions about reading comprehension strategies*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
Meloth, M. S., & Deering, P. D. (1999). The role of the teacher in promoting cognitive processing during collaborative learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 235–255). Mahwah, NJ: Erlbaum.
Nelson-Le Gall, S. (1992). Children's instrumental help-seeking: Its role in social acquisition and construction of knowledge. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 49–70). New York: Cambridge University Press.
Newman, F., & Thompson, J. A. (1987). *Effects of cooperative learning on achievement in secondary schools: A summary of research*. Madison: University of Wisconsin, National Center on Effective Secondary Schools.
Peterson, P. L., Wilkinson, L. C., Spinelli, F., & Swing, S. R. (1984). Merging the process-product and the sociolinguistic paradigms: Research on small-group

process. In P. L. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction* (pp. 126–152). Orlando, FL: Academic Press.

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

Sheskin, D. J. (1997). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FL: CRC Press.

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

Stipek, D. (1988). *Motivation to learn: From theory to practice*. Englewood Cliffs, NJ: Prentice Hall.

Tsui, A. B. (1994). *English conversation*. Oxford, England: Oxford University Press.

Tucker, L. R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10.

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

Webb, N., & Kenderski, C. M. (1984). Student interaction and learning in small group and whole-class settings. In P. L. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes* (pp. 153–170). Orlando, FL: Academic Press.

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: Macmillan.

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

Weiner, B. (1980). May I borrow your class notes? An attributional analysis of judgments of help giving in an achievement-related context. *Journal of Educational Psychology, 72,* 676–681.

Wilkinson, L. C. (1985). Communication in all-student mathematics groups. *Theory into Practice, 24,* 8–13.

Manuscript received October 8, 2001
Revision received April 3, 2003
Accepted March 31, 2004

APPENDIX A

Polyserial Correlation Matrix for All Variables at the TI Level (N = 108)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Post-TI PS (1 minute)	1.00																
2. Post-TI TOT rating (1 minute)	.86	1.00															
3. Pre-TI overall progress	.32	.18	1.00														
4. Student specific initiation of TI	.42	.23	.63	1.00													
5. Teacher agreement percentage	.63	.35	−.25	.54	1.00												
6. Teacher disagreement percentage	.40	−.52	.57	−.20	.36	1.00	1.00										
7. Teacher help content	−.62	−.38	−.41	−.37	−.29	−.69	−.61	1.00									
8. Teacher command percentage	−.68	−.06	−.48	−.27	−.38	.44	.78	−.58	1.00								
9. Teacher Command Percentage × Pre-TI Progress	−.58	.22	−.32	−.20	−.33	.24											
10. Post-TI TOT rating (2 minutes)	.84	.18	.70	.47	.39	−.68	−.59	−.56	.97	1.00							
11. Post-TI TOT rating (3 minutes)	.85	.17	.71	.42	.39	−.68	−.56	−.54	.95	.97	1.00						
12. Post-TI TOT rating (4 minutes)	.84	.18	.67	.41	.39	−.69	−.53	−.52	.92	.94	.97	1.00					
13. Post-TI TOT rating (5 minutes)	.78	.17	.65	.42	.36	−.71	−.52	−.50	.89	.91	.94	.97	1.00				
14. Post-TI PS (2 minutes)	.76	.10	.19	.43	.34	−.45	−.45	−.29	.62	.57	.57	.56	.53	1.00			
15. Post-TI PS (3 minutes)	.78	.02	.07	.44	.37	−.52	−.49	−.34	.62	.58	.58	.55	.52	.94	1.00		
16. Post-TI PS (4 minutes)	.78	.03	.22	.37	.40	−.52	−.52	−.34	.61	.57	.57	.53	.51	.90	.96	1.00	
17. Post-TI PS (5 minutes)	.70	−.08	.25	.28	.41	−.40	−.52	−.36	.55	.52	.50	.47	.46	.83	.89	.94	1.00

Note. TI = teacher intervention; PS = problem solving; TOT = time on-task.

APPENDIX B

Polyserial Correlation Matrix for Variables at the Group Level (*N* = 55)

Variable	1	2	3	4
1. Highest math score	1.00			
2. Mean math score	.73	1.00		
3. Mean time-on-task rating	.55	.43	1.00	
4. Solution score	.67	.69	.69	1.00