Short Circuits or Superconductors? Examining Factors That Encourage or Undermine Group Learning and Collaboration Among High-Ability Students

CSE Technical Report 541

Noreen M. Webb, Mari Welner, and Stephen Zuniga
CRESST / University of California, Los Angeles

May 2001

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 1.1 Models-Based Assessment Design: Individual and Group Problem Solving—Science,
Noreen Webb, Project Director, CRESST/UCLA

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

SHORT CIRCUITS OR SUPERCONDUCTORS? EXAMINING FACTORS THAT ENCOURAGE OR UNDERMINE GROUP LEARNING AND COLLABORATION AMONG HIGH-ABILITY STUDENTS

Noreen Webb, Kariane Mari Welner, and Stephen Zuniga
CRESST/University of California, Los Angeles

Abstract

Peer-directed small-group collaboration is featured prominently in debates about good classroom instruction and in the promotion of school reform. Although many cooperative learning methods advocate grouping students heterogeneously in order to maximize the diversity of perspectives, skills, and backgrounds, past research shows that while heterogeneous grouping generally benefits low-ability students, it does not necessarily benefit high-ability students. This study investigates the effects of group ability composition (homogeneous vs. heterogeneous) on group processes and outcomes for high-ability students completing science performance assessments. The results show that group ability composition does not have straightforward effects on achievement. While high-ability students working in homogeneous groups uniformly performed well, high-ability students in some heterogeneous groups performed better than high-ability students in other heterogeneous groups. The quality of group functioning served as the strongest predictor of high-ability students' achievement. High-ability students in groups that were responsive to group members' need for help and did not engage in debilitating social behavior performed well, whereas high-ability students in groups with poorer functioning did not. Whereas homogeneous groups consistently showed beneficial group functioning, only some heterogeneous groups exhibited these traits. These results show that achievement of high-ability students cannot be predicted from a simple homogeneous-heterogeneous grouping contrast and that the level of group functioning provides the key to understanding group performance.

Debates about good classroom instruction and the promotion of school reform prominently feature peer-directed small-group collaboration. Group collaboration can increase learning and achievement, social skills, self-esteem, and attitudes toward classmates and school. Past research demonstrates that lower ability students usually benefit from collaborative learning, at least when working in groups heterogeneous in ability or achievement. In such groups, lower ability students can learn from their more able peers. The benefits of group collaboration for high-ability students, however, are less clear. Previous research on the benefits of collaborative work for high-ability students and the relative advantages of different group compositions for these students yields inconsistent results. Moreover, current research tells us little about the processes operating in collaborative groups that may work to the benefit or detriment of high-ability students' learning. The present study, then, examined high-ability 8th-grade science students' performance during and after solving electric circuit building problems individually and/or in collaborative groups. This study specifically investigated (a) group processes that positively or negatively influenced the achievement of high-ability students, and (b) the influence of group composition on group processes and achievement outcomes.

Group Processes Hypothesized to Promote Learning

Piaget's model of socio-cognitive conflict and learning provides a useful theoretical perspective regarding the mechanisms by which learning occurs within group contexts. Interaction with others may produce discrepancies between a child's views and new information, giving rise to cognitive conflict within the child, and leading the child to try out new ideas. Internal cognitive conflict may manifest itself in interaction with others through overt conflict or controversy in which individuals recognize the differences between their beliefs. Overt conflict encourages individuals to explain and justify their own positions, raises uncertainties about their beliefs, encourages individuals to seek new information to help resolve their disagreements and arguments, and helps them understand alternative points of view (Brown & Palincsar, 1989; Johnson & Johnson, 1979).

The co-construction of knowledge that some or all students did not have prior to collaboration provides a second way students can benefit from group work (Damon & Phelps, 1989). When a group has a student with high subject-matter knowledge, that student can help less knowledgeable peers carry out tasks or solve problems that they could not perform without assistance, a process called scaffolding (Wood, Bruner, & Ross, 1976). During this process, the learners practice skills that they then internalize, which increase their understanding and enable them to accomplish more individually (Tudge & Rogoff, 1989; Vygotsky, 1981). Even if a group contains no expert, students can build on each other's knowledge and understanding to solve problems and develop new knowledge that none could accomplish alone (Forman & Cazden, 1985; Saxe, 1992).

Students working collaboratively can also learn by helping each other. From a theoretical perspective, both the help-giver and the help-receiver stand to benefit from elaborated help, that is, detailed explanations about how to solve problems or complete tasks. Giving explanations encourages the explainer to clarify and reorganize the material in new ways in order to make it understandable to others (Bargh & Schul, 1980). This cognitive restructuring may help the explainer to understand the material better, develop new perspectives, and recognize and fill in gaps in her understanding. By tailoring explanations to the difficulties of other students, helpers may construct more elaborate conceptualizations than they would when solving the problems for themselves (Benware & Deci, 1984; Yackel, Cobb, & Wood, 1991). Receiving explanations can benefit the receiver by filling in gaps in her understanding, correcting misconceptions, and strengthening connections between new information and previous learning (Mayer, 1984; Wittrock, 1990). Students have the potential to explain effectively because peers share a similar language and can translate difficult vocabulary and expressions into words that fellow students can understand (Noddings, 1985). Further, because explainers and other students often learn material at the same time, this enables explainers to tune into other students' misconceptions and give relevant explanations (Vedder, 1985). Students can also control the pace of explanations (as well as the pace of group work) to better promote understanding. Students experience the benefits of receiving explanations most strongly when receivers have, and use, opportunities to apply the explanations to solve problems or perform tasks for themselves (Vedder, 1985; Mastergeorge, Webb, Roc, & Baure, 2000; Webb, Troper, & Fall, 1995). Carrying out further activity after receiving explanations may (a) help learners internalize principles and construct-specific inference rules for solving problems (Chi & Bassock, 1989; Chi, Bassock, Lewis, Reimann, & Glaser, 1989), (b) help learners monitor their own understanding (Chi & Bassock, 1989), and (c) help make the group aware of a student's continued misunderstandings (Shavelson, Webb, Stasz, & McArthur, 1988).

The beneficial social-emotional processes that take place in collaborative groups can also enhance learning, especially if students receive incentives to cooperate, such as rewards for high group performance or effort. According to the motivational theory of Deutsch (1949), when groups work toward a common goal, students will praise, encourage, and support each other's efforts, resulting in greater effort and an increased liking of the task and other students (see also Johnson & Johnson, 1985).

However, not all groups function in ways that promote learning. A variety of debilitating processes may inhibit group functioning. For example, a number of studies found that receiving non-elaborated help (e.g., the answer without any description of how to obtain it) in response to questions or errors negatively affects achievement (see Webb & Palincsar, 1996). Another debilitating process concerns the failure of students to seek help when they need it. Insights into when students do and do not seek help come from Nelson-Le Gall's (1981, 1985, 1992; Nelson-Le Gall, Gumerman, & Scott-Jones, 1983) comprehensive, five-step model of children's help-seeking. For students to benefit from group work, they must realize that they need help, be willing to seek help, identify someone who can provide help, use effective strategies to elicit help (e.g., ask explicit, precise, and direct questions; Peterson, Wilkinson, Spinelli, & Swing, 1984; Wilkinson, 1985; Wilkinson & Calculator, 1982a, 1982b; Wilkinson & Spinelli, 1983; see also Webb et al., 1995; Webb & Kenderski, 1984), and be willing to reassess their strategies for obtaining help.

Debilitating socio-emotional processes can also occur within groups, decreasing the chance that students will benefit from the group-work experience. Salomon and Globerson (1989) observed four such processes among pairs of students working on reading and writing tasks at the computer. In the "free rider" effect (see also Kerr & Bruun, 1983), also called "diffusion of responsibility" (Slavin, 1990), some group members sat back and let others do the work. This sometimes produced a "sucker effect" in which the group members doing all of the work discover that their fellow students took them for a free ride and therefore start to contribute less to group work to avoid the label of "sucker." The dominance of some group members over others constituted a third effect, with high-status students having undue influence over group functioning, thereby preventing low-status students from making contributions or obtaining the help that they needed (see Dembo & McAuliffe, 1987). Finally, Salomon and Globerson observed some pairs deciding to go through the motions without actually performing the assigned task, feigning busyness. Deering (1989) observed additional disabling processes, such as aggressiveness and hostility leading to unconstructive and bitter arguments, and passivity and acquiescence leading to premature agreement on answers.

Nearly all research investigating the relationships between group processes and learning outcomes combines all ranges of students' ability. Therefore, this body of work tells us little about the processes that specifically predict high-ability students' learning outcomes. Yet, the small amount of research that focuses specifically on high-ability students demonstrates similar patterns in what makes for a positive group experience. Swing and Peterson (1982) found that on two out of three outcome measures, giving conceptual or sequencing explanations related positively, and carrying out off-task behavior related negatively, to high-ability students' achievement.

Achievement of High-Ability Students in Homogeneous and Heterogeneous Groups

Teachers often ask how they should assign students to collaborating groups within their classrooms. Many cooperative learning methods recommend that teachers form groups heterogeneously to reflect the diversity of ability, gender, and ethnic backgrounds in the class. The reasons for creating such groups include maximizing the opportunities for peer tutoring and support, improving cross-race and cross-gender relations, ensuring that every group has at least one student who can do the work (Kagan, 1992), and making groups comparable for fair inter-group competition (Slavin, 1990). While much research has found that lower ability students gain more from working in heterogeneous groups than working in homogeneous groups or individually (see Webb & Palincsar, 1996), fewer studies have conducted these comparisons for high-ability students, where researchers typically define high-ability as the top 25% or 33% of a population or sample. Studies comparing individual and group work for these high-ability students yield inconsistent results. In the comparison of homogeneous grouping, heterogeneous grouping, and individual work from which the present study stems, Webb, Nemer, Chizhik, & Sugrue (1998) found that high-ability students performed better in homogeneous groups than in heterogeneous groups, with high-ability students working individually performing in the middle (but not significantly different from students working in either heterogeneous or homogeneous groups). Two other studies, in contrast, found advantages of heterogeneous grouping over individual work for high-ability students for one out of three computer learning posttests (Hooper, Temiyakarn, & Williams, 1993) and on both outcome measures of reading recall and higher level reasoning (Johnson & Johnson, 1993).

A number of other studies compare achievement of high-ability students in heterogeneous and homogeneous groups. In a large majority of these studies, high-ability students performed equally well on achievement tests after working in heterogeneous and homogeneous groups (e.g., Armstrong, 1994; Azmitia, 1988; Carter & Jones, 1994; Hooper & Hannafin, 1988, 1991; Hooper, Ward, Hannafin, & Clark, 1989; Jones & Carter, 1994; Jones, Carter, & Rua, 1998; Lawrenz, & Munch, 1985; Melser, 1999; Sherman & Klein, 1995; Skon, Johnson, & Johnson, 1981). Most of the remaining studies found higher achievement in homogeneous groups than in heterogeneous groups on some or all outcomes (Baron, 1994; Fuchs, Fuchs, Hamlett, & Karns, 1998; Knupfer, 1993). A notable exception lies in a study that demonstrated that high-ability students learned more in heterogeneous groups than in homogeneous groups (Webb, 1980) because, in heterogeneous groups, high-ability students often assumed the role of teacher and explained the material to other group members (typically the low-ability student). In homogeneous groups, high-ability students exchanged relatively few explanations, apparently because they assumed that everyone could master the material without help.

Some studies reported differences in the behavior of high-ability students in the two group compositions, even when no differences in achievement appeared. Carter and Jones (1994) found that high-ability students in heterogeneous pairs working on activities related to balance interacted in more productive ways than those in homogeneous pairs. The high-ability students in heterogeneous pairs spoke more, took more turns speaking, and exhibited more helping behavior. Similarly, Jones and Carter (1994) found that high-ability students in heterogeneous pairs modeled thinking processes and strategies for manipulating equipment, and guided and monitored low-ability students' progress. Fuchs, Fuchs, Hamlett, and Karns (1998), in contrast, found that high-ability students in heterogeneous pairs did not adopt a teaching role and, for the most part, did the work themselves, possibly because the task proved too complex to involve low-ability students' participation.

The participants in the present study came from a project investigating the effects of group ability composition on group processes and outcomes in science performance assessments (Webb et al., 1998). In that study, students in 21 eighth-grade classes worked on a science assessment in heterogeneous groups, homogeneous groups, or individually, and subsequently worked on another science assessment individually. Below-average students who worked with above-average students showed higher achievement than did below-average students who worked in groups without above-average students. Below-average students learned from the more accurate and higher quality answers and explanations given in heterogeneous groups. Yet unlike their low-ability peers, high-ability students working in homogeneous groups outperformed high-ability students who worked in heterogeneous groups. The analyses of behavior and experiences of high-ability students in that study, however, failed to uncover explanations for the differences in achievement across group compositions. For example, the quality of the group discussion did not seem to correlate with outcome measures for high-ability students. Furthermore, high-ability students did not engage in more high-level participation (e.g., making suggestions, paraphrasing others' suggestions) in one group composition than in the other, a variable that positively related to learning outcomes in previous research (e.g., Webb et al., 1995). To understand the performance results found in the previous study, the present study focused exclusively on the experiences and outcomes of high-ability students and carried out intensive analyses of high-ability students' interactions with other group members.

Method

Sample

The sample in the original study (Webb et al., 1998) consisted of 662 seventh-grade and eighth-grade students (21 classes) from five schools in Los Angeles County. The present study focuses on the top 25% of that sample, designated as high-ability (described in more detail below). Two schools had small numbers of high-ability students, and no homogeneously high groups, so we did not include them in the analyses. A third school had a small number of homogeneous and heterogeneous groups, but because its overall performance ranked considerably below the two remaining schools, we did not include it in the analyses. The sample for analysis, then, consisted of the high-ability students from the remaining two schools (9 classes; n = 99). Of these students, 83 students worked in 47 three-person groups, and 16 students worked individually. The samples from the two schools were predominantly White (89%), with small numbers of Black students (2%), Hispanic students (2%) and Asian students (6%). Females comprised approximately half of the sample (49%).

Procedures

At the beginning of the study, prior to instruction in electricity and electric circuits, all students completed three pretests: vocabulary (Vocabulary Test V-1, Buros, 1974), verbal reasoning (the New Jersey Test of Verbal Reasoning Form B, Shipman, 1985), and nonverbal reasoning (Raven's Progressive Matrices). After the administration of these tests, all teachers conducted a 3-week unit on electricity and electric circuits in their classrooms. We gave the teachers the freedom to teach the topic in their own way, using their own instructional materials. We told teachers that the assessments administered at the end of instruction would target students' knowledge of the concepts of voltage, resistance, and current, as well as the relationships among them, but teachers did not know in advance what form our evaluation of the students' knowledge of these concepts would take.

At the end of the instructional unit, students completed a hands-on and a written assessment measuring their understandings of voltage, resistance, current, and the interrelationships between these concepts. Students finished both activities individually, without help from other students or from the teacher. The hands-on component required students to assemble manipulatives (batteries, bulbs, wires, and resistors) into electric circuits, draw diagrams of their circuits, and then answer questions (multiple choice and written explanations). Proctors videotaped all assembled circuits. A subset of the paper-and-pencil written component consisted of items analogous to the hands-on element (see Appendix) but required students to draw the circuits instead of assemble them. The other items on the written component did not have direct analogues in the hands-on component, and therefore we did not include them in the present analyses.

One month later, with no intervening instruction or review, students re-completed the same two science assessments (hands-on and written). Students completed the hands-on component first and the written module the next day. For the hands-on component, approximately 80% of the students in each class worked in collaborative three-person groups; as a control, the remaining 20% of the students worked individually at separate desks with no interaction with others. All students completed the written component individually.

When students worked in collaborative groups, each group received one set of manipulatives (batteries, bulbs, wires, resistors), but students had to complete their own written responses (although they could work together on composing the answers). Students received instructions to work together to complete the tasks, to help each other understand how to solve the problems, to ask questions of each other, and to assist group members who seemed confused or appeared to need help. Of the seven or eight groups in each class, proctors videotaped five groups while they completed the test.

Design

The design, then, had three phases. In the first phase, students completed the verbal and nonverbal pretests, and the first administration of the hands-on and paper-and-pencil assessments individually. In the second phase, students completed the hands-on component again either in collaborative three-person groups or individually. In the third phase, all students individually completed the paper-and-pencil component again. Table 1 summarizes the design of the study.

Scoring of Science Achievement Assessments

Hands-on component. The hands-on component consisted of two tasks. For each task, students received a bag of materials containing 1.5-volt and / or 9-volt batteries, wires, bulbs, and graphite resistors (Task 1: two 9-volt batteries and two 1.5-volt batteries, three bulbs in bulb holders, and seven wires with alligator clips; Task 2: two 9-volt batteries, two bulbs in bulb holders, three graphite resistors, and seven wires with alligator clips). Students received instructions to assemble pairs of circuits so that the bulb in one circuit was brighter (or dimmer) than the bulb in the other circuit. After circuit construction, we asked students to draw diagrams of their circuits, explain why one circuit was brighter than the other, and answer three multiple-choice items about which of their two circuits had higher voltage, resistance, and current. Further, the test asked students to write an explanation to justify each multiple-choice answer (see Appendix).

Students assembled a variety of circuits. Because different pairs of circuits gave rise to different correct answers, we scored the multiple-choice items and the explanations according to the circuits that students assembled (as shown on the

Table 1

Design of Study

Phase 1	Phase 2	Phase 3
INDIVIDUAL	GROUP	INDIVIDUAL
Pretests	Hands-on science	Paper-and-pencil science
Verbal reasoning		
Vocabulary		
Nonverbal reasoning		
	OR	
Hands-on science	INDIVIDUAL	
Paper-and-pencil science	Hands-on science	

videotapes of their circuits). For example, if a student assembled two circuits that each contained a nine-volt battery and a 1.5-volt battery, both circuits would have equal voltage. If a student assembled one circuit with one 9-volt battery and two 1.5-volt batteries and the other with only one 9-volt battery, the first circuit would have the higher voltage.

Each multiple-choice item received a score of correct or incorrect (0,1) based on the circuits assembled. Coders scored justifications on a 0-to-1 scale according to accuracy and completeness. For example, when asked "Why was voltage in Circuit A higher than in Circuit B?" coders assigned the following scores:

- 1 if a student mentioned the relative number of batteries in the two circuits in addition to the relative power or voltage generated by the batteries;

- 0.67 if a student mentioned the relative number of batteries in the two circuits but not the strength or voltage of the batteries;

- 0.33 if a student mentioned batteries but not the relative number or relative strength of them; and

- 0 if the explanation was irrelevant or if it displayed confusion over cause and effect (for example, "the voltage is higher because it is brighter").

Two coders scored all justifications. The correlation between coders' ratings exceeded 0.95 for all justification items.

Paper-and-pencil written component. The portion of the paper-and-pencil written component presented in these analyses consists of items exactly analogous to the hands-on component except that students received pictures of the equipment (batteries, bulbs, resistors) instead of actual manipulatives (see Appendix). Otherwise, on this portion of the test the students received identical instructions. We asked students to construct two circuits (draw diagrams using the items given in order to make the bulb in one circuit brighter (or dimmer) than the bulb in the other) and answer the same multiple-choice and explanation questions.

Scores Used in the Analyses

Hands-on component score. The hands-on component score used for analysis consisted of the mean of the three multiple-choice items and four justification items for each of the two tasks. This totaled 14 items (equally weighted). We calculated one hands-on score for Phase 2 (internal consistency alpha = .86) and one for Phase 3 (alpha = .83).

Paper-and-pencil analogue score. As described above, we analyzed the subset of the written component analogous to the hands-on component. The only difference between these two modules was that students received pictures of the batteries, bulbs, and resistors instead of the actual equipment. We combined scores on the paper-and-pencil analogue items in the same way as the hands-on component scores to form a composite score (14 items, equally weighted). We calculated one written analogue score for Phase 2 (alpha = .85) and one for Phase 3 (alpha = .85).

Determining Ability Levels

We computed a composite score based on both the pretests and the science tests administered during Phase 1 to serve as a measure of students' competence prior to group-work. In the ability composite, we weighted the science scores more heavily than the verbal and nonverbal reasoning scores because, among students who worked individually in all phases of the study (the total sample), science scores in Phases 2 and 3 had a higher correlation with science scores in Phase 1 (*r*s ranged from .68 to .69) than with the verbal and nonverbal scores in Phase 1 (*r*s ranged from .40 to .60). We call the composite measure *ability* to distinguish it from the achievement scores in Phases 2 and 3. Although we use the term ability here, we should note that the composite combines science achievement, verbal and nonverbal reasoning ability, and vocabulary, and so it has a heavy achievement component.

We classified students as low, low-medium, medium-high, or high in ability based on the composite ability score. One quarter of the sample comprised each ability level.

Group Compositions

A variety of group compositions included high-ability students. For the present analyses, we describe the group compositions according to the lowest ability level in the group: low (6 groups), low-medium (10 groups), medium-high (23 groups), and high (8 groups). Proctors videotaped 6 low groups, 9 low-medium groups, 13 medium-high groups, and 8 high groups. Regardless of the ability mix in the group, we formed groups so that they would be heterogeneous on gender and ethnic background to reflect the class mix of student characteristics.

We used stratified random sampling from the different ability levels in order to select the 20% of students in each class who worked individually during the second phase. In the two schools analyzed here, 16 high-ability students worked individually during that phase.

Coding of Videotapes

Group discussions. Proctors videotaped groups as the students worked collaboratively on the hands-on component. Coders then assigned two group discussion scores using the same scoring scheme applied to the hands-on test: the accuracy of students' multiple-choice answers and the quality of their justifications. For each multiple-choice item on the test, coders gave the group a 1 if anyone in the group voiced the correct answer and a 0 otherwise. For each item, coders gave the group a score for the highest level of justification articulated based on the contributions of the three students in the group. For example, when looking at a justification about why one circuit had higher current than the other, if one student explained that Circuit A had higher current because it had more voltage and another student explained that lower resistance also led to higher current, coders combined the two justifications (one about voltage and one about resistance) to form a single justification that coders then scored. The justification score, then, reflected the maximum level of discussion in the group and could have reflected the contributions of one, two, or all three students.

To check interrater agreement, two coders scored the group discussion variables for 42 groups (2 per classroom). The raters' codes correlated at 1.00 for accuracy of answers and 0.99 for quality of justifications.

It should be noted that the coding of these group discussion variables differed somewhat from the coding in prior analyses (Webb et al., 1998). In the earlier study, if a group did not provide any verbal response for an item, we coded the accuracy of the response 0. In the present study, if a group did not provide a verbal response for an item, we did not code the accuracy; the code would be missing for that item. The accuracy of group discussion codes, then, reflects only the accuracy of the responses given verbally.

Individual behavior of students in groups. In order to evaluate how working in collaborative groups impacted students, we looked at four categories of student behavior: (1) the responsiveness of the high-ability student to other students' need for help (as indicated by questions, errors, or nonverbal signals of confusion), (2) the responsiveness of the other members of the group to the high-ability student's need for help, (3) the contributions made by the high-ability student in the absence of any indication of a need for help on the part of other group members, and

(4) debilitating processes occurring in the group (initiated or received by the high-ability student).

For categories 1 and 2, responsiveness to need for help, we took a two-pronged approach to coding group interactions. We coded the signal that indicated a need for help and coded the response given. Signals included nonverbal behavior (looking confused), general statements of confusion ("I don't know how to do this"), questions that signaled confusion, a lack of comprehension, or a need for information (requests for the answer, an explanation, information, or a justification for why someone gave the answer he or she did), and errors (incorrect multiple-choice answers, incorrect or incomplete justifications, incorrect suggestions for circuit building or diagram drawing). Responses included answers (multiple choice or justification), explanations, information, suggestions, corrections of error, agreement, a non-content response ("I don't know"), and negative responses (insulting, dismissing requests for help). We coded all responses for accuracy, completeness, level of elaboration, and whether the help-giver could have provided a better or more complete response based on his or her knowledge as demonstrated in the Phase 1 measures.

We then placed each combination of signal and response along two continua of responsiveness. The first continuum looked at the content of the response, specifically the accuracy and level of elaboration of the content provided in response to a need for help. The continuum ranged from +3 (fully accurate and elaborated) to –3 (inaccurate and elaborated). Descriptions of each point on the content continuum appear in Table 2. The second continuum took into account the motivation of the student in giving a particular response, that is, the intended helpfulness of the response from the perspective of the help-giver. We coded a response as high on the intended helpfulness continuum if the help-giver gave an elaborated response, whether or not it was correct. This continuum also ranged on a scale from +3 (highly responsive) to –3 (insulting a fellow group member). Descriptions of each point on the intended helpfulness continuum also appear in Table 2.

Category 3 consisted of high-ability students' behaviors when no one in the group overtly signaled a need for help. These behaviors included offering suggestions, answers, explanations, and information; listening to other students make such contributions; agreeing with others' contributions; ignoring others' contributions; and making no contribution when no one else in the group contributed either. We used the frequency of each kind of behavior in the analyses.

Table 2

Continua of Student Responsiveness to Other Students' Need for Help

Level	Description
	Content continuum
+3	Correct answer, complete with explanation
+2	Correct answer that completely answers the question, but with little or no explanation)
+1	Partially correct answer
0	No content
–1	Probably incorrect, very unclear
–2	Incorrect answer with little explanation
–3	Incorrect answer with explanation
	Intended helpfulness continuum
+3	Correct or incorrect response with explanation
+2	Correct or incorrect response with minimal explanation
+1	Correct or incorrect response with no explanation
0	Neutral
–1	Providing a response at a lower level than the capacity of the help-giver
–2	Ignoring or dismissing need for help
–3	Insulting

Finally, category 4 consisted of debilitating behavior: ordering others to do something or to stop doing something, commandeering equipment (including preventing others from participating in circuit construction), giving or receiving insults, interrupting others or being interrupted, and carrying out off-task conversation or behavior. The analyses used the frequencies of these behaviors.

Results and Discussion

Overview of Analysis Strategy

The analyses proceeded in four phases. The first phase examined performance differences across group ability compositions. We looked at group ability composition in three ways. (1) We examined performance differences across high-ability students in the four group compositions defined by the lowest ability in the group (low, low-medium, medium-high, high). (2) We collapsed all heterogeneous groups (groups with high-ability students and at least one student at another ability level—low, low-medium, or medium-high) and compared the performance of high-ability students in the overall heterogeneous group category with high-ability students in homogeneous groups. (3) We separated high-ability students who worked in heterogeneous groups into two subgroups: (a) high-performing high-ability students in heterogeneous groups (high-ability students who performed as well as high-ability students in homogeneous groups), and (b) low-performing high-ability students in heterogeneous groups (high-ability students who performed worse than high-ability students in homogeneous groups). We then compared the performance of high-ability students across homogeneous groups and these two subgroups of high-ability students who worked in heterogeneous groups.

The second phase in the analysis investigated the predictors of student performance. The predictors used (in correlational and regression analyses) included behavior variables at both the student and group levels: quality of group discussion, group members' responsiveness to each other's need for help, unsolicited correct work, and debilitating behavior.

The third phase examined the relationships among the individual and group behavior variables, also using correlational analyses. In particular, this phase used as outcomes the behavior variables found in the second phase to be most strongly related to performance scores and used the remaining behavior variables as predictors.

The fourth phase in the analysis examined group process differences across group compositions. The group compositions compared in this phase are the homogeneous groups and the two subgroups of heterogeneous groups—high-performing high-ability students in heterogeneous groups and low-performing high-ability students in heterogeneous groups.

This sequence of analyses demonstrates the distinctions among group compositions that most impacted student performance, determines which behavior variables (at the individual and group levels) most strongly predicted student performance, and shows whether group compositions that differed in performance also differed on the critical behavior variables. The overall intent is to understand which kinds of student and group behaviors account for differences in high-ability students' performance across group compositions.

Performance Differences Across Group Ability Compositions

Table 3 (see also Figure 1) gives the mean ability scores and performance scores for high-ability students according to the composition of their group during the

Table 3

Performance of High-Ability Students Who Worked in Different Group Compositions

	Student ability			Performance scores					
	Phase 1			Phase 2			Phase 3		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Worked individually	.71	.08	16	.79	.12	16	.82	.10	15
Lowest ability of level in the group									
Low	.74	.11	7	.76	.11	7	.82	.14	7
Low-medium	.68	.08	14	.73	.19	14	.78	.13	14
Medium-high	.70	.08	38	.71	.12	38	.72	.13	37
High	.75	.07	24	.89	.06	24	.88	.07	22

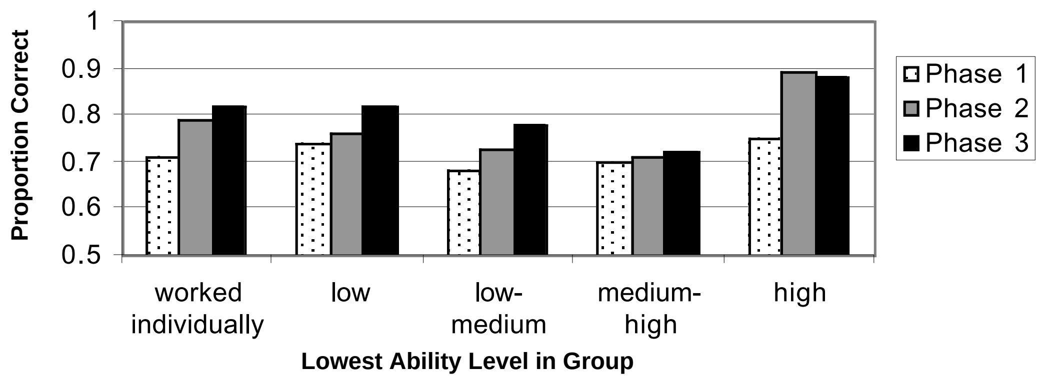


Figure 1. Mean scores for high-ability students in various group compositions.

second phase. We used four group compositions: homogeneously high and three types of heterogeneous groups according to the lowest ability level in the group (low, low-medium, or medium-high). The results also include the performance of the comparison students who worked individually in all phases of the study. Due to similar results in the separate analyses conducted for each school and problems of low power influencing statistical significance (small sample sizes in some analyses), the reported results combine both schools.

When we analyzed differences in performance scores between the four different group compositions based on the lowest ability level present, we found significant differences between them in both Phases 2 and 3. Students who worked in homogeneous groups performed significantly better than students who worked in all three types of heterogeneous groups ($F = 9.11$, $p < .001$ and $F = 6.96$, $p < .001$, in Phases 2 and 3, respectively, from analyses of variance with student ability scores as the covariate). However, we found no statistically significant differences between the three heterogeneous group compositions. High-ability students performed similarly whether the lowest ability level in the group was low, low-medium, or medium-high. As can be seen in Table 3 (and Figure 1), high-ability students who worked individually throughout the study scored partway between high-ability students in homogeneous groups and those in heterogeneous groups.

Because the performance of high-ability students in the three types of heterogeneous groups (lowest ability level in the group being low, low-medium, or medium-high) did not differ, we combined the three types of heterogeneous groups into one general heterogeneous group category and contrasted that category with homogeneous groups. The contrast between homogeneous and heterogeneous groups appears in Table 4 (see also Figure 2).

Close inspection of the data in Table 4 revealed substantial variability in performance within the heterogeneous group category, variability that was larger than in the homogeneous groups (see the larger standard deviations for heterogeneous groups than for homogeneous groups in Table 4). Some high- ability students in heterogeneous groups performed well—as well as high-ability students in homogeneous groups—while others performed much less well. To examine the variability of high-ability students' performance in the heterogeneous group category more clearly, we calculated the average of the two performance scores from Phases 2 and 3. The distribution of this average performance score in the

Table 4

Performance of High-Ability Students in Homogeneous and Heterogeneous Groups

	Student ability			Performance scores					
	Phase 1			Phase 2			Phase 3		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Heterogeneous	.70	.08	59	.72	.14	59	.75	.13	58
Homogeneous	.75	.07	24	.89	.06	24	.88	.07	22

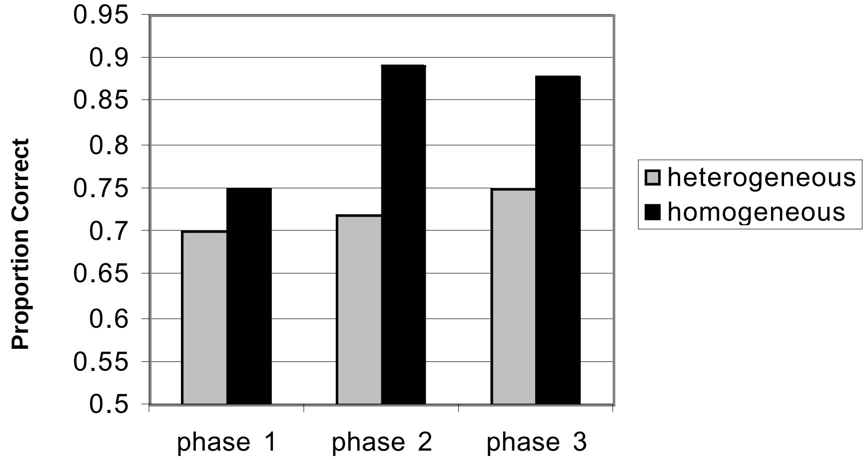


Figure 2. Mean Performance of high-ability students in homogeneous and heterogeneous groups.

heterogeneous group category revealed two distinct subgroups of high-ability students who worked in heterogeneous groups. One subgroup of high-ability students in heterogeneous groups obtained scores entirely within the range of scores of students in the homogeneous group composition; the other subgroup of high-ability students in heterogeneous groups obtained scores entirely below the range of scores in homogeneous groups (see Figure 3). Based on this result, we separated the students working in heterogeneous groups into two subgroups for analysis: One subgroup consisted of the high-performing high-ability students in heterogeneous groups and the other subgroup consisted of the low-performing high-ability students in heterogeneous groups.

Table 5 (see also Figures 4 and 5) gives the mean scores for ability and performance for the three subgroups of high-ability students: the high-ability students who worked in homogeneous groups, the high-performing high-ability students who worked in heterogeneous groups, and the low-performing high-ability students who worked in heterogeneous groups. The bottom two rows in Table 5 (see also Figure 4) show that high-ability students in homogeneous groups and high-performing high-ability students in heterogeneous groups have very similar mean performance scores for both Phases 2 and 3. Low-performing high-ability students in heterogeneous groups have considerably lower mean performance scores ($F = 55.02$, $p < .001$ and $F = 40.43$, $p < .001$).

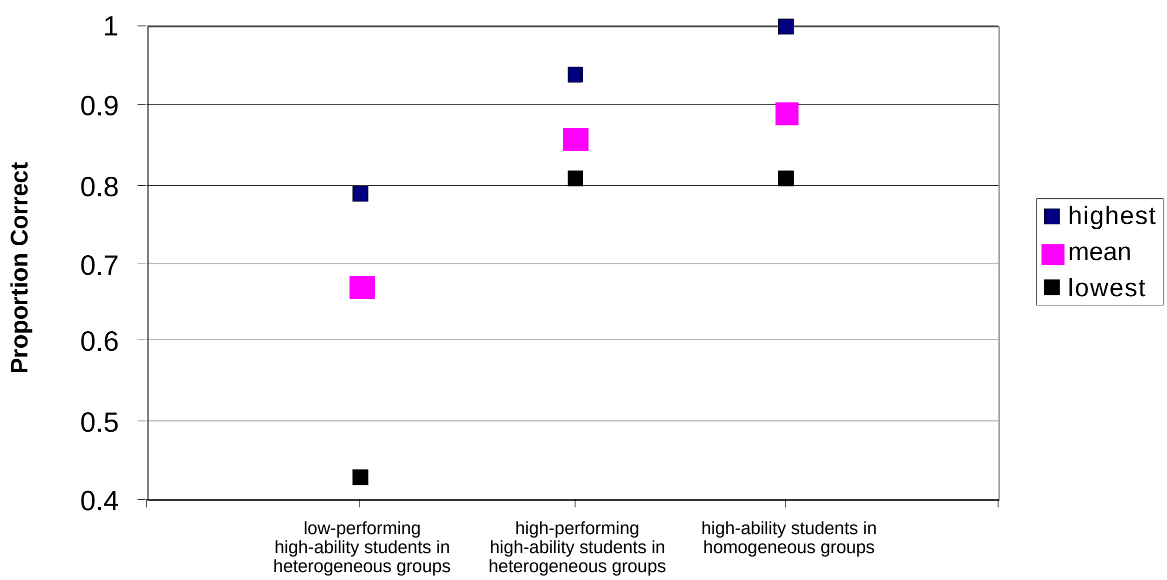


Figure 3. Range of performance scores for high-ability students in heterogeneous and homogeneous groups.

Table 5

Performance of High-Ability Students in Homogeneous Groups and High-Performing and Low-Performing Students in Heterogeneous Groups

	Heterogeneous					
	Low-performing high-ability students (*n* = 40)		High-performing high-ability students (*n* = 20)		Homogeneous (*n* = 24)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Student ability						
Student's own ability	.67	.06	.75	.10	.75	.07
Mean of other group members' ability	.54	.11	.52	.11	.75	.04
Maximum of other group members' ability	.62	.13	.62	.15	.79	.05
Minimum of other group members' ability	.46	.11	.41	.11	.71	.05
Performance score (Phase 2)	.65	.12	.85	.05	.89	.06
Performance score (Phase 3)	.68	.11	.87	.06	.88	.07

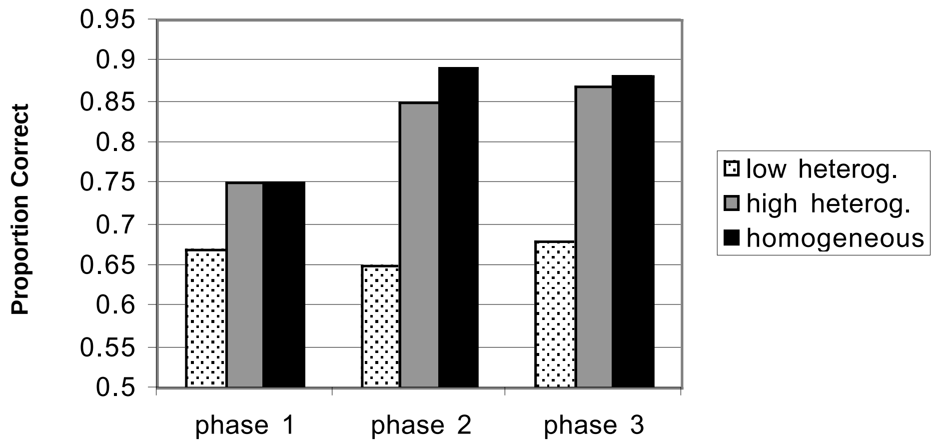


Figure 4. Mean scores of high-ability students in different group compositions.

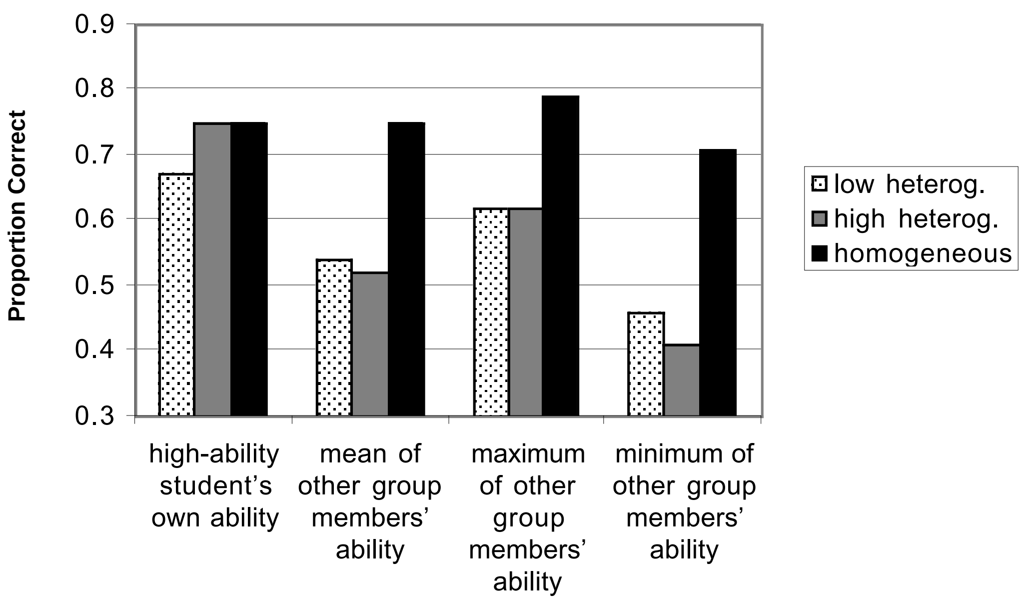


Figure 5. Phase 1 ability scores of students in different group compositions.

The upper rows of Table 5 show that neither students' own ability scores nor the ability level of the other students in their groups totally explains the pattern of performance score results across the three subgroups. First, we contrast the two subgroups of high-ability students in heterogeneous groups: low-performing high-ability students and high-performing high-ability students. While low-performing high-ability students in heterogeneous groups had lower ability (Phase 1) than high-performing high-ability students in heterogeneous groups ($F = 11.17$, $p < .001$), the ability level of high-ability students' groupmates did not differ between the two subgroups. For the two subgroups of high-ability students in heterogeneous groups, high-ability students' groupmates had very similar mean, maximum, and minimum ability levels. This finding shows that the two subgroups of high-ability students who worked in heterogeneous groups did not differ in their access to intellectual resources as defined by the ability levels of the other students in the group.

Second, we contrast high-ability students in homogeneous groups and high-performing high-ability students in heterogeneous groups to examine whether patterns of student ability (either the student's own ability or the ability level of students' groupmates) were consistent with the similar performance scores in these two subgroups (see Table 5 and Figure 5). First, the high-ability students in these two subgroups had identical mean ability levels (means of .75 in both subgroups).However, differences emerged in the ability level of students' groupmates. High-ability students in homogeneous groups had access to higher intellectual resources than did the high-performing high-ability students in heterogeneous groups, as indicated by the higher mean, maximum, and minimum ability scores of their groupmates in homogeneous groups compared to the means of those variables in the high-performing heterogeneous subgroup ($F = 44.96$, $p < .001$; $F = 7.96$, $p < .001$; $F = 67.32$, $p < .001$). These results show that, although high-performing high-ability students in heterogeneous groups had access to lower intellectual resources than did students in homogeneous groups, this did not have a detrimental effect on their performance as indicated by their Phase 2 and 3 performance scores, which remained very similar to those of students in homogeneous groups.

All further analyses presented here focus on the three subgroups of students: high-ability students who worked in homogeneous groups, high-performing high-ability students who worked in heterogeneous groups, and low-performing high-ability students who worked in heterogeneous groups. In subsequent sections we focus on understanding why the high-performing high-ability students in heterogeneous groups performed so much better than the low-performing high-ability students in heterogeneous groups. Further analyses of student and group characteristics revealed no differences between the two heterogeneous subgroups. High-performing and low-performing students in heterogeneous groups had similar proportions of female and male students, had similar distributions in terms of the exact group composition of the group (the ability levels of the three students in the group), and had similar proportions of students from the two schools (all statistical comparisons were nonsignificant). The analyses presented in subsequent sections, then, will focus on differences in student behavior and group functioning as variables to explain differences in performance between the two subgroups of high-ability students who worked in heterogeneous groups.

Predictors of Performance: Student Behavior and Group Functioning

In trying to explain which student and group behaviors predict performance outcomes and differentiate the subgroups, we examined variables related to the quality of group discussion, the responsiveness of high-ability students to others' need for help, the responsiveness of other students to high-ability students' need for help, the tendency of students to provide unsolicited correct answers, and negative or disruptive behavior. Table 6 gives the partial correlations between the behavior variables and performance scores (controlling for student ability).

A number of variables correlated significantly with performance scores. First, in contrast with previous analyses of these students (Webb et al., 1998), the accuracy of justifications provided in the group significantly correlated with high-ability students' performance scores. This results from the change in the method used to code the quality of group discussion. As noted earlier, we did not code the accuracy of group discussion in the present study if a group did not articulate the responses verbally. The high correlations for this variable show that, of the justifications articulated verbally in the group, high-ability students in groups that articulated accurate justifications performed better than high-ability students in groups that articulated less accurate justifications. These significant correlations parallel those previously reported for students of lower ability (see Webb et al., 1998). As demonstrated in those previous analyses for lower ability students, the results reported here (Table 6) show that high-ability students can also benefit from working in a group that presents accurate justifications: They can use the

Table 6

Partial Correlations Between Student Behavior and Group Functioning Variables and Student Performance (Controlling for Student Ability)

	Performance scores	
Student behavior and group functioning variables	Phase 2 (Group)	Phase 3 (Individual)
Quality of group discussion		
Accuracy of multiple-choice responses	.18	.12
Accuracy of justifications	.62****	.56****
High-ability students' responsiveness to other students' need for help		
Average level of content provided	.43***	.27
Average level of intended helpfulness	.38**	.33**
Responsiveness of group to high-ability students' need for help		
Average level of content received	.66****	.52****
Average level of intended helpfulness	.18	.28*
Unsolicited correct justifications		
Provided by high-ability student	.17	.04
Provided by another student	.26*	.32*
Debilitating behavior of high-ability student		
Commandeers equipment	–.04	–.08
Gives insult	–.07	–.22
Receives insult	–.21	–.21
Off-task	–.17	.01

* $p < .05$. ** $p < .01$. *** $p < .001$. **** $p < .0001$.

justifications they hear to correct their own misconceptions and gaps in understanding, and they can solidify initially tenuous understanding.

Second, the responsiveness of high-ability students to other students' need for help (questions and errors) significantly related to their own performance scores. This result confirms the notion discussed above that high-ability students can benefit from providing help to others (see above for more detail).

Third, the responsiveness of other group members to high-ability students' need for help (questions and errors) significantly related to high-ability students' performance scores. In particular, high-ability students benefited most from receiving accurate and elaborated help. Receiving accurate and elaborated help enabled high-ability students to correct their misconceptions and fill in gaps in their understanding. Interestingly, the level of help that other students intended to give high-ability students correlated only weakly with high-ability students' performance scores. The result can be explained by the less-than-perfect correlation between the intended helpfulness of the help-giver and the average level of content that high-ability students actually received (r = .56). While a number of students tried to give helpful explanations to high-ability students (as seen by the level of elaboration in their responses), they sometimes gave incomplete or inaccurate answers. That is, help-givers, although well meaning, did not always have the requisite knowledge or understanding to give high-ability students accurate or complete explanations.

Fourth, the frequency of unsolicited correct answers given—those volunteered even when no one else in the group asked questions or made errors—significantly correlated with high-ability students' performance scores. In particular, the number of accurate and complete justifications provided by other students correlated significantly with the performance of high-ability students. If a high-ability student did not know the answer, or did not fully understand it, listening to another student articulate the justification would provide an opportunity for the high-ability student to learn the material.

Unsolicited correct justifications provided by high-ability students, on the other hand, did not correlate with their own performance scores. This result corroborates previous research showing that the intent to explain to another person constitutes a key component of giving help. This intent influences learning because (a) the prospect of having to explain material to someone else leads to more differentiated, complex, unified, and organized cognitive structures than does merely learning the material for oneself (Zajonc, 1960), and (b) vocalizing to a peer (presumably with the intent to teach that person) produces greater concept attainment than vocalizing to the experimenter (presumably only to demonstrate mastery of the material; Durling & Schick, 1976). The purpose of the communication—helping others versus providing an answer one already knows—has more impact on cognitive restructuring than the act of verbalizing material itself. When high-ability students gave unsolicited correct answers, they likely provided answers to the items on the assessment without specifically intending to help another group member. This result contrasts with the significant correlations described above for the help provided by high-ability students in response to other students' need for help. The significant correlations for help provided when offered in response to another student's need in conjunction with the nonsignificant correlations for answers offered without any apparent need for help provides further support for the claim that help-givers benefit most from providing answers when they intend to help or explain to another.

Finally, debilitating behavior did not correlate significantly with the performance scores of high-ability students. While most of the correlations for negative processes including commandeering equipment, giving and receiving insults, and engaging in off-task conversation had an expected negative effect, none were statistically significant.

To identify the behavior variables most predictive of performance scores, we carried out multiple regression analyses. As can be seen in Table 7, two process variables provided the strongest prediction of high-ability students' performance scores: accuracy of justifications provided in the group and the average level of content that high-ability students received in response to their own need for help (questions and errors). Together, these two variables predicted performance scores more strongly than did high-ability students' own ability level.

Predictors of Group Process Variables

To understand the possible influences on the two process variables most predictive of performance scores, we examined the correlations between those two process variables and the other variables coded. The zero-order correlations appear in Table 8. As can be seen in Table 8, the accuracy of justifications provided in the

Table 7

Regression Coefficients Predicting Performance Scores From Student Ability and Behavior Variables

	Performance scores					
	Phase 2 (Group)			Phase 3 (Individual)		
	b	*B*	R^2 Chg.	*b*	*B*	R^2 Chg.
Student ability	.73***	.45***	.14**	.69**	.35**	.26***
Behavior variables			.47***			.29***
Quality of group discussion: Accuracy of justifications	.30***	.36***		.23**	.35**	
Responsiveness of group to high-ability students' need for help: Average level of content received	.07***	.43***		.04*	.28*	
Total R^2			.61***			.54***

* $p < .05$. **$p < .01$. ***$p < .001$.

Table 8

Correlations With Behavior Variables Significantly Predicting Performance Scores

Student behavior and group functioning variables	(1) Accuracy of justifications in group	(2) Average level of content received by high-ability student
Quality of group discussion		
Accuracy of multiple-choice responses	.28*	.04
Accuracy of justifications (1)	—	.51****
High-ability students' responsiveness to other students' need for help		
Average level of content provided	.54***	.38**
Average level of intended helpfulness	.65****	.42***
Responsiveness of other students to high-ability students' need for help		
Average level of content received (2)	.51****	—
Average level of intended helpfulness	.23	.56****
Unsolicited correct justifications		
Provided by high-ability student	.20	–.13
Provided by other students	.44***	.27*
Debilitating behavior of high-ability student		
Commandeers equipment	–.09	–.29*
Gives insult	–.30*	–.17
Receives insult	–.38**	–.10
Off-task	.07	–.20

* $p < .05$. ** $p < .01$. *** $p < .001$. **** $p < .0001$.

group significantly correlated with the help provided by high-ability students in response to other students' need for help, help received by high-ability students in response to their own need for help, unsolicited correct work, and several debilitating behavior variables. The finding of stronger correlations for the variables involving help offered in response to a need for help than for unsolicited answers suggests that students provide more accurate and complete responses in response to someone's question or error than when merely responding to the question asked on the assessment. The negative correlations for giving and receiving insults suggests that insulting behavior may inhibit students from voicing accurate and complete responses. The mechanisms by which this process may operate will be explored in later sections of this paper.

The average level of content received by high-ability students in response to their questions and errors positively correlated with their own responsiveness to other students' need for help and negatively correlated with their tendency to commandeer the equipment that the group used to set up their electric circuits. The positive correlation between the content that high-ability students received and their responsiveness to others suggests a link between students' mutual responsiveness. Whether high-ability students received help because they helped others, whether they helped others because they received help, or both in a reciprocally reinforcing system will be explored in later sections. The negative correlation for commandeering equipment is likely a unidirectional effect. Instances of commandeering equipment generally occurred early in group interaction as groups set up their circuits. High-ability students who grabbed and controlled the equipment may have negatively influenced other students' willingness to help them later on. This hypothesis will also be explored in subsequent sections.

Group Process Differences Across Group Compositions

As we described earlier, some high-ability students working in heterogeneous groups (the subgroup of students labeled "high-performing high-ability students in heterogeneous groups") performed as well as high-ability students working in homogeneous groups, while other high-ability students working in heterogeneous groups (the subgroup of high-ability students labeled "low-performing students in heterogeneous groups") did not. This section explores differences in student behavior and group functioning across the three of subgroups of students. Table 9 (see also Figures 6 to 10) presents means for all process variables by subgroup of group composition. First, the three subgroups differed significantly on the accuracy of justifications given in the group (but not in the accuracy of multiple-choice answers). Homogeneous groups gave the most accurate justifications; heterogeneous groups with low-performing high-ability students groups gave the least accurate justifications.

Second, low-performing high-ability students in heterogeneous groups responded least to other students who needed help. They gave the lowest level content in their responses and tended to be the least responsive from a motivational perspective. To show these results more clearly, Table 10 gives breakdowns of responses according to whether or not they were correct (without regard to the amount of elaboration in the response) and whether they had a positive or negative level of intended helpfulness (without regard to severity). Correct content included any responses at level 3 (correct answer, complete with explanation), level 2 (correct answer that completely answers the question, but with little or no explanation) or

Table 9

Group Functioning and Behavior of High-Ability Students in Homogeneous Groups and High-Performing and Low-Performing Students in Heterogeneous Groups

	Heterogeneous					
	Low-performing high-ability students (*n* = 18)[a]		High-performing high-ability students (*n* = 17)[a]		Homogeneous (*n* = 21)[a]	
Process variables	*M*	*SD*	*M*	*SD*	*M*	*SD*
Quality of group discussion						
Accuracy of multiple-choice responses	.85	.14	.79	.20	.92	.10
Accuracy of justifications*	.58	.13	.67	.12	.82	.13
High-ability students' responsiveness to other students' need for help						
Average level of content provided*	.56	.41	1.09	.40	1.07	.95
Average level of intended helpfulness*	–.38	.61	.22	.72	.46	1.40
Responsiveness of group to high-ability students' need for help						
Average level of content received*	.40	.61	1.03	.53	1.11	.89
Average level of intended helpfulness	.46	.70	.30	1.12	.49	1.27
Unsolicited correct justifications						
Provided by high-ability student	.39	.61	1.29	1.53	.95	1.02
Provided by another student	.33	.59	.59	.94	1.05	1.36
Debilitating behavior by high-ability student						
Commandeers equipment	.61	1.09	.24	.56	1.05	2.50
Gives insult	1.11	2.56	.47	1.37	.33	.91
Receives insult	.72	1.56	.29	.99	.29	.71
Off-task	1.67	3.16	.59	1.00	2.19	2.34

[a] Only students who were videotaped during group work are included here; consequently the sample sizes are smaller than in Table 5.

*Significant *F* statistics for one-way analysis of variance comparing subgroups.

level 1 (partially correct answer). Responses with no content included (level 0) included non-content answers or saying "I don't know." Incorrect content included any responses at level –1 (probably incorrect, very unclear), level –2 (incorrect answer with little or no explanation) or level –3 (incorrect answer with explanation). For intended helpfulness, positive responses included those at level 3 (correct or

Table 10

Detailed Breakdown of Behavior of High-Ability Students in Homogeneous Groups and High-Performing and Low-Performing Students in Heterogeneous Groups

	Heterogeneous		
	Low-performing high-ability students	High-performing high-ability students	Homogeneous
Process variables	%[a]	%[a]	%[a]
High-ability students' responsiveness to other students' need for help			
Content provided			
Correct	42	62	58
No content	47	34	35
Incorrect	11	4	8
Intended helpfulness			
Positive	36	51	58
Neutral	1	1	0
Negative	63	48	42
Responsiveness of group to high-ability students' need for help			
Content received			
Correct	49	64	59
No content	30	34	37
Incorrect	20	2	4
Intended helpfulness			
Positive	57	54	57
Neutral	3	1	0
Negative	39	46	43

[a]Average percentage of responses in each category.

incorrect response with explanation), level 2 (correct or incorrect response with minimal explanation), or level 1 (correct or incorrect response with no explanation). Neutral (level 0) included saying "I don't know." Negative responses included those at level –1 (providing a lower level response than the help-giver was capable of, as determined by his or her Phase 1 responses), level –2 (ignoring or dismissing the question), or level –3 (insulting the person in need of help).

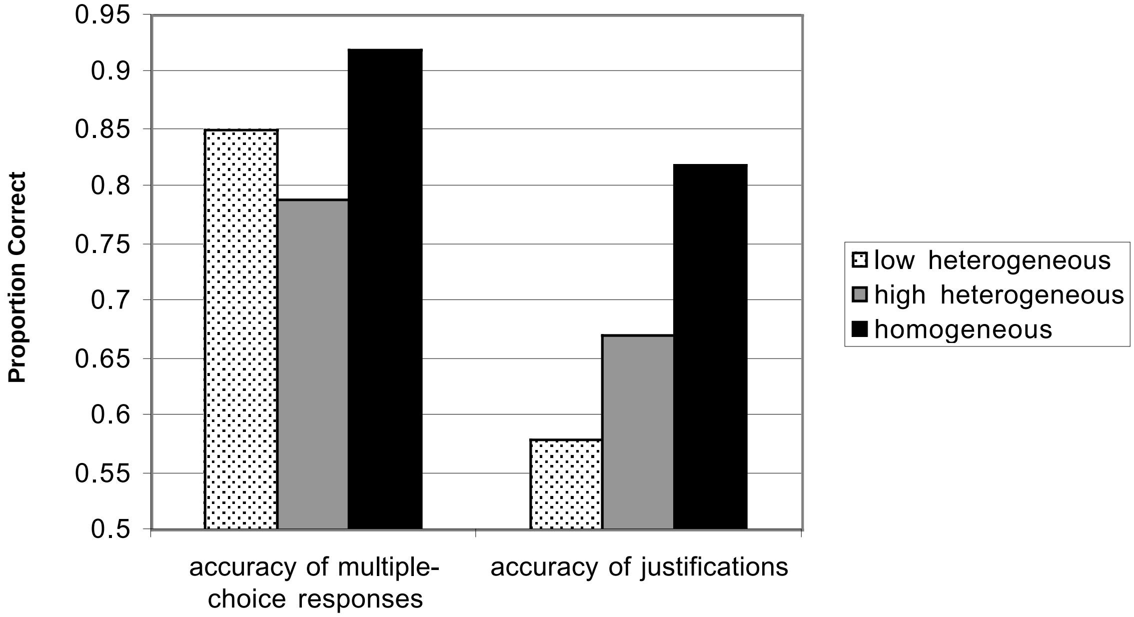


Figure 6. Mean quality of group discussion across group types.

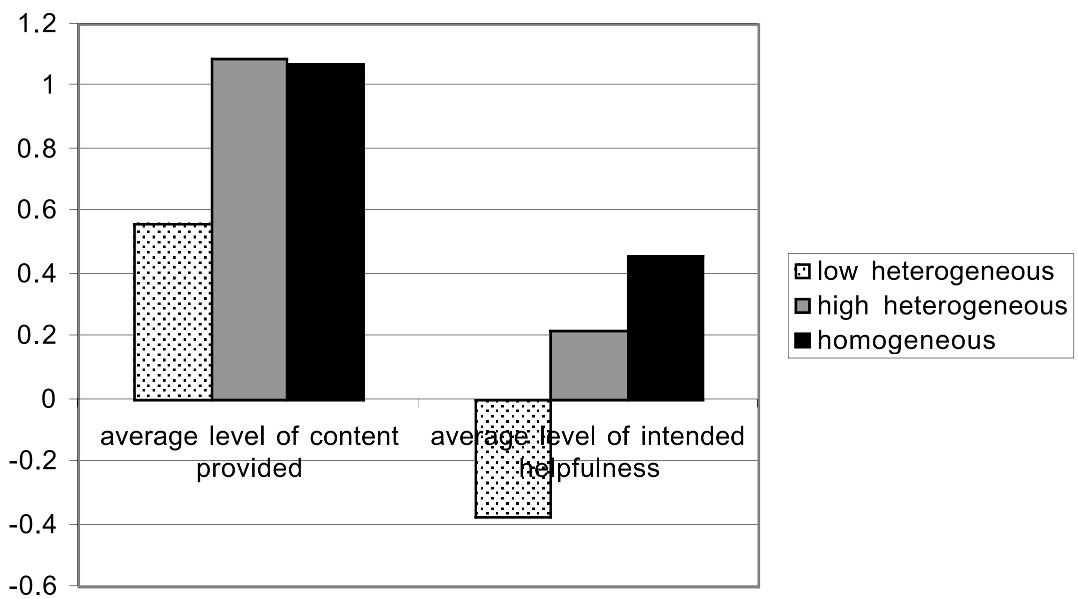


Figure 7. High-ability students' responsiveness to other students' need for help.

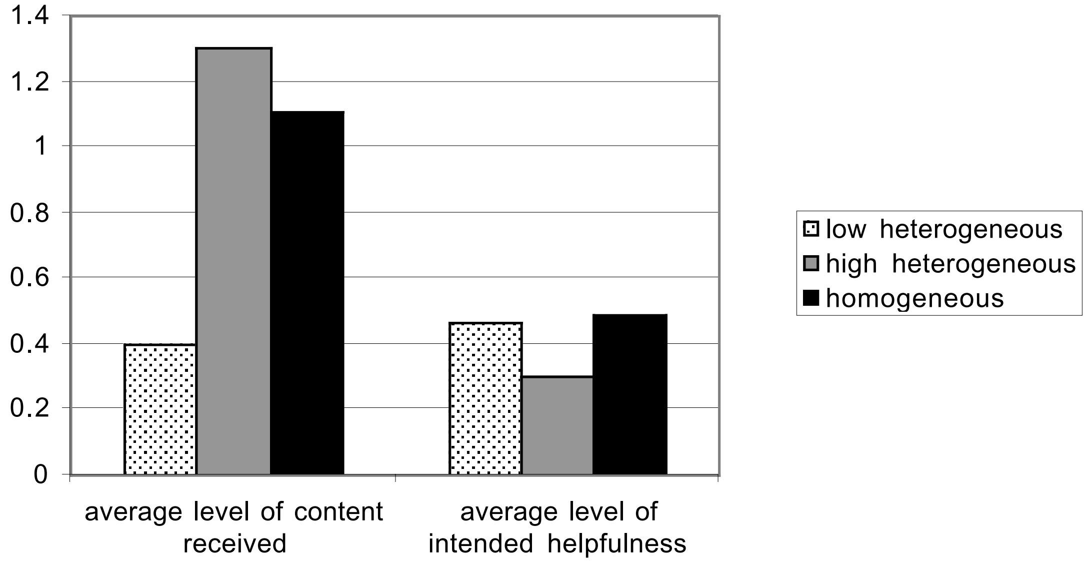


Figure 8. Responsiveness of group to high-ability students' need for help.

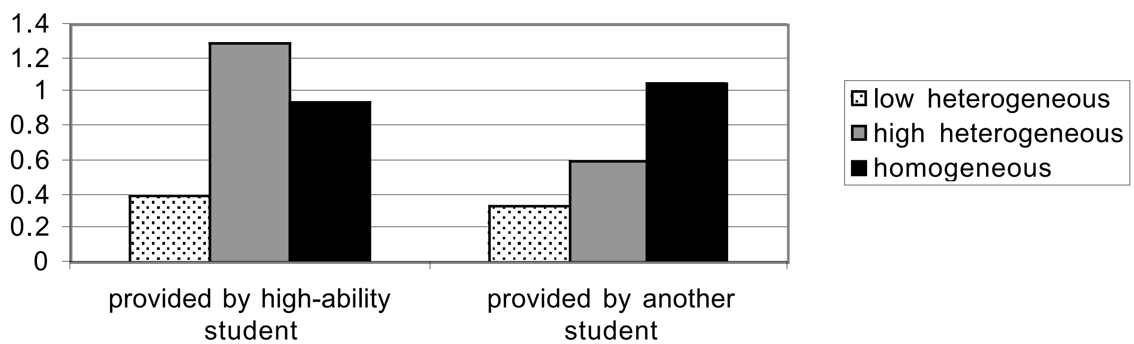


Figure 9. Unsolicited correct justifications across groups.

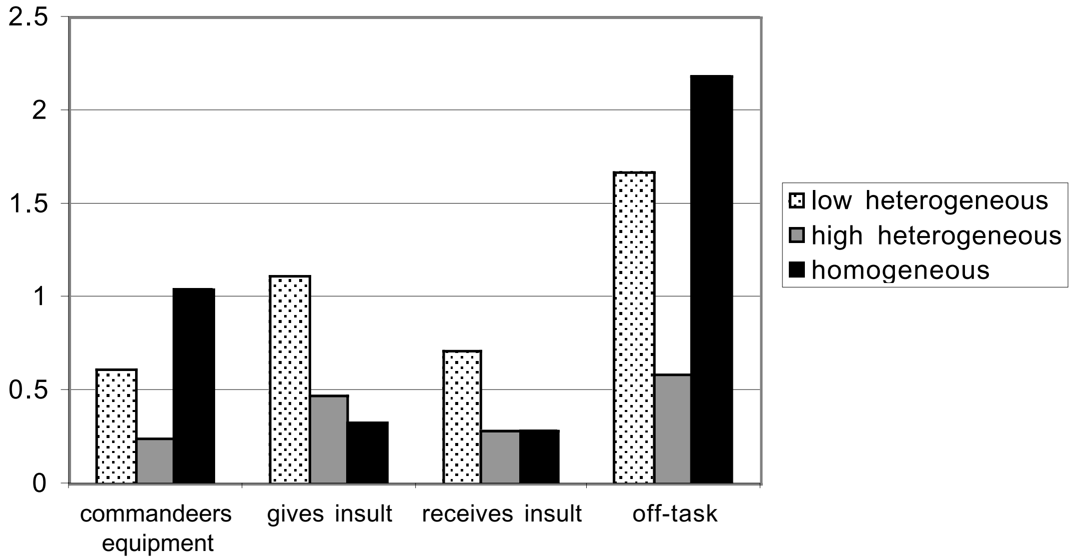


Figure 10. Debilitating behavior by high-ability students.

As seen in Table 10, high-ability students in homogeneous groups and high-performing high-ability students in heterogeneous groups for the most part gave correct responses to other students, with most of the rest of their responses having no content. They gave only a small percentage of incorrect responses. Low-performing high-ability students in heterogeneous groups, in contrast, gave just as many no-content responses as correct responses. Students in homogeneous groups and high-performing high-ability students in heterogeneous groups more likely gave positive responses (responses that at least partially answered the question) than negative responses (responses that had an implicit or explicit refusal to fully answer the question), whereas low-performing high-ability students in heterogeneous group were more likely to give negative responses than positive responses.

Differences also appeared among subgroups in the responsiveness of other students to high-ability students' need for help. High-ability students in homogeneous groups and high-performing high-ability students in heterogeneous groups received mostly correct responses from other students, whereas low-performing high-ability students in heterogeneous groups received incorrect responses or responses with no content as often as they received correct responses.

The knowledge level of the high-ability students' groupmates probably did not cause this outcome. As noted in Table 5, whether in low-performing or high-performing heterogeneous groups, high-ability students' groupmates had similar mean, maximum, and minimum ability levels.

Consistent with the nonsignificant differences shown in Table 9 for average level of intended helpfulness in responses to high-ability students, analyses across subgroups showed no significant differences in the patterns of intended helpfulness in the responses received by high-ability students. In all subgroups, high-ability students' peers tried to be helpful more than half of the time.

The remaining portions of Table 9 present the differences between the three subgroups in unsolicited correct justifications and debilitating behavior. Although none of the statistical comparisons were significant, low-performing high-ability students in heterogeneous groups showed a tendency to give fewer unsolicited correct justifications, to give more insults, and to receive more insults than high-ability students in high-performing heterogeneous groups and high-ability students in homogeneous groups.

Conclusions

This study produced three main findings: (a) High-ability students performed well in homogeneous groups; (b) high-ability students in some heterogeneous groups performed well whereas high-ability students in other heterogeneous groups did not; and (c) the types of group interactions that occurred during group work strongly influenced performance. In homogeneous high-ability groups, students provided accurate answers to the task's questions, responded to other students' need for help (as indicated by the questions these students asked and the errors they made), and contributed answers to the task's questions even when no one requested the answer. These types of behavior significantly and strongly correlated with performance scores. In heterogeneous groups in which the high-ability student had similar experiences, high-ability students performed as well as those students in homogeneously high groups. In heterogeneous groups in which the high-ability student did not experience positive group functioning, high-ability students performed significantly worse. In the latter groups, high-ability students infrequently responded to their groupmates' need for help, seldom received much help when they asked questions or made errors, rarely volunteered unsolicited answers to the tasks' questions, and more often engaged in debilitating social behavior (especially giving and receiving insults).

Importantly, the particular ability composition of heterogeneous groups did not relate to either group functioning or the performance of high-ability students in those groups. Heterogeneous groups in which high-ability students performed well and heterogeneous groups in which high-ability students performed poorly were equally likely to have students with low ability, and had very similar mixes of ability levels in the group.

The results of this study help resolve the apparent inconsistencies in previous research by comparing the performance of high-ability students in different group compositions. As described earlier, some studies found that high-ability students performed better in homogeneous groups than in heterogeneous groups, while others found no differences in the performance of high-ability students in either group type. The current results show that heterogeneous groups should not be considered as a single, undifferentiated category of group composition. This study revealed two categories of high-ability students who worked in heterogeneous groups: high-performing and low-performing according to the group dynamics occurring in the group. The mix of high-functioning and low-functioning heterogeneous groups in previous studies could explain the discrepant results. Studies in which heterogeneous groups functioned in productive ways would be expected to produce similar performance of high-ability students in homogeneous and heterogeneous groups. Studies in which heterogeneous groups functioned poorly would be expected to show an advantage for homogeneous grouping. Importantly, group functioning, not group structure, predicts the outcomes for high-ability students.

A final note concerns the designation "high-ability" and the relevance of the current results to gifted populations. The current study defined high-ability students as the top 25% of a sample representing a fairly wide cross-section of the 8th-grade population in Los Angeles metropolitan area public middle schools. Although the high-ability students in this study may have included a limited number of students designated as gifted or talented (usually defined as the top 3% of the population on achievement and/or intelligence measures), many or most of the students were probably not designated as gifted or talented. Consequently, the results of this study should not be generalized to the experiences or outcomes of gifted students. Moreover, the results of this study generalize only to the situations in which teachers expose all students to the same curriculum and instruction, not to situations in which some highly able students receive an accelerated curriculum or special instruction or faster-paced curriculum.

Because the present study did not shed light on *why* and *how* productive group dynamics developed in some heterogeneous groups but not in others, our ongoing, further analyses examine (a) the evolution of each group's dynamics over time, and (b) the relationship of students' behavior to self-report measures about their goal orientation (a focus on obtaining correct answers versus a focus on learning the material) and other attitudes toward the task. In these in-depth analyses, we hope to uncover the reasons why some heterogeneous groups produced dynamics conducive to learning while others did not.

References

Armstrong, N. A. (1994). *The effects of cooperative learning on gifted students in heterogeneous and homogeneous groups (ability grouping).* Unpublished doctoral dissertation, Ball State University, Muncie, IN.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59,* 87-96.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefit of teaching. *Journal of Educational Psychology, 72*, 593-604.

Baron, J. B. (1994, April). *Using multi-dimensionality to capture verisimilitude: Criterion-referenced performance-based assessments and the ooze factor.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal, 21,* 755-765.

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

Buros, O. K. (1974). *Tests in print II.* Highland Park, NJ: Gryphon Press.

Carter, G., & Jones, M. G. (1994). Relationship between ability-paired interactions and the development of fifth graders' concepts of balance. *Journal of Research in Science Teaching*, *31*, 847-856.

Chi, M. T. H., & Bassock, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 251-282). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., Bassock, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Damon, W., & Phelps, E. (1989). Critical distinctions among three methods of peer education. *International Journal of Educational Research, 13,* 9-19.

Deering, P. (1989, November). *An ethnographic approach for examining participants' construction of a cooperative learning classroom culture.* Paper presented at the annual meeting of the American Anthropological Association, Washington, DC.

Dembo, M. H., & McAuliffe, T. J. (1987). Effects of perceived ability and grade status on social interaction and influence in cooperative groups. *Journal of Educational Psychology*, *79*, 415-423.

Deutsch, M. (1949). An experimental study of the effects of cooperation and competition upon group process. *Human Relations*, *2,* 199-231.

Durling, R., & Schick, C. (1976). Concept attainment by pairs and individuals as a function of vocalization. *Journal of Educational Psychology, 68,* 83-91.

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

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.

Hooper, S., & Hannafin, M. J. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research, 4,* 413-424.

Hooper, S., & Hannafin, M. J. (1991). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Educational Technology Research and Development, 39,* 27-40.

Hooper, S., Ward, T. J., Hannafin, M. J., & Clark, H. T. (1989). The effects of aptitude composition on achievement during small group learning. *Journal of Computer-Based Instruction*, *16*, 102-109.

Hooper, S., Temiyakarn, C., & Williams, M. D. (1993). The effects of cooperative learning and learner control on high- and average-ability students. *Educational Technology Research and Development*, *41,* 5-18.

Johnson, D. W., & Johnson, R. T. (1979). Conflict in the classroom: Controversy and learning. *Review of Educational Research*, *49*, 51-70.

Johnson, D. W., & Johnson, R. T. (1985). The internal dynamics of cooperative learning groups. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 103-124). New York: Plenum Press.

Johnson, D. W., & Johnson, R. T. (1993). Gifted students illustrate what isn't cooperative learning. *Educational Leadership, 50*, 60-61.

Jones, M. G., & Carter, G. (1994). Verbal and nonverbal behavior of ability-grouped dyads. *Journal of Research in Science Teaching*, *31*, 603-619.

Jones, M. G., Carter, G., & Rua, M. (1998, April). *The effects of partner's ability on the achievement and conceptual organization of high achieving fifth-grade students*. Paper presented at the National Association of Research in Science Teaching Conference, San Diego, CA.

Kagan, S. (1992). *Cooperative learning.* San Juan Capistrano, CA: Resources for Teachers (27128 Paseo Espada, Suite 622).

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*, 78-94.

Knupfer, N. N. (1993). Logo and transfer of geometry knowledge: Evaluating the effects of student ability grouping. *School Science and Mathematics, 93*, 360-368.

Lawrenz, F., & Munch, T. W. (1985). Aptitude treatment effects of laboratory grouping method for students of differing reasoning ability. *Journal of Research in Science Teaching, 22*, 279-287.

Mastergeorge, A., Webb, N. M., Roc, C., & Baure, G. (2000, April). *Understanding collaborative learning environments: The development of students' mathematical thinking.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Mayer, R. E. (1984). Aids to prose comprehension. *Educational Psychologist, 19*, 30-42.

Melser, N. A. (1999). Gifted students and cooperative learning: A study of grouping strategies. *Roeper Review, 21*, 315.

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. V. 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 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.

Nelson-Le Gall, S., Gumerman, R. A., & Scott-Jones, D. (1983). Instrumental help-seeking and everyday problem-solving: A developmental perspective. In J. D. Fisher, A. Nadler, & B. M. DePaulo (Eds.), *New directions in helping* (Vol. 2, pp. 265-283). New York: Academic Press.

Noddings, N. (1985). Small groups as a setting for research on mathematical problem solving. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving* (pp. 345-359). Hillsdale, NJ: Lawrence Erlbaum Associates.

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 the way they ought to. *International Journal of Educational Research, 13,* 89-99.

Saxe, G. B. (1992). Studying children's learning in context: Problems and prospects. *The Journal of the Learning Sciences, 2,* 215-234.

Shavelson, R. J., Webb, N. M., Stasz, C., & McArthur, D. (1988). In R. Charles & E. Silver (Eds.), *Teaching and assessing mathematical problem-solving: A research agenda* (pp. 203-231) Hillsdale, NJ: Lawrence Erlbaum Associates.

Sherman, G. P., & Klein, J. D. (1995). The effects of cued interaction and ability grouping during cooperative computer-based science instruction. *Educational Technology Research and Development, 43,* 5-24.

Shipman, V. C. (1985). *The New Jersey Test of Reasoning Skills.* Totowa, NJ: Totowa Board of Education.

Skon, L., Johnson, D. W., & Johnson, R. T. (1981). Cooperative peer interaction versus individual competition and individualistic efforts: Effects on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology, 73,* 83-92.

Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice-Hall.

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.

Tudge, J., & Rogoff, B. (1989). Peer influences on cognitive development: Piagetian and Vygotskian perspectives. In M. H. Bornstein & J. S. Bruner (Eds.), *Interaction in human development* (pp. 17-40). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vedder, P. (1985). *Cooperative learning. A study on processes and effects of cooperation between primary school children*. Westerhaven Groningen, Netherlands: Rijkuniversiteit Groningen.

Vygotsky, L. S. (1981). The genesis of higher mental functioning. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 144-188). Armonk, NY: Sharpe.

Webb, N. M. (1980). A process-outcome analysis of learning in group and individual settings. *Educational Psychologist, 15,* 69-83.

Webb, N. M., & 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). New York, NY: Academic Press.

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

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

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

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

Wilkinson, L. C., & Calculator, S. (1982a). Effective speakers: Students' use of language to request and obtain information and action in the classroom. In L. C. Wilkinson (Ed.), *Communicating in the classroom*. New York: Academic Press.

Wilkinson, L. C., & Calculator, S. (1982b). Requests and responses in peer-directed reading groups. *American Educational Research Journal, 19,* 107-120.

Wilkinson, L. C., & Spinelli, F. (1983). Using requests effectively in peer-directed instructional groups. *American Educational Research Journal, 20,* 479-502.

Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24,* 345-376.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology Psychiatry, 17,* 89-100.

Yackel, E., Cobb, P., & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education*, *22*, 390-408.

Zajonc, R. B. (1960). The process of cognitive tuning in communication. *Journal of Abnormal and Social Psychology, 61,* 159-167.

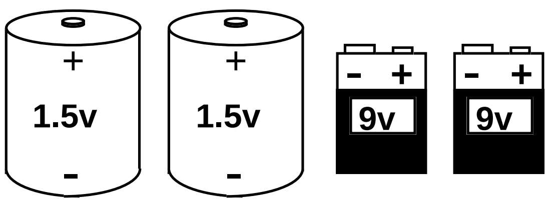
APPENDIX

Example Test Questions

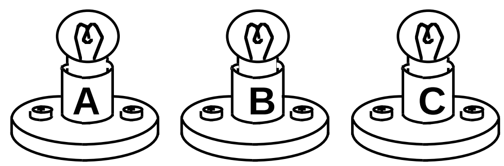
1. (a) Use the items drawn below (batteries and bulbs) to draw **two circuits** in the boxes labeled Circuit A and Circuit B. **Follow these rules:**

- **Bulb A should be in Circuit A. Bulb B should be in Circuit B.**
- **Bulb A should be brighter than Bulb B**.
- **There should be one 9-volt battery in each circuit.**
- **You must draw the wires needed to connect up the items in each circuit.**
- **Use all of the items but do not use any item more than once.** For example, if you put Bulb C in Circuit A, you cannot also put it in Circuit B.

Batteries



Bulbs



Draw the circuits in these boxes:

Circuit A (brighter)	Circuit B (dimmer)

1. (b) Why will Bulb A in Circuit A be brighter than Bulb B in Circuit B? (Try to use scientific terms in your answer.)

1. (c) Which of the two circuits you drew has the **highest voltage**?

Circle one: **CIRCUIT A** **CIRCUIT B** **BOTH CIRCUITS HAVE THE SAME VOLTAGE**

Why? (Try to use scientific terms in your answer.)

____

____

____

1. (d) Which of the two circuits you drew has the **highest resistance**?

Circle one: **CIRCUIT A** **CIRCUIT B** **BOTH CIRCUITS HAVE THE SAME RESISTANCE**

Why? (Try to use scientific terms in your answer.)

____

____

____

1. (e) Which of the two circuits you drew has the **highest current**?

Circle one: **CIRCUIT A** **CIRCUIT B** **BOTH CIRCUITS HAVE THE SAME CURRENT**

Why? (Try to use scientific terms in your answer.)

____

____