

Collaboration: A Literature Review

Research Report

Emily R. Lai

June 2011

About Pearson

Pearson, the global leader in education and education technology, provides innovative print and digital education materials for preK through college, student information systems, and learning management systems, teacher licensure testing, teacher professional development, career certification programs, and testing and assessment products that set the standard for the industry. Pearson's other primary businesses include the Financial Times Group and the Penguin Group. For more information about the Assessment & Information group of Pearson, visit <http://www.pearsonassessments.com/>.

About Pearson's Research Reports

Pearson's research report series provides preliminary dissemination of reports and articles prepared by TMRS staff, usually prior to formal publication. Pearson's publications in .pdf format may be obtained at: <http://www.pearsonassessments.com/research>.

Abstract

Collaboration is the “mutual engagement of participants in a coordinated effort to solve a problem together.” Collaborative interactions are characterized by shared goals, symmetry of structure, and a high degree of negotiation, interactivity, and interdependence. Interactions producing elaborated explanations are particularly valuable for improving student learning. Nonresponsive feedback, on the other hand, can be detrimental to student learning in collaborative situations. Collaboration can have powerful effects on student learning, particularly for low-achieving students. However, a number of factors may moderate the impact of collaboration on student learning, including student characteristics, group composition, and task characteristics. Although historical frameworks offer some guidance as to when and how children acquire and develop collaboration skills, there is scant empirical evidence to support such predictions. However, because many researchers appear to believe children can be taught to collaborate, they urge educators to provide explicit instruction that encourages development of skills such as coordination, communication, conflict resolution, decision-making, problem-solving, and negotiation. Such training should also emphasize desirable qualities of interaction, such as providing elaborated explanations, asking direct and specific questions, and responding appropriately to the requests of others. Teachers should structure tasks in ways that will support the goals of collaboration, specify “ground rules” for interaction, and regulate such interactions. There are a number of challenges in using group-based tasks to assess collaboration. Several suggestions for assessing collaboration skills are made.

Keywords: collaboration, collaborative learning, cooperation, group assessment

Acknowledgements

The author thanks Janet Fowler for assistance in conducting literature searches and the following reviewers for their helpful comments and suggestions on an earlier draft of this paper: Ha Phan, Jolana Rivas, Tian Song, and Changjiang Wang.

Collaboration: A Literature Review

Educators in a variety of educational settings—from K12 to the university classroom—have long used collaborative approaches to teaching and assessing students. More recently, educators and policy makers have identified the ability to collaborate as an important outcome in its own right rather than merely a means to an end. For example, the Partnership for 21st Century Skills has identified collaboration as one of several learning and innovation skills necessary for post-secondary education and workforce success. In addition, the newly-created Common Core State Standards reflect collaboration as a communication skill vital for college and employment. The purposes of this literature review are to (a) explore how researchers have defined collaboration; (b) investigate how collaboration skills develop; (c) learn how teachers can encourage development of collaboration skills in their students; and (d) review best practices in assessing collaboration skills.

Definition of Collaboration

Theoretical Perspectives

Collaborative learning is broadly defined as “a situation in which two or more people learn or attempt to learn something together,” and more specifically as joint problem solving (Dillenbourg, 1999, p. 1). Roschelle and Teasley define collaboration more specifically as “mutual engagement of participants in a coordinated effort to solve a problem together,” (as cited in Dillenbourg et al., 1996, p. 2). Dillenbourg notes the difficulty of agreeing on a definition of collaborative learning, even among experts. Ambiguity in the meaning of collaborative learning stems from several sources. First, the scale of such interactions may range from two people to thousands, with different theoretical tools needed to analyze interactions occurring at different

levels. Second, the question of what constitutes learning is a source of uncertainty. As Dillenbourg (1999) explains, researchers use “learning” to refer to several different types of activities:

- students studying course materials together for a test;
- joint problem solving in which learning is assumed to occur as a by-product of interactions;
- learning as a “biological and/or cultural process” that takes place over several years (p. 4); and
- “learning from collaborative work, which refers to the lifelong acquisition of expertise within a professional community” (p. 4).

Dillenbourg notes that nothing is inherently instructive about working with more than one person on a task; rather, interaction triggers learning processes. Collaborative learning situations require instructions, a physical setting, and other kinds of performance constraints. These elements do not guarantee collaboration; they only make it more likely.

Roschelle (1992) frames collaboration as an exercise in convergence or construction of shared meanings and notes that research on conversational analysis has identified features of interactions that enable participants to reach convergence through the construction, monitoring, and repairing of shared knowledge. Convergence occurs gradually, but tends to include four elements: a) construction of an abstract understanding of the problem’s deep structure; b) the interplay of metaphors; c) an iterative cycle of displaying, confirming, and repairing conceptions; and d) application of progressively higher standards of evidence for convergence. Similarly, Roschelle and Teasley (1995) define collaboration as “coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem”

(p. 70). They define the joint problem space as the shared knowledge structure that supports problem-solving by integrating goals, descriptions of the current problem state, and awareness of potential strategies, as well as the links between these things. According to Roschelle and Teasley, collaboration takes place within this joint problem space, which provides the structure needed to allow meaningful conversations about the problem. To construct a joint problem space, partners must have ways to introduce and accept knowledge, monitor exchanges for evidence of divergent meanings, and repair any divergences identified.

As Van Boxtel, et al. (2000) explain, collaborative learning activities allow students to provide explanations of their understanding, which can help students elaborate and reorganize their knowledge. Social interaction stimulates elaboration of conceptual knowledge as group mates attempt to make themselves understood, and research demonstrates that providing elaborated explanations improves student comprehension of concepts. Once conceptual understandings are made visible through verbal exchange, students can negotiate meaning to arrive at convergence, or shared understanding.

Qualities of Collaborative Learning

Collaboration is sometimes distinguished from cooperative learning in that cooperation is typically accomplished through the division of labor, with each person responsible for some portion of the problem solving. Collaboration, on the other hand, involves participants working together on the same task, rather than in parallel on separate portions of the task. However, Dillenbourg et al. (1996) note that some spontaneous division of labor may occur during collaboration. Thus, the distinction between the two is not necessarily clear-cut. According to Dillenbourg et al. (1996), in cooperation, the task is split hierarchically into independent sub-

tasks and coordination is only required for “assembling partial results.” Collaboration, on the other hand, may divide cognitive processes into intertwined layers, but coordination occurs throughout.

As Dillenbourg (1999) notes, there are several qualities that characterize truly collaborative interactions. First, collaboration is characterized by a relatively symmetrical structure, however that symmetry is accomplished. For example, in situations with symmetry of action, each participant has access to the same range of actions. This contrasts with the typical division of labor in cooperative learning structures; partners split up the work, solve sub-tasks individually, and then put their respective contributions together. Symmetry of knowledge occurs when all participants have roughly the same level of knowledge, although they may have difference perspectives. Symmetry of status involves collaboration among peers rather than interactions involving supervisor/subordinate relationships. Finally, symmetry of goals involves common group goals rather than individual goals that may conflict (Dillenbourg, 1999).

Another marker of true collaboration is the quality of interactions, especially the degree of interactivity and negotiability (Dillenbourg, 1999). Interactivity refers to the extent to which interactions influence participants’ thinking. Negotiability refers to the extent to which no single group member can impose his view unilaterally on all others, but rather all group members must work toward common understanding. Dillenbourg (1999) points out that trivial, obvious, and unambiguous tasks provide few opportunities to observe negotiation because there is nothing about which to disagree. Moreover, misunderstandings may actually be important from a learning standpoint; they force participants to construct explanations, give reasons, and justify their positions.

Historical Perspectives

Much of the research on collaborative and cooperative learning is rooted in the work of Piaget and Vygotsky (Dillenbourg et al., 1996). For example, socio-constructivists borrow Piaget's system of developmental stages describing children's cognitive progress, as well as ideas related to cognitive conflict, which refers to the sense of dissonance experienced when one becomes aware of a discrepancy between one's existing cognitive framework and new information or experiences. According to the socio-constructivist approach, cognitive conflict is critical in triggering growth. Social interactions help to facilitate such conflict to the extent that students interact with peers at more advanced developmental levels. Within this school of thought, group heterogeneity is an important consideration, as group mates are expected to possess different knowledge, different knowledge representation schemes, and different reasoning mechanisms (as reviewed in Dillenbourg et al., 1996). For example, research in the Piagetian tradition suggests that when conservers (i.e., children who realize that pouring a glass of water into another glass that is differently-sized and differently-shaped does not change the quantity of water) are paired with non-conservers on a conservation task, non-conserving members are highly likely to reach conservation as a result of interaction, whereas the regression of conserving members is rare (as summarized in Tudge, 1992). Dillenbourg et al. (1996) point out that this approach is probably too mechanistic, that disagreement and conflict in and of themselves are not as important as the communication they engender.

Vygotsky's work placed more emphasis on the value of social interaction itself for causing individual cognitive change, as opposed to being merely stimulated by it (as reviewed in Dillenbourg et al., 1996). In this formulation, social interaction is internalized, which causes conceptual changes as participants appropriate new understandings. Like Piaget, Vygotsky

emphasized the importance of heterogeneous groupings of collaborators. According to Vygotsky, the zone of proximal development is the distance between what a student can accomplish individually and what he/she can accomplish with the help of a more capable “other.” Whereas Piagetian studies typically pair children from different developmental stages to facilitate cognitive conflict, studies in the Vygotskian tradition frequently pair children with adults. Rather than focusing on cognitive conflict as a trigger for conceptual change, socio-culturalists view collaborative learning as learning that occurs within the zone of proximal development (as summarized in Dillenbourg et al., 1996).

More recently, the shared or situated cognition approach—informed by researchers in sociology, anthropology, and even computer science—emphasizes the social structures in which interactions occur (Dillenbourg et al., 1996). This approach sees the environment as an integral part of cognitive activities associated with collaboration. Accordingly, attempts to investigate collaboration that ignore social structures are likely to be biased. Under this view, knowledge is not something that is handed down from one partner to another. Rather, knowledge is co-constructed through interactions among collaborators. This approach emphasizes that the whole of group behavior is more than the sum of its individual parts. In other words, group interactions evolve in ways that are not necessarily predictable based on the inputs of group members. This latter insight suggests that viewing the group rather than individual group members as the unit of analysis could produce qualitatively different conclusions about collaboration (Dillenbourg et al., 1996).

Since the late 1990s, a new strand of research on collaborative learning focusing on new technologies for mediating, observing, and recording interactions during collaboration has emerged (Kreijns et al., 2003). This new strand of research, commonly called computer-

supported collaborative learning (CSCL), typically uses online networks for facilitating and recording online interactions among two or more individuals who may be geographically and/or temporally dispersed. Much of this research has grown in parallel to new technologies for supporting distance interactions, such as email, chat, instant-messaging capability and more recently, resources for synchronous video conferencing (such as Skype).

These different historical perspectives have led to different research paradigms. For example, Dillenbourg et al. (1996) characterize research stemming from the Piagetian, Vygotskian, and shared cognition approaches as the “effect” paradigm, the “conditions” paradigm, and the “interactions” paradigm, respectively. The latest variant of the interactions paradigm might be called the “computer-supported” paradigm. Each of these paradigms is explored separately below.

The “effect” paradigm.

Those working in the “effect” paradigm tend to examine outcomes of collaboration rather than the collaborative process itself, comparing group performance with individual performance. Research in this tradition suggests that collaborative classroom structures can have powerful effects on student learning and performance. For example, seventh-grade students working in groups of 3–4 on computational math problems earned significantly higher scores working in groups than equivalent-ability students working individually (Webb, 1993). Furthermore, among students working in groups, students who received help during collaboration and who tried to understand the assistance they received earned much higher math scores at post-test than did students who passively received assistance. Behavior during group collaboration was significantly related to ability, with high-ability students more likely than others to correctly

solve problems aloud with little or no assistance. However, behavior during group work was a more salient predictor of subsequent performance on the individual post-test than was ability.

Saner et al. (1994) report the results of a study that administered hands-on science tasks to students in grades 5 and 8. Students completed the first portion, answering short-answer questions about relevant content, individually (Part 1). The second portion of the task included hands-on science activities, such as observing, recording data, and carrying out experiments. This portion was completed in pairs (Part 2). The third portion of the task was again completed individually, and entailed interpreting results obtained from the group portion of the task and applying those results to an unfamiliar context (Part 3). Saner et al. (1994) conclude that at both grade levels, higher- and lower-ability students were affected differently by the collaboration. Higher ability students who performed well as individuals on Part I tended to perform similarly well on Part 3. In addition, the best predictor of Part 3 scores for higher-ability students was their own score on Part 1 of the task. In contrast, lower-ability students exhibited a carry-over effect of the collaboration on their Part 3 scores, which tended to be higher than their Part 1 scores. In fact, the best predictor of Part 3 scores for lower-ability students was their score on Part 2. Thus, collaboration tends to benefit lower-ability students, whereas there appears to be no carry-over effect for higher-ability students.

Finally, Fall et al. (1997) summarize the results of a study in which 500 students in grade 10 participated in a multi-stage collaborative assessment. During the first stage, students read a story and individually answered a few questions, interpreting salient themes from the readings. During the second stage, half of the students were randomly selected to discuss the story in 3-person groups for 10 minutes. The other half continued to work independently. During the third stage, students who collaborated completed the remainder of their questions as individuals.

Results suggest that students who discussed the story improved their understanding of facts and their interpretations. After the discussion opportunity, these students presented more correct facts, more interpretations, and higher-quality interpretations than before the discussion and compared to the no-discussion group. Similar to results from Saner et al. (1994), the effects of discussion varied according to student ability, with low-ability students benefitting more from discussion than high-ability students.

On the other hand, Tudge (1992) studied the performance of student pairs on a science task, concluding that collaboration was as likely to diminish performance as to improve it. In Tudge's study, 153 students aged 5 to 9 worked in pairs on a series of tasks involving a balance beam. Researchers manipulated weights applied to the balance beam, as well as the distance from the fulcrum. Students were asked to predict which side the beam would tilt toward when the supports were removed. Importantly, researchers did not provide students with any feedback about the correctness of their predictions, and because the supports were never removed, students were unable to test and modify their predictions. Children's responses were coded into several categories according to the degree of sophistication of the decision rule used to support their prediction. Researchers tested four different types of student configuration: individual students, students who demonstrated equivalent decision rules at pre-test, and students demonstrating unequal decision rules at pre-test—one student adopting a higher rule, and the other adopting a lower rule. During collaboration, students took turns making predictions and justifying their predictions. When their predictions conflicted, students were asked to explain their reasons and agree on one prediction.

Results from the Tudge study (1992) suggest that collaboration had a strong impact on student performance, with changes in reasoning persisting for several weeks afterwards. Tudge

found that the group of less-competent students (those using inferior decision rules) was the only group that improved significantly on post-test. However, the more competent students (those using superior decision rules) were the only group that declined significantly post-test. Student pairs that used an equivalent decision rule tended to earn the same score at post-test, as did students who worked alone, regardless of the score earned on the pre-test. Thus, collaboration between students of different ability levels did change student performance, but collaboration was as likely to decrease as increase student performance. Moreover, these latter results suggest that this finding is not simply an artifact of regression to the mean. By way of explanation, the authors point to the fact that students received no feedback on their performances. Thus, students using more sophisticated reasoning, particularly if they were not confident in their responses, were susceptible to persuasion by their less sophisticated partners in the absence of confirming or disconfirming evidence.

The “conditions” paradigm.

Researchers working in the “conditions” paradigm generally attempt to determine the conditions moderating the effectiveness of collaboration on learning, such as individual characteristics of group members, group heterogeneity and size, and task features (Dillenbourg et al., 1996). For example, Webb (1991) found significant differences in the collaborative learning experiences of boys and girls. Boys were more likely than girls to give and receive elaborated explanations, and their explanations were more likely to be accepted by group mates than girls’ explanations. Boys were also more likely to receive responses to requests for help, perhaps because they asked direct and specific questions more frequently. Boys also benefitted more from collaboration, outperforming girls on subsequent achievement measures, despite the fact that the two groups did not differ on ability. Tudge (1992) found that boys tended to demonstrate

no performance differences as a result of collaboration, whereas girls demonstrated significant declines in performance after collaborating.

Participation patterns may also vary by the ability level of the student. For example, Webb (1991) found that high-ability students tend to provide more explanations and give more information, whereas low-ability students are more likely to be off task. Interestingly, the relative ability level of the student was more important in this regard than absolute ability. Thus, moderate achievers placed in groups with lower-performing students demonstrated the same pattern of interaction as students with high absolute ability.

The composition of the group, particularly members' genders and abilities, is also important in moderating the effects of collaboration. For example, Webb (1991) meta-analyzed 17 studies investigating collaborative group work, finding that both the patterns of interaction and the effects of collaboration varied across groups with different ability-level compositions. She categorized interactions within five different types of student groups: heterogeneous groups with a wide ability range (including high-, medium-, and low-ability students), heterogeneous groups with a narrow ability range (combining high- with medium-ability or medium- with low-ability), homogeneous high-ability, homogeneous moderate-ability, and homogeneous low-ability groups. In mixed groups featuring a wide range of abilities, high- and low-ability students tended to form teacher-student relationships, leaving medium-ability students out. Medium-ability students in these groups provided fewer explanations than the other participants. In mixed groups featuring a narrow range of abilities, medium-ability students fared much better. In this type of group, all students tended to participate actively, there were more questions eliciting help, and medium-ability students provided more explanations and demonstrated higher achievement. Homogeneous ability groups also had mixed success. In homogeneous high-ability

groups, students often assumed they all knew how to solve the problems. Such students provided fewer explanations and they ultimately performed worse on tasks than high-ability students in mixed groups. In homogeneous low-ability groups, students lacked sufficient skills and could not provide correct explanations to one another, performing worse than low-ability students in mixed groups.

In a more recent study, Webb et al. (1998) investigated how grade 7 and grade 8 students performed on a science assessment involving electrical circuits that was designed to measure students' conceptual understanding of voltage, resistance, current, and the relationships among them. Some of the students completed tasks individually, and some completed tasks in groups. Results were somewhat different from those observed in the meta-analysis. First, similar to the meta-analysis, results support the conclusion that for students at most ability levels, working in a heterogeneous group with at least one able member is more beneficial than working in a homogeneous group. In contrast with results from the meta-analysis, however, medium-ability students in heterogeneous groups did not perform worse than those in homogeneous groups. Webb et al. (1998) speculate that perhaps this is because moderate-ability students participated just as much in these groups as their high- and low-ability group mates. In the Webb et al. (1998) study, the effects of collaboration appear to be strongest for low-ability students, particularly when they are matched with high-ability classmates. Such students performed significantly better on both the group task and the individually-completed post-test than low-ability students working alone or working with less capable peers. Moreover, for below-average students the quality of group discussion, as indicated by the accuracy of answers and the quality of explanations, was a significant predictor of subsequent achievement. Such quality contributed more to their performance on the individual post-test than their actual ability. In contrast, high-

ability students performed better in homogeneous groups than they did in heterogeneous groups, and for high-ability students the quality of group interaction was unrelated to subsequent achievement. However, for high-ability students working in heterogeneous groups, group composition did not affect their performance. Thus, working with a low-ability student does not appear to significantly impair high-ability students. On the basis of these results, Webb et al. (1998) recommend the use of heterogeneous groups.

Group composition with respect to gender may also impact interaction patterns and moderate the effects of collaboration on student learning. For example, Webb (1991) found that in equally-balanced groups (i.e., groups with equal numbers of boys and girls), there were no differences in the interaction patterns of boys and girls. However, in majority male groups, girls were less successful in getting their questions answered than boys, and the researcher noticed that in these groups, boys tended to ignore the girls. Interestingly, even in majority female groups, girls tended to obtain less help because they tended to direct their requests for assistance to the boy in the group, who often ignored them. In both majority male and majority female groups, boys outperformed girls on subsequent individual achievement measures, despite the fact that boys and girls did not differ on ability.

Finally, some have suggested that task characteristics may moderate the effect of collaboration on group learning. For example, Mercer (1996) argues that whether the group task requires students to collaborate and communicate with one another in order to solve the problem will affect the quality of group discourse. Tasks should require planning, decision-making, and interpreting feedback. Partners should have to talk with one another to complete the task, and cooperative rather than competitive incentive structures should be used. Webb (1991) concurs, arguing that group reward structures—in which grades, tangible prizes, praise, or recognition are

based on the group's performance—are more likely to promote helping behaviors than individual reward structures. Dillenbourg et al. (1996) note that what is considered a desirable task feature may differ depending on what paradigm one is working within. For example, from a Vygotskian perspective, the most useful tasks will involve skill acquisition, joint planning, categorization, and memory. In contrast, from a Piagetian perspective, tasks for measuring conservation and coordination should involve perspective taking, planning, and problem solving.

The “interactions” paradigm.

The “interactions” paradigm developed in response to the complexities associated with the former paradigm and attempts to identify mediating mechanisms between collaboration and learning outcomes (Dillenbourg et al., 1996). In particular, this strand of research attempts to isolate characteristics and processes of interactions through which collaboration effects learning. For example, one proposed mediator of the effect of collaboration on learning is the extent to which social interactions produce elaborated explanations (Dillenbourg et al., 1996). Mercer (1996) argues that interactions producing elaborated explanations enable students to learn the principles underlying practical procedures and strategies, which can result in learning that is more generalizable and transferrable to new situations. Webb (1991) was one of the first researchers to collect evidence on the types of student interactions that occurred during collaborative learning. She meta-analyzed 17 studies of how collaborative or cooperative learning structures affected student learning outcomes in math. She found that content-related (elaborated) explanations positively and significantly correlated with subsequent math achievement in a majority of studies, with partial correlations ranging from 0.07–0.53. Receiving such explanations did not significantly improve subsequent achievement. Providing other kinds of assistance, such as giving the answer or solution without explaining or providing procedural

information about the task, was not related to math achievement, whereas receiving these other types of assistance was significantly and negatively related to achievement. Thus, the effect of collaborative learning on student achievement depends on the quality of the interactions. Verbal disagreements among students were found to have a U-shaped relationship with achievement, suggesting that either very low or very high levels of verbal disagreement tend to diminish subsequent achievement, whereas a moderate amount of such disagreements improves subsequent math performance.

Webb (1991) notes that the success of providing and receiving elaborated explanations to improving student learning depends on several factors, including whether the student receiving the explanation actually needs help, the relevance and timeliness of the information provided, whether the assistance is understood, whether the student has opportunities to practice new skills independently, and whether the student takes advantage of those opportunities. Although working in small groups with peers makes it more likely that assistance will be timely and understandable (i.e., in terms the student can relate to), other factors needed for successful interactions may not be present. For example, the student providing help may not allow the student receiving help to practice new skills independently. Alternatively, the person receiving help may lack the motivation to attempt to solve problems individually after receiving assistance (Webb, 1991).

The “computer supported” paradigm.

Finally, research emerging from the computer-supported collaborative learning paradigm has generally attempted to determine whether the theoretical benefits of collaborative learning in face-to-face settings can be realized through computer-mediated or computer-assisted interactions that are limited to asynchronous, text-based interactions. For example, Kreijns et al.

(2003) caution that, in contrast to face-to-face interactions, social interaction should not be taken for granted just because the technology to support interaction is there. With respect to asynchronous, text-based interactions, research suggests that there are a few differences between face-to-face collaborations and those conducted at a distance. First, certain interaction patterns may be more or less evident in one medium versus the other. For example, Curtis and Lawson (2001) found that in the online medium, there were fewer exchanges in which students challenged one another and more exchanges related to planning. The authors speculate that because students did not know one another prior to interacting, they may have felt less comfortable challenging the ideas of others. In addition, the online medium appeared to have made planning activities more important for coordinating work. Despite these differences between traditional and technology-enabled collaboration, the authors conclude that successful student collaboration in an online medium is possible.

Relationship to Other Concepts

Collaborative learning approaches are related to a number of other, so-called 21st century skills, including critical thinking, metacognition, and motivation. For example, a number of researchers have linked collaborative learning to the development of critical thinking (Bailin et al., 1999; Bonk & Smith, 1998; Heyman, 2008; Nelson, 1994; Paul, 1992; Thayer-Bacon, 2000). Definitions of critical thinking vary widely, but common elements of most definitions include the following component skills:

- analyzing arguments (Ennis, 1985; Facione, 1990; Halpern, 1998; Paul, 1992);
- making inferences using inductive or deductive reasoning (Ennis, 1985; Willingham, 2007; Paul, 1992; Facione, 1990);

- judging or evaluating (Case, 2005; Ennis, 1985, Facione, 1990; Lipman, 1988; Tindal & Nolet, 1995); and
- making decisions or solving problems (Ennis, 1985; Halpern, 1998; Willingham, 2007).

In addition to skills or abilities, critical thinking also entails dispositions. These dispositions, which can be seen as attitudes or habits of mind, include things such as open- and fair-mindedness, a propensity to seek reason, inquisitiveness, a desire to be well-informed, flexibility, and respect for and willingness to entertain diverse viewpoints (Bailin et al., 1999; Ennis, 1985; Facione, 1990; Halpern, 1998; Paul, 1992).

Collaborative or cooperative learning structures are argued to trigger critical thinking skills and likewise, students with strong critical thinking skills and dispositions may be better collaborators. Proponents of collaborative or cooperative learning include Thayer-Bacon (2000), who emphasizes the importance of students' relationships with others in developing critical thinking skills. Supporters also include Bailin et al. (1999), who argue that critical thinking skills involve the ability to respond constructively to others during group discussion, which implies interacting in pro-social ways by encouraging and respecting the contributions of others. Similarly, Heyman (2008) notes that social experiences can shape children's reasoning about the credibility of claims. Nelson (1994) provides some clues as to how collaboration can prompt cognitive development among college students. According to Nelson, students' misconceptions interfere with their ability to acquire new knowledge, despite appropriate instruction. Collaborations create opportunities for disagreements and misconceptions to surface and to be corrected. Collaboration also provides a vehicle for students to attain necessary acculturation to the college environment and helps to make tacit disciplinary expectations more explicit for students.

Collaborative learning is also related to metacognition, which has been defined most simply as “thinking about thinking.” Other definitions include the following:

- “the knowledge and control children have over their own thinking and learning activities” (Cross & Paris, 1988, p. 131);
- “awareness of one’s own thinking, awareness of the content of one’s conceptions, an active monitoring of one’s cognitive processes, an attempt to regulate one’s cognitive processes in relationship to further learning, and an application of a set of heuristics as an effective device for helping people organize their methods of attack on problems in general” (Hennessey, 1999, p. 3); and
- “the monitoring and control of thought” (Martinez, 2006, p. 696).

Several researchers have recommended collaborative learning approaches to teaching metacognition (Cross & Paris, 1988; Hennessey, 1999; Kramarski & Mevarech, 2003; Kuhn & Dean, 2004; Martinez, 2006; McLeod, 1997; Paris & Winograd, 1990; Schraw & Moshman, 1995; Schraw et al., 2006). Proponents include Cross and Paris (1988), who identify group discussions about the use of reading strategies as one critical feature of an intervention targeted at improving students’ metacognition during reading. Hennessey (1999) points out that collaborative approaches promote metacognitive discourse among students and stimulate cognitive conflict. Such conflict can lead to clarifications of students’ beliefs and concepts. Similarly, Kramarski and Mevarech (2003) attribute the superior performance of students working in collaborative group settings to the higher quality of discourse observed among students working together. Students participating in cooperative learning expressed their mathematical ideas in writing more ably than did those who worked alone. Moreover, as Schraw and Moshman (1995) note, peer interaction can encourage the construction and refinement of

metacognitive theories, which are frameworks for integrating cognitive knowledge and cognitive regulation. Kuhn and Dean (2004) argue that social discourse can cause students to “interiorize” processes of providing elaborations and explanations that have been associated with improved learning outcomes. In turn, those with strong metacognitive skills may model self-regulated learning better for their group mates than those with poor metacognitive skills.

Finally, collaborative approaches may be related to motivation. Paraphrasing Gredler, Broussard and Garrison define motivation as “the attribute that moves us to do or not to do something” (2004, p. 106). Gottfried defines academic motivation in particular as the “enjoyment of school learning characterized by a mastery orientation; curiosity; persistence; task-endogeny; and the learning of challenging, difficult, and novel tasks” (1990, p. 525). Bossert (1988) argues that motivation is one of the potential mediating processes whereby cooperative learning affects achievement. According to Bossert, peer encouragement may improve task engagement, and the novelty of collaborative learning tasks causes students to shift attentional resources. Hidi and Harackiewicz (2000) frame the issue in terms of situational interest. According to this perspective, working with others is a way to enhance situational interest that can ultimately trigger personal or individual interest. As Turner (1995) argues, collaboration provides opportunities for students to experience disequilibrium, which can spur curiosity and interest. Second, collaboration provides opportunities for peer modeling, and models of successful student performance can be more motivating to students than teacher modeling. Finally, working with others promotes academic engagement through the added responsibility of group performance, which causes individuals to persist at difficult tasks longer than they normally would.

Development of Collaboration Skills

This section reviews the limited theoretical and empirical literature on development of collaborative capacities, including how collaboration skills first appear and develop over time.

According to Tudge (1992), early work by Piaget and Vygotsky is informative with respect to development of collaboration skills in young children. From a Piagetian perspective, children younger than 7 may lack the developmental skills to benefit from collaboration because they have not reached the concrete operational stage, or the stage at which logical reasoning first appears and children begin to apply mental operations to concrete problems such as conservation tasks. Once children have attained this stage, however, they appear to benefit from collaboration. For example, research in the Piagetian tradition suggests that when conservers are paired with non-conservers on a conservation task, non-conserving members are highly likely to reach conservation, whereas conserving members are very unlikely to regress as a result of interaction (as summarized in Tudge, 1992). Dillenbourg et al. (1996) similarly observe that Piaget's theory leads to specific expectations for development of collaboration skills. For example, pre-operational children may lack the ability to benefit from collaboration because they cannot de-center from their own perspective, suggesting they may have difficulty recognizing the views of others. Similarly, preschool-age children may lack the ability to sustain discussions of alternative hypotheses. For collaborations to produce the interactions necessary to support learning, children serving as "tutors" must be skilled at the task and must be able to reflect on their own performance. The authors point out that even if young children are able to serve as skilled tutors to their less able peers, 5- and 6-year-olds may not have the ability to inhibit their own actions enough to allow someone else to learn something they themselves already know how to do (Dillenbourg et al., 1996).

As Tudge (1992) notes, Vygotsky did not identify particular stages at which children would be ready to collaborate. Rather, the Vygotskian tradition emphasizes the benefits of collaboration whenever participants differ in terms of their initial skill levels. Typically, those working in this paradigm have conceptualized collaboration as occurring between a child and a more competent peer, either another child or an adult (as summarized by Tudge, 1992). Any collaborative learning activity within the child's zone of proximal development will potentially improve learning. In Tudge's study, significant age effects were found with respect to the sophistication of decision rules invoked by student pairs. Thus, older children uniformly used more sophisticated reasoning than younger children. However, the effects of collaboration were the same for 5-year-olds as for 9-year-olds, suggesting that young children are as likely as older children to be impacted by collaboration, and the outcomes of such interactions tend to be similar.

Instructional Implications

This section reviews both general and specific instructional recommendations for helping students improve their ability to collaborate.

General Approaches to Teaching Collaboration

Few studies investigate whether students can be successfully trained to collaborate. As Bossert (1988) observes, "specific training in cooperative roles is not offered in most studies of cooperative learning methods: The activity itself constitutes the training" (p. 227). However, many researchers recommend providing explicit instruction in collaboration skills (Fall et al., Webb, 1995). For example, educators are urged to devote explicit instruction to developing collaboration skills. Such training could include instruction in effective communication, how to

seek help, and how to provide help to others (Fall et al., 1997). Similarly, Webb (1991 and 1995) recommends training students in general interpersonal and teamwork skills, including coordination, communication, conflict resolution, decision making, problem solving, and negotiation. Such training could emphasize how to give explanations, how to directly and explicitly ask for help, and how to respond appropriately to others' requests for help. Teachers should also provide ample opportunities for students to practice collaboration skills, using tasks that are similar to those used during group-based assessments. Teachers should encourage students to actively participate during group work (Fall et al., 1997). Teachers should also emphasize that multiple skills are necessary to complete group tasks and each person in the group is going to be skilled in at least one area (Webb, 1995).

In addition to preparing students for collaboration by providing explicit instruction, teachers should also structure tasks to support collaboration (Bossert, 1988; Dillenbourg, 1999; Mercer, 1996; Webb, 1995). For example, teachers can embed specific roles within tasks (Dillenbourg, 1999; Webb, 1995). These roles can be based on knowledge complementarities or on conflicting viewpoints. Dillenbourg points out, however, that decomposition of the task into independent sub-tasks reduces the level of collaboration required. Thus, individual student roles should define horizontal rather than vertical division of labor. For example, one student may assume responsibility for the task level, whereas the other group member oversees meta-task aspects (e.g., planning). Webb (1995) describes roles such as "learning leader," responsible for summarizing and recounting the main points of the material, and "learning listener," responsible for detecting errors or omissions in the summary and asking questions to clarify the material.

Teachers can also specify rules for interaction, requiring, for example, that every group-member ask at least one question (Dillenbourg, 1999). Mercer (1996) argues that when teachers

establish “ground rules” for collaboration, student motivation and performance improve. Such ground rules can include sharing all relevant information and suggestions, providing reasons to back up assertions and suggestions, asking for reasons where appropriate, agreeing about what action to take, and accepting that the group rather than the individual is ultimately responsible for decisions and actions. In the Mercer study, when teachers provided such ground rules for guiding interactions, student interactions featured higher-quality discourse in comparison to classrooms without such rules. In particular, students were more likely to engage critically and constructively with one another, making their reasoning and justifications visible for others to evaluate.

Teachers should also monitor and regulate such interactions. For example, Fall et al. (1997) found that when teachers actively circulate among groups and encourage students to share their ideas, students are more engaged and discussion is more fruitful. In addition, Tudge’s study demonstrates the importance of providing groups with feedback to confirm or disconfirm the group’s direction (1992). In the absence of tools for monitoring interactions at different times and places (e.g., an automated computer interface) teachers are encouraged to provide group members with tools to monitor and evaluate their own interactions, a topic that will be explored more fully in the section on assessing collaboration (Dillenbourg, 1999).

Specific Instructional Techniques

There are a variety of ways to structure collaborative learning activities. For example, Palincsar (1987) describes reciprocal teaching as “an interactive teaching procedure in which the teachers and students collaborate in the joint construction of text” (p. 2). During reciprocal teaching, teachers and students take turns assuming “leader” and “respondent” roles, with the

leader employing several strategies to direct discussion: asking questions, summarizing responses, clarifying misunderstandings, and supporting predictions about upcoming text content.

Bossert (1988) describes a number of specific formats, which use various reward and task structures to induce student collaboration. “Learning together,” for example, has heterogeneous groups of 4–5 students work together to complete a single group lesson. Students are instructed to help each other until the entire group has learned the material. Bossert refers to this approach as “pure” cooperation because student groups do not compete against one another. In contrast, the “jigsaw” technique breaks academic material to be learned into portions. Each student learns a portion of the material in an “expert group” comprised of others assigned to the same material. Students then return to their original groups and share what they have learned. In this approach, students can be tested as a group or individually on their knowledge of the material. “Group investigation” is similar to “jigsaw,” except students have more freedom to choose which topics they will pursue and assign their own division of labor. In “student team learning,” students work in small, mixed-ability groups to help prepare one another for quizzes that student teams compete on against one another. Students contribute to their team’s performance through their own performance on the quiz. Teams with high levels of mean group performance are recognized. In “team assisted individualization,” each student progresses at his or her own pace, but can be helped by group mates. Students must take mastery tests before progressing to more difficult material, and their scores on these tests contribute to their team’s scores. A number of other methods also structure group interactions. The important point to note is that teachers should manipulate reward and task structures in ways that will accomplish the specific learning goals they have for students. This point will be elaborated in the next section.

Assessment Implications

This section reviews challenges in assessing collaboration, describes extant methods of assessing or measuring collaboration, and reviews guidelines from the literature for measuring collaboration.

Challenges in Assessing Collaboration

There are a number of challenges in attempting to measure students' collaboration skills. First, educators are usually interested in obtaining individual student scores or grades. Collaborative learning, because it occurs in a group, can obscure individual contributions, making it difficult to isolate individual student scores. Often teachers will assign a single score to a group based on completion of a group product, and this group score in turn is assigned to each individual group member. However, to the extent that individual student contributions are not comparable or equivalent, this does not necessarily reflect individual effort, knowledge, or skills (Race, 2001; Saner et al., 1994; Webb, 1995). Research shows that group-level assessments may not yield scores that are predictive of individual-level ability, even when individual students turn in separate products. In particular, scores from group work tend to over-estimate individual performance, and exhibit both ceiling effects and range restriction (Webb, 1993). Fall et al. (1997) demonstrate that the effect of collaboration on performance is evident even with very limited interaction (e.g. a 10-minute discussion regarding themes in a reading passage). Even when individual group mates turn in separate products, scores between group members tend to be correlated at very high levels (.66–.83), suggesting that they are not independent of one another (Saner et al., 1994). In addition, there is a carry-over effect, with scores on subsequent individual work affected by previous group collaboration, particularly for low achievers (Saner

et al., 1994; Webb, 1998). If this higher performance reflects real learning, then group assessment scores are not necessarily invalid. However, if students obtain higher scores from group assessments simply because their more able group mates complete work for them, then using group scores as indicators of individual student learning is problematic.

Webb (1995) details four potential purposes of group-based assessments. First, educators may be trying to measure individual student learning, as evidenced by student knowledge or skill, in the context of a group activity. Second, the goal may be to assess an individual student's ability to learn from collaboration, which is typically accomplished by including both individual and group assessment components. A third goal is to assess group productivity, as evidenced by the quality or quantity of a product completed collaboratively. Finally, educators may be trying to measure a student's collaboration skills, such as coordination, communication, decision-making, conflict resolution, and negotiation.

As Webb notes, processes that facilitate good performance under one goal may be counter-productive for another. For example, group processes such as co-construction of ideas, conflict, giving and receiving elaborated help, equality of participation, division of labor, and free riding or social loafing are more or less desirable depending on whether the goal of group assessment is to measure individual student learning or group productivity. If the goal is to measure individual learning, then group processes of co-constructing ideas, resolving cognitive conflicts, giving and receiving elaborated explanations, and ensuring equality of participation should be encouraged through the structure of group tasks, because these processes are associated with student learning. On the other hand, if the goal is to measure group productivity, then the above group processes would actually be counter-productive to the extent that they impede progress on completing a high-quality group product within a set amount of time. As

Webb explains, if the task is well-defined and has one correct answer or solution, attempting to involve low-ability students may actually slow the group down. In this case, “it may be more effective for students to work separately instead of together, for one or more group members to do most of the work while others contribute little, for one student to take control of group work if group members cannot agree, and to have minimal helping behavior” (p. 249). In other words, group processes that are considered maladaptive for learning, such as free riding or social loafing, can actually be effective for groups if the goal is to maximize productivity. Webb cautions that educators need to be very clear about the intended purposes of group-based assessments, and assessment tasks should be structured to support those goals.

Extant Assessment Methods

Researchers studying collaboration have used a variety of observational tools to capture student interactions. Dillenbourg et al. (1996) base their framework on conversation models developed in the field of linguistics. The authors are primarily interested in student negotiation, which they define as a process by which students attempt to attain agreement on aspects of the task or interaction. They identify four types of negotiation behaviors that can be observed during interaction: mutual adjustment occurs when partners refine their positions; competitive argumentation occurs when one partner tries to convince the other to adopt his position; “standing pat” is when one student uses another as a resource by eliciting a proposal; and negotiation of meaning is evidenced by “repair sequences,” in which misunderstandings become evident and are explicit targets of discussion. Roschelle’s coding system for describing collaborative interactions has four components: construction of a “deep-featured” situation at an intermediate level of abstraction, the interplay of metaphors, an iterative cycle of displaying, confirming, and repairing conceptions, and the application of progressively higher standards of

evidence for convergence of understandings (1992). Within this progression of evidence, simple affirmative acknowledgement is the lowest level, followed by verbatim recitation of a concept or inference rule, collaborative completion of a mutually satisfactory explanation, and finally, completion of partners' utterances, paraphrasing, and elaboration of emerging thoughts.

Roschelle and Teasley (1995) have delineated a number of conversational strategies for achieving shared understanding. These strategies are said to indicate deep levels of collaborative interaction and include taking turns, socially distributed productions, repairs, narrations, and nonverbal actions or gestures. For example, socially distributed productions occur when a compound sentence is distributed over conversational partners. One example is a collaborative completion of an "if-then" statement. These types of interactions are particularly useful for constructing shared knowledge. Repairs are the methods by which periodic breakdowns in intelligibility are reconciled, and are a major means for consolidating understanding. Narrations occur when one partner "thinks aloud" to another to make thinking and reasoning processes explicit. Finally, nonverbal actions indicative of converging understandings can include hand gestures, such as pointing, or production of an appropriate action that signals understanding of an ambiguous utterance.

Mercer (1996) recorded around 60 hours of classroom talk with 50 children between the ages of 5 and 13 while they worked in small groups on collaborative tasks. He categorized student talk into three types: disputational talk, cumulative talk, and exploratory talk. Disputational talk is characterized by disagreement and individualized decision-making, with few attempts to pool resources, or to offer or accept constructive criticism. Typically, this type of interaction is exemplified by short exchanges consisting of assertions and counter-assertions. In cumulative talk, speakers build positively but uncritically on what the other has said. This type of

interaction is characterized by repetitions, confirmations, and elaborations. Finally, exploratory talk is when partners engage critically but constructively with one another. Students engaging in such talk offer statements and suggestions for joint consideration. These may be challenged, but all arguments are justified and reasons are provided.

Finally, Webb has developed systems for coding types of student communication as well as for categorizing students on the basis of their group behavior. First, Webb (1991) distinguishes between responsive feedback and nonresponsive feedback during collaborative group work. Responsive feedback entails substantive corrections, elaborations, and explanations. Nonresponsive feedback, on the other hand, occurs when student errors are not corrected, when students are told the correct answer with no elaboration, or when student requests for help receive no response at all. Webb (1993) proposes four categories of student behavior. The first type of student solves problems correctly aloud with little or no assistance from others. The second type of student expresses difficulty with the problems, either by making errors or asking questions that indicate they are confused. Students in the third category copy other students' work without doing it themselves. Students in category four do not contribute verbally to the group discussion at all. In a more recent study, Webb et al. (1998) code contributions of individual students to group discussion according to their cognitive level. Thus, high-level participation includes making or defending suggestions for how to answer a particular item, asking questions about a suggestion, or paraphrasing a suggestion. Medium-level participation includes copying someone else's suggestion, repeating it verbatim, or agreeing with what was said without further elaboration. Low-level participation entails listening or watching without making any substantive contribution or inquiry. Finally, students who manipulate materials or write answers without referencing other group members are coded as working alone.

General Suggestions for Designing Group-Based Assessments

Webb's (1995) framework provides several suggestions for designing group-based assessments, depending on the particular purpose. For example, if the purpose of assessment is to measure individual student learning, then group-based assessments should not be used at all. The best measure of individual student knowledge and skills is an individual student assessment, and even a small amount of collaboration will invalidate measures of individual learning. If the purpose is to measure an individual's ability to learn from collaboration, then assessments should be multi-staged, with both individual and group portions. Ideally, individual portions will occur both before and after group portions, creating a collaboration sandwich of sorts. In addition, Webb urges educators to stress individual accountability, require all students to be prepared to summarize, explain, and justify group work, and inform students that actively participating in group work (e.g., asking questions and explaining their thinking) will result in better learning. If the purpose of group assessment is to measure group productivity, educators should provide a task to complete and focus evaluation on the completed task rather than individual student contributions or student interactions. Finally, if the purpose is to measure student ability to collaborate, students should be encouraged to exchange ideas, opinions, and knowledge, and to help one another, work together, actively seek help, justify and explain ideas, and give elaborated explanations. Further, evaluation should focus on qualities of student interactions rather than the quality or quantity of the group product or solution.

If the focus of the group-based assessment is on an individual's ability to learn from others or collaborate, then educators need to create opportunities to observe and score individual performances. For the former, this could involve combining group and individual portions of the task. For example, in one study, the task was divided into three stages. During Stage 1, students

worked individually to brainstorm factors influencing the role of yeast in food. During Stage 2, students worked in collaborative small groups on a complex task that involved designing and conducting an experiment investigating the behavior of yeast in foods. Finally, in Stage 3, students again worked individually on a similar task, with each student evaluating and critiquing a lab report describing an experiment conducted by one of the other groups. Structuring tasks in this way provides educators with information regarding students' initial knowledge or skill, their performance in a group context, and their knowledge or skill after having collaborated. This type of information corresponds directly to Vygotsky's notion of the zone of proximal development, in that it identifies both what an individual can accomplish alone, and what they can accomplish with the help of more competent peers.

For the latter situation, in which educators are trying to measure individual student ability to collaborate, educators may find it difficult to observe all students during group interactions, as this would require being in multiple places at once. In some situations, automated computer interfaces, such as those used in some interactive learning environments, can be used to monitor and regulate interactions between students during group work (Dillenbourg, 1999). However, in other situations this type of automated monitoring is not feasible. In this case, both Dillenbourg (1999) and Race (2001) recommend supplementing teacher observations with peer- and/or self-assessments of collaborative skills. For example, Race notes the benefits of pairing group-based assessment with either self- or peer-assessment of learning. First, students already assess themselves and their peers naturally. Second, sometimes students are in a better position to judge the quality of their own and others' contributions because the learning processes are more salient to them. Evaluating oneself or one's peers can improve learning, to the extent that evaluating others' work is like providing an elaborated explanation. Moreover, being able to take stock of

their own learning is a skill that is important for lifelong learning, and students may receive more feedback from peers than from their teacher. To have students assess their own ability to collaborate, students could be given a questionnaire to complete and submit along with any products of group work. This questionnaire could include questions such as the following:

- What do you think is a fair score for your contribution?
- What was the thing you did best during collaboration?
- What was the thing you did least well during collaboration?
- What was the hardest part of collaborating?
- What was the most important thing you learned?

Peer assessment of collaboration skills could be accomplished by having each student rate his or her group mates, with ratings remaining anonymous. Criteria for rating peer performance should reflect important collaborative learning outcomes, such as taking turns, being a good listener, and communicating respectfully.

Educators are also urged to use reward structures that support the particular assessment goals they want to accomplish. Bossert (1988) defines reward structures as the degree to which students are dependent on one another for reinforcement or recognition. Reward structures essentially establish incentives for students that will motivate them to engage in desirable or undesirable group processes to varying degrees. Bossert identifies several different types of reward structures. For example, cooperative reward structures are when individuals obtain rewards in direct proportion to other members of their group. As Bossert notes, “cooperative reward structures are supposed to motivate group members to help one another, foster positive

group pressures that maintain each member's involvement in task completion, and reinforce students' perception that they share a common goal” (p. 227). Bossert distinguishes two types of cooperative structures: those in which group rewards are independent of individual contributions and those in which group rewards are contingent on individual performance. In contrast to cooperative reward structures, competitive structures occur when individuals obtain rewards in inverse proportion to other members of their group, whereas individualistic structures are those in which individual rewards are uncorrelated with the rewards received by group mates (Bossert, 1988). Bossert argues that combining contingent rewards with individual accountability counteracts the tendency toward maladaptive group processes, such as social loafing or free riding. Webb (1991) concurs, arguing that tasks should utilize group reward structures that reward students based on the achievement of all group members. Such structure promotes helping behaviors and creates incentives for all group members to learn the material.

Race (2001) offers a number of different approaches to scoring the products of collaboration that vary in the extent to which the group product is emphasized relative to individual contributions. The simplest approach is to assign all members of the group the same score, but this can often be perceived as unfair if group members make different contributions. Another approach is to completely divide the group task into portions and score each person's individual contribution. However, as Race points out, it can be difficult to divide a single group task in a way that ensures all sub-tasks are equally difficult. Moreover, dividing the task in this way makes the interaction less collaborative. Another method involves assigning a score for the overall group product, and then negotiating differentials with individual students by asking the group to divide up the points. This approach is usually perceived by students to be fair, but should not be implemented with immature students. A related approach is to assign a single score

for the overall group product, and then ask individuals to peer-assess one another's individual contributions. Teachers could also assign additional tasks to each member of the group beyond the group task. For example, individual students can be tested on material to be learned, either orally or in writing. As Race argues, when students know they will be individually accountable for learning, they tend to engage in more desirable group processes.

Although intra-group competitive reward structures are argued to be detrimental to cooperative learning, inter-group competition may help foster greater cooperation and cohesion among group members (Bossert, 1988). In other words, when group activities are structured so that groups or teams compete against one another, particularly when mixed groups of roughly equal ability are used, cooperation and collaboration can actually increase. Bossert observes that there is a great deal of debate surrounding the value of inter-group competition, with critics arguing that such competition will always result in some "losing groups," which can damage motivation and self-esteem. Furthermore, such structures emphasize the extrinsic value of participating in learning activities rather than the intrinsic value. However, proponents argue that such structures are necessary to achieve truly deep levels of cooperation and collaboration. Summarizing the research on both sides of the debate, Bossert concludes that there is not enough evidence to resolve the question.

Another suggestion for creating group assessments concerns group composition. Educators are urged to carefully consider group composition when creating collaborative groups or teams (Fall et al., 1997; Webb, 1991, 1993, and 1995). Given research results suggesting that factors such as student ability and gender can impact both patterns of interaction and the effects of collaboration, educators should attempt to create groups of students that can work productively together. First, the bulk of research on collaboration suggests that heterogeneous

groups, or groups with a range of ability levels, function best. However, depending on the type of task, groups with a narrow range of ability—combining high-ability students with moderate-ability and moderate-ability students with low-ability—may be more effective than groups featuring a wide range of ability (i.e., combining high-, moderate-, and low-ability students in a single group). Moreover, groups with balanced genders may function better than groups that are male- or female-dominated. Finally, other characteristics may also affect interaction patterns and the success of collaborative learning. These factors, called status characteristics, can include race, popularity, attractiveness, and perceived intelligence or achievement (Webb, 1995). Webb notes that it is very difficult to construct groups that are completely balanced, because it requires the teacher to keep track of status variables and to consider combinations of status variables simultaneously. This becomes even more challenging in relatively homogeneous classrooms with small pockets of student diversity. Webb argues that at the very least, teachers should attempt to ensure that each group has at least one able member. Training students to communicate and collaborate productively with one another may help to mitigate the negative consequences of having unbalanced groups (Webb, 1995).

One final recommendation concerns characteristics of group tasks that will likely provoke more desirable group interactions. First, tasks should create incentives for all group members to participate in order to combat the problem of free riding (Salomon & Globerson, 1989; Webb, 1995). Webb et al. (1998) point out that most studies of group productivity have used simple tasks with one correct solution that can easily be solved by a single, competent group member. As Dillenbourg (1999) observes, such trivial, obvious, and unambiguous tasks provide little need for students to negotiate, and thus offer few opportunities to observe real collaboration. In contrast, two types of tasks that might encourage all members to participate are

ill-structured tasks and “group” or additive tasks. Ill-structured tasks are those with very little structure, no clear solution strategies, and more than one correct solution. Group tasks are those that require knowledge, information, skills, and strategies that no single individual is likely to possess. Similarly, Salomon and Globerson (1989) define additive tasks as those “where performance depends on the maximal contribution of all members,” giving the example of a tug-of-war game. Both group or additive tasks and ill-structured tasks are more likely to encourage full group participation to the extent that the task cannot be completed by a single, competent person. In fact, Webb et al. (1998) observe that when complex, ill-structured tasks are used, all group members are more likely to participate actively, even in groups with a range of abilities.

Summary

Roschelle and Teasley define collaboration as the “mutual engagement of participants in a coordinated effort to solve a problem together” (as cited in Dillenbourg et al., 1996, p. 2). Collaboration can be seen as “coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” (Roschelle & Teasley, 1995, p. 70). Research on collaboration has developed within three distinct strands: research that compares group performance to individual performance, studies identifying the conditions under which collaboration is more or less effective, and research investigating the characteristics of interactions that mediate the impact of collaboration on learning, including use of new technologies that facilitate asynchronous text-based interactions. Such research suggests that collaborative interactions are characterized by shared goals, symmetry of structure, and a high degree of negotiation, interactivity, and interdependence. Interactions producing elaborated explanations are particularly valuable for improving student learning, especially for the student

providing such explanations. Nonresponsive feedback, on the other hand, can be detrimental to student learning in collaborative situations.

Collaboration can have powerful effects on student learning, particularly for low-achieving students. These effects are seen in the form of higher scores on work completed collaboratively, even when students turn in separate products. In addition, there appears to be a carry-over effect, such that individual performance on subsequent measures of achievement tends to be higher for students exposed to collaborative learning. However, a number of factors may moderate the impact of collaboration on student learning, including student characteristics, group composition, and task characteristics. For example, patterns of interaction as well as the effects on subsequent performance vary across males and females, with boys participating more actively and appearing to benefit more from collaborative learning than girls. Similarly, high-ability students may participate more actively than low-ability students. Group composition, with respect to gender and ability, is also an important factor. Thus, heterogeneous groups featuring a narrow ability range appear most successful, as do groups that have a balance of girls and boys. Finally, task characteristics, such as the degree of role interdependence, and task and reward structures can impact the types of group processes used.

Collaboration is linked to a number of important educational outcomes, including critical thinking, metacognition, and motivation. Collaborative learning structures are argued to spur development of critical thinking, to the extent that they stimulate cognitive conflict and disequilibrium. Likewise, students with strong critical thinking skills and dispositions, including the ability to consider multiple perspectives, may be better collaborators. Collaborative approaches also promote metacognitive discourse among students to the extent that students are able to interiorize processes of providing elaborated explanations and make their thinking and

reasoning visible. In turn, students with strong metacognitive skills can serve as powerful models of self-regulated learning for their group mates. Finally, collaborative learning may enhance motivation, because collaborative tasks signal novelty that shifts attention, and working with others triggers situational interest and curiosity.

Although historical frameworks, such as those provided by Piaget and Vygotsky, offer some guidance as to when and how young children acquire and develop collaboration skills, there is scant empirical evidence to support such predictions. Available research suggests that children as young as 5 are as susceptible to influence through collaborative learning, as are older students. Similarly, few studies investigate whether students can be successfully trained to collaborate well. However, many researchers appear to believe they can, and they urge educators to provide explicit instruction encouraging development of collaboration skills, such as coordination, communication, conflict resolution, decision making, problem solving, and negotiation. Such training should also emphasize desirable qualities of interaction, such as providing elaborated explanations, asking direct and specific questions, and responding appropriately to requests from others. Teachers should structure tasks in ways that support the goals of collaboration. Such structure can be accomplished by embedding specific roles within tasks that decompose the task horizontally, into, for example, task and meta-task levels. Teachers should also specify “ground rules” for interaction and monitor and regulate such interactions.

There are a number of challenges in assessing collaboration or assessing learning in collaborative group settings. First, educators are typically interested in obtaining individual student scores, but group assessment, by its very nature, obscures individual contributions. A wealth of empirical evidence demonstrates that work completed collaboratively cannot be considered equivalent to work completed individually, as scores from group work tend to over-

estimate individual performance. Even when group mates turn in separate work products and even with a limited amount of collaboration, scores between group members cannot be considered independent of one another. There is also a carry-over effect, particularly for low-achievers, who tend to earn higher scores on subsequent achievement measures than similar-ability students working alone.

Another challenge associated with assessments occurring in collaborative contexts is that they can fulfill different purposes, and group processes that facilitate good performance under one goal can be counterproductive for another goal. For example, if assessment aims to measure student ability to learn from collaboration, then group processes such as co-construction of ideas, conflict, giving and receiving elaborated help, and equality of participation should all be encouraged. In contrast, these processes may be counter-productive if the goal of group assessment is to measure group productivity. In this case, group processes that are considered maladaptive from a student learning perspective—such as division of labor, social loafing, or free riding—may actually enhance group performance. Thus, educators are urged to be clear about the purposes of group assessment, and to encourage processes that will support the intended goals.

Educators should align assessment design with intended purposes and goals. Group-based assessments should not be used as indicators of individual student learning, because empirical research suggests that group products are not representative of individual student knowledge, skill, or ability. In addition, if teachers are interested in obtaining individual student scores—either student ability to collaborate or ability to learn from collaboration—they will need to create opportunities within the task for observing and scoring individual student performances. For the former, teachers can supplement their own observations of group interactions with peer-

and/or self-assessment of collaboration skills. For the latter, teachers should select or create tasks that entail individual and group elements—ideally, group performance should be “sandwiched” between individual performance. Structuring tasks this way provides teachers with information relevant to the “zone of proximal development,” highlighting both what students can accomplish individually and what they can accomplish with help from a more competent peer.

Teachers should also use reward structures that support the particular goals they are trying to accomplish. Cooperative reward structures in which individuals obtain rewards in direct proportion to other members of their group appear to facilitate better group processes. In particular, combining contingent rewards with individual accountability for learning counteracts the tendency toward maladaptive group processes, such as free riding. Such structures promote helping behaviors and create incentives for all group members to learn the material. Although intra-group competition should be minimized to promote cooperative behaviors, inter-group competition may be used to spur student motivation and group cohesion, although empirical evidence is inconclusive with respect to the efficacy of such competition. Teachers need to carefully consider group composition in terms of gender and ability when using collaborative learning. Heterogeneous groups featuring a narrow range of ability and a rough balance between males and females appear to be most supportive of desirable group processes. In addition, educators should select complex, ill-structured, and/or so-called “group” tasks that cannot be solved by a single group member. Simple tasks with one correct solution that can easily be solved by a single, competent group member should not be used in collaborative settings.

References

- Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies, 31*(3), 285–302.
- Bonk, C. J. & Smith, G. S. (1998). Alternative instructional strategies for creative and critical thinking in the accounting curriculum. *Journal of Accounting Education, 16*(2), 261-293.
- Bossert, S. T. (1988). Cooperative activities in the classroom. *Review of Research in Education, 15*(1988-1989), 225–250.
- Broussard, S. C., & Garrison, M. E. B. (2004). The relationship between classroom motivation and academic achievement in elementary school-aged children. *Family and Consumer Sciences Research Journal, 33*(2), 106–120.
- Case, R. (2005). Moving critical thinking to the main stage. *Education Canada, 45*(2): 45–49.
- Cross, D. R. & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology, 80*(2), 131–142.
- Dillenbourg, P. (1999). What do you mean by 'collaborative learning?' In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1–19). Oxford: Elsevier.
- Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189-211). Oxford: Elsevier.
- Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership, 43*(2), 44–48.

- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: The California Academic Press.
- Fall, R. Webb, N., & Chudowsky, N. (1997). *Group discussion and large-scale language arts assessment: Effects on students' comprehension*. CSE Technical Report 445. Los Angeles, CRESST.
- Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology, 82*(3), 525–538.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist, 53*(4), 449–455.
- Hennessey, M. G. (1999, March). Probing the dimensions of metacognition: Implications for conceptual change teaching-learning. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA.
- Heyman, G. D. (2008). Children's critical thinking when learning from others. *Current Directions in Psychological Science, 17*(5), 344–347.
- Hidi, S. & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*(2), 151–179.
- Kramarski, B. & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal, 40*(1), 281–310.
- Kuhn, D. & Dean, D. (2004). A bridge between cognitive psychology and educational practice. *Theory into Practice, 43*(4), 268–273.
- Lipman, M. (1988). Critical thinking—What can it be? *Educational Leadership, 46*(1), 38–43.

- Martinez, M. E. (2006). What is metacognition? *Phi Delta Kappan*, 87(9), 696–699.
- McLeod, L. (1997). Young children and metacognition: Do we know what they know they know? And if so, what do we do about it? *Australian Journal of Early Childhood*, 22(2), 6–11.
- Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, 6(4), 359–377.
- Nelson, C. E. (1994). Critical thinking and collaborative learning. *New Directions for Teaching and Learning*, 1994(59), 45–58.
- Palincsar, A. S. (1987, April). Collaborating for collaborative learning of text comprehension. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.
- Paris, S. G. & Winograd, P. (1990). Promoting metacognition and motivation of exceptional children. *Remedial and Special Education*, 11(6), 7–15.
- Paul, R. W. (1992). Critical thinking: What, why, and how? *New Directions for Community Colleges*, 1992(77), 3–24.
- Race, P. (2001). *A briefing on self, peer, and group assessment. Assessment Series Number 9*. York, UK; Learning and Teaching Support Network.
- Roschelle, J. (1992) Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences*, 2, 235–276.
- Roschelle, J. & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem-solving. In C.E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–97). Berlin: Springer-Verlag.

Salomon, G. & Globerson, T. (1989). When teams do not function the way they ought to.

International Journal of Educational Research, 13(1), 89–100.

Saner, H., McCaffrey, D., Stecher, B., Klein, S., & Bell, R. (1994). The effects of working in pairs in science performance assessments. *Educational Assessment*, 2(4), 325–338.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36, 111–139.

Schraw, G. & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371.

Thayer-Bacon, B. J. (2000). *Transforming critical thinking: Thinking constructively*. New York: Teachers College Press.

Tindal, G. & Nolet, V. (1995). Curriculum-based measurement in middle and high schools: Critical thinking skills in content areas. *Focus on Exceptional Children*, 27(7), 1–22.

Tudge, J. R. H. (1992). Processes and consequences of peer collaboration: A Vygotskian analysis. *Child Development*, 63(6), 1364–1379.

Turner, J. C. (1995). The influence of classroom contexts on young children's motivation for literacy. *Reading Research Quarterly*, 30(3), 410–441.

Van Boxtel, C., Van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction*, 10(4), 311–330.

Webb, N.M. (1991). Task-related verbal interaction and mathematical learning in small groups. *Research in Mathematics Education*, 22(5), 366–389.

Webb, N. M. (1993). Collaborative group versus individual assessment in mathematics: Processes and outcomes. *Educational Assessment*, 1(2), 131–152.

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis, 17*(2), 239–261.

Webb, N. M., Nemer, K. M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. *American Educational Research Journal, 35*(4), 607–651.

Willingham, D. T. (2007). Critical thinking: Why is it so hard to teach? *American Educator, 31*(2), 8–19.