Cooperative Learning Methods: A Meta-Analysis

David W. Johnson, Roger T. Johnson, and Mary Beth Stanne

University of Minnesota

60 Peik Hall

159 Pillsbury Drive, S.E.

Minneapolis, Minnesota 55455

May, 2000

Running Head: Cooperative Learning Methods

Abstract

Cooperative learning is one of the most widespread and fruitful areas of theory, research, and practice in education. Reviews of the research, however, have focused either on the entire literature which includes research conducted in non-educational settings or have included only a partial set of studies that may or may not validly represent the whole literature. There has never been a comprehensive review of the research on the effectiveness in increasing achievement of the methods of cooperative learning used in schools. An extensive search found 164 studies investigating eight cooperative learning methods. The studies yielded 194 independent effect sizes representing academic achievement. All eight cooperative learning methods had a significant positive impact on student achievement. When the impact of cooperative learning was compared with competitive learning, Learning Together (LT) promoted the greatest effect, followed by Academic Controversy (AC), Student-Team-Achievement-Divisions (STAD), Teams-Games-Tournaments (TGT), Group Investigation (GI), Jigsaw, Teams-Assisted-Individualization (TAI), and finally Cooperative Integrated Reading and Composition (CIRC). When the impact of cooperative lessons was compared with individualistic learning, LT promotes the greatest effect, followed by AC, GI, TGT, TAI, STAD, Jigsaw, and CIRC. The consistency of the results and the diversity of the cooperative learning methods provide strong validation for its effectiveness.

Cooperative Learning Methods: A Meta-Analysis

Cooperative learning is one of the most remarkable and fertile areas of theory, research, and practice in education. Cooperative learning exists when students work together to accomplish shared learning goals (Johnson & Johnson, 1999). Each student can then achieve his or her learning goal if and only if the other group members achieve theirs (Deutsch, 1962). In the past three decades, modern cooperative learning has become a widely used instructional procedure in preschool through graduate school levels, in all subject areas, in all aspects of instruction and learning, in nontraditional as well as traditional learning situations, and even in after-school and non-school educational programs. There is broad dissemination of cooperative learning through teacher preparation programs, in-service professional development, and practitioner publications. The use of cooperative learning so pervades education that it is difficult to find textbooks on instructional methods, teachers' journals, or instructional materials that do not mention and utilize it. While a variety of different ways of operationalizing cooperative learning have been implemented in schools and colleges, there has been no comprehensive review of the research evidence validating the cooperative learning methods. The purpose of this review, therefore, is to examine the empirical support validating the effectiveness of the different methods of cooperative learning. In order to do so, it is first helpful to discuss why cooperative learning is so widely used.

The widespread use of cooperative learning is due to multiple factors. Three of the most important are that cooperative learning is clearly based on theory, validated by research, and operationalized into clear procedures educators can use. First, cooperative learning is based solidly on a variety of theories in anthropology (Mead, 1936) , sociology (Coleman, 1961), economics (Von Mises, 1949), political science (Smith, 1759), psychology, and other social sciences. In psychology, where cooperation has received the most intense study, cooperative learning has its roots in social interdependence (Deutsch, 1949, 1962; Johnson & Johnson, 1989), cognitive-developmental (Johnson & Johnson, 1979; Piaget, 1950; Vygotsky, 1978), and behavioral learning theories (Bandura, 1977; Skinner, 1968). It is rare that an instructional procedure is central to such a wide range of social science theories.

Second, the amount, generalizability, breath, and applicability of the research on cooperative, competitive, and individualistic efforts provides considerable validation of the use of cooperative learning, perhaps more than most other instructional methods (Cohen, 1994a; Johnson, 1970; Johnson & Johnson, 1974, 1978, 1989, 1999a; Kohn, 1992; Sharan, 1980; Slavin, 1977, 1991). There are over 900 research studies validating the effectiveness of cooperative over competitive and individualistic efforts. This body of research has considerable generalizability since the research has been conducted by many different researchers with markedly different orientations working in different settings and countries and in eleven different decades, since research participants have varied widely as to cultural background, economic class, age, and gender, and since a wide variety of research tasks and measures of the dependent variables have been used.

The research on cooperative efforts, furthermore, has unusual breath, that is, it has focused on a wide variety of diverse outcomes. Over the past 100 years researchers have focused on such diverse outcomes as achievement, higher-level reasoning, retention, time on task, transfer of

learning, achievement motivation, intrinsic motivation, continuing motivation, social and cognitive development, moral reasoning, perspective-taking, interpersonal attraction, social support, friendships, reduction of stereotypes and prejudice, valuing differences, psychological health, self-esteem, social competencies, internalization of values, the quality of the learning environment, and many other outcomes. There may be no other instructional strategy that simultaneously achieves such diverse outcomes.

The diverse and positive outcomes that simultaneously result from cooperative efforts have sparked numerous research studies on cooperative learning focused on preventing and treating a wide variety of social problems such as diversity (racism, sexism, inclusion of handicapped), antisocial behavior (delinquency, drug abuse, bullying, violence, incivility), lack of prosocial values and egocentrism, alienation and loneliness, psychological pathology, low self-esteem, and many more (see reviews by Cohen, 1994a; Johnson & Johnson, 1974, 1989, 1999a; Johnson, Johnson, & Maruyama, 1983; Kohn, 1992; Sharan, 1980; Slavin, 1991). For preventing and alleviating many of the social problems related to children, adolescents, and young adults, cooperative learning is the instructional method of choice.

The third factor contributing to the widespread use of cooperative learning is the variety of cooperative learning methods available for teacher use, ranging from very concrete and prescribed to very conceptual and flexible. Cooperative learning is actually a generic term that refers to numerous methods for organizing and conducting classroom instruction. Almost any teacher can find a way to use cooperative learning that is congruent with his or her philosophies and practices. So many teachers use cooperative learning in so many different ways that the operationalizations cannot all be listed here. In assessing the effectiveness of specific cooperative learning methods, however, there are a number of "researcher-developers" who have developed cooperative learning procedures, conducted programs of research and evaluation of their method, and then involved themselves in teacher-training programs that are commonly credited as the creators of modern-day cooperative learning. The following ten have received the most attention (see Table 1): Complex Instruction (CI) (Cohen, 1994b), Constructive Controversy (CC) (Johnson & Johnson, 1979), Cooperative Integrated Reading and Composition (CIRC) (Stevens, Madden, Slavin, & Farnish, 1987), Cooperative Structures (CS) (Kagan, 1985), Group Investigation (GI) (Sharan & Sharan, 1976, 1992), Jigsaw (Aronson, et al., 1978), Learning Together (LT) (Johnson & Johnson, 1975/1999), Student Teams Achievement Divisions (STAD) (Slavin, 1978), Teams-Games-Tournaments (TGT) (DeVries & Edwards, 1974), and Team Assisted Individualization (TAI) (Slavin, Leavey, & Madden, 1982).

Table 1: Modern Methods Of Cooperative Learning

Researcher-Developer	Date	Method
Johnson & Johnson	Mid 1960s	Learning Together & Alone
DeVries & Edwards	Early 1970s	Teams-Games-Tournaments (TGT)
Sharan & Sharan	Mid 1970s	Group Investigation
Johnson & Johnson	Mid 1970s	Constructive Controversy

Aronson & Associates	Late 1970s	Jigsaw Procedure
Slavin & Associates	Late 1970s	Student Teams Achievement Divisions (STAD)
Cohen	Early 1980s	Complex Instruction
Slavin & Associates	Early 1980s	Team Accelerated Instruction (TAI)
Kagan	Mid 1980s	Cooperative Learning Structures
Stevens, Slavin, & Associates	Late 1980s	Cooperative Integrated Reading & Composition (CIRC)

This combination of theory, research, and practice makes cooperative learning a powerful learning procedure. Knowing that cooperative learning can have powerful effects when properly implemented does not mean, however, that all operationalizations of cooperative learning will be effective or equally effective in maximizing achievement. While many different cooperative learning methods are being advocated and used, educators have very little guidance as to which specific cooperative learning methods will be most effective in their situation. The purpose of this review, therefore, is to examine the empirical support validating the effectiveness of the different methods of cooperative learning in maximizing achievement. More specifically, four issues will be investigated.

The first issue is how much research has been conducted to validate specific cooperative learning procedures. While the voluminous research on cooperation has been summarized in a various books and articles (Cohen, 1994a; Johnson & Johnson, 1989; Sharan, 1980; Slavin, 1977), the majority of the research studies on cooperation do not directly test the effectiveness of specific cooperative learning procedures. Many of the research studies that have been conducted may be classified as efficacy studies (i.e., laboratory studies of short-term effects) as opposed to effectiveness studies (i.e., real-world studies of how cooperative learning is actually delivered and what the outcomes are like). Effectiveness studies can be divided into studies aimed at testing theory as well as the effectiveness of a cooperative learning method and curriculum evaluation case studies that have little theoretical relevance but demonstrate that a cooperative learning method worked in a specific situation. The two types of studies complement each other. While a number of people have reviewed the research supporting their cooperative learning methods (e.g., Cohen & Lotan, 1997; Sharan & Sharan, 1992, Slavin, 1991), there has never been a comprehensive and complete review of the effectiveness studies on all the different cooperative learning methods. It is unknown, therefore, how much of the existing research specifically focuses on cooperative learning methods and achievement.

The second issue investigated is how many different cooperative learning methods have been evaluated. As noted earlier, cooperative learning is a generic term referring to numerous methods for organizing and conducting classroom learning. It is used in many different variations, most of which have never been evaluated. There has never been a comprehensive assessment of how many of cooperative learning methods have been empirically tested. The methods most frequently referred to in the research and educational methods literatures are listed in Table 1.

The third issue investigated is how effective are the different cooperative learning methods in maximizing achievement. Once it is known how much research has been conducted on how many of the cooperative learning methods, the next issue is the strength of the empirical support for each method. In order to determine the size of the effect of each cooperative learning method on student achievement, a meta-analysis must be conducted. Meta-analysis is a method of statistically combining the results of a set of independent studies that test the same hypothesis and using inferential statistics to draw conclusions about the overall result of the studies (Cohen, 1987; Cooper, 1989). The meta-analysis process basically consists of a literature search and the calculation of effect sizes.

The fourth issue investigated is what are the characteristics of the more effective cooperative learning methods. Methods of cooperative learning may be placed on a continuum from direct to conceptual. More direct cooperative learning methods consist of very specific and well-defined techniques that teachers can learn in a few minutes and apply immediately. Teachers are trained to use direct procedures in a lock-step way that is the same in all situations. More conceptual cooperative learning methods consist of conceptual frameworks teachers learn and use as a template to restructure current lessons and activities into cooperative ones. Teachers are trained to create cooperative lessons to fit their specific circumstances. Direct methods may initially be more appealing and seem more user friendly, while conceptual methods (once they are mastered) may be integrated into teachers' teaching repertoires and used throughout their career (Antil, et al., 1998; Berman, 1980; Berman & McLaughlin, 1976; Fullan, 1981; Griffin & Barnes, 1984; Johnson, 1970, 1979; Johnson, Druckman, & Dansereau, 1994; Johnson & Johnson, 1994a, 1994b; Joyce & Showers, 1980, 1982; Smith & Keith, 1982). More specifically, more direct methods tend to be easy to learn (and require less training time), tend to be easily implemented, are often focused on specific subject areas and grade levels (i.e., nonrobust), are easy to discontinue as interest wanes, and are not easily adapted to changing conditions. More conceptual methods tend to be difficult to learn and use initially, may be used in lessons in any subject area for any age student (i.e., robust), and become internalized and routinely used and thus difficult to discontinue, and are highly adaptable to changing conditions. While the considerable research on direct and conceptual innovations documents their strengths and weaknesses affecting implementation and institutionalization, there is almost no research on the important issue of the relative impact of direct and conceptual innovations on achievement and productivity. In this review, therefore, methods of cooperative learning will be classified on a direct-conceptual continuum and correlated with the size of each method's effect on student individual achievement.

Methods

Literature Search

The studies included in this meta-analysis were identified through a thorough search for relevant published and unpublished studies. Methods included conducting computer searches (Educational Resources Information Center [ERIC], Psychological Abstracts [PA], Dissertation Abstracts International [DAI], and the Social Sciences Citation Index [SSCI]), examining relevant bibliographies, searching reference sections of the studies included in the meta-analysis to identify further relevant studies, and contacting relevant researchers and

organizations. We also examined the bibliographies of previous relevant meta-analyses. Over 900 studies on social interdependence were located. The criterion for inclusion in the meta-analysis was that the study evaluated the impact of a specific method of cooperative learning on student achievement. A total of 164 studies met the criteria. Since some reports contained multiple studies, the total number of reports was 158. Since studies that compared multiple cooperative learning methods or had more than one control condition are listed more than once, the tables present 194 separate comparisons of cooperative learning and control methods.

Independent Variables

The first independent variable is method of cooperative learning. Method of cooperative learning was defined by the author(s) of each article. If the author stated that the method used was STAD or Jigsaw it was noted as such. In addition, the operationalization of the method had to include positive interdependence. Examples are positive goal interdependence (mutual goals), positive reward interdependence (joint rewards), resource interdependence (each group member has different resources that must be combined to complete the assignment), and role interdependence (each group member is assigned a specific role). Studies that included intergroup competition as part of operationalizing cooperation were included among the cooperative conditions.

Cooperative learning is compared with competitive or individualistic learning. Competition was operationally defined as the presence of negative goal or reward interdependence. Participants worked alone or with a minimum of interaction and rewards were given on a norm-referenced basis or by ranking participants from best to worst. All studies in this analysis focused on competition among group members, not competition between groups. Individualistic efforts were operationally defined as the lack of social interdependence between participants. Participants worked alone or with a minimum of interaction and rewards were given according to set criteria so there was little opportunity for social comparison. When the control condition was labeled as traditional instruction, the condition was coded as either competitive or individualistic depending on the description of the condition.

The second independent variable was the classification of cooperative learning methods on a continuum of direct to conceptual. More direct cooperative learning methods consist of well-defined procedures that teachers are supposed to follow in an exact, lock-step way while more conceptual cooperative learning methods consist of conceptual frameworks teachers use as a template to overlay lessons and activities they structure to fit their specific circumstances. Each cooperative learning method was rated by two psychology professors on five criteria, each of which was defined as a five-point scale. The ratings on the scales were added together to get a total score. Ease of learning (how quickly the method can be learned) was rated on a five-point scale from "a simple procedure easy to understand and remember" to "a conceptual system difficult to understand and apply." Ease of initial use (the effort required to implement the method initially) was rated on a five-point scale from "a simple procedure easy to do perfectly the first time" to "general, conceptual guidelines that are applied to specific lessons and activities." Ease of maintaining its use over time (once implemented, how difficult it is to discontinue) was rated on a five-point scale from "a procedure that is not integrated into basic

teaching patterns" to "conceptual framework that is integrated into basic teaching patterns." Robustness (applicable to specific subject area and age level) was rated on a five-point scale from "aimed specifically at a subject area and grade levels" to "can be applied to any subject area and grade level." Adaptability (how difficult it is to modify cooperative learning to ensure its effectiveness in changing conditions) was rated on a five-point scale from "lock-step, specific procedures that have to be done the same way every time" to "conceptual system that can be modified and changed to meet changing conditions." A direct cooperative learning method, for example, may be easy to learn, easy to use initially, can be performed without integrating framework into basic teaching patterns, aimed at a specific subject area and grade level, and difficult to adapt to changing conditions. Conceptual method, on the other hand, may be hard to learn, difficult to implement initially, integrated into basic teaching patterns and thus maintained long-term, applicable to all subject areas and grade levels, and easy to adapt to changing conditions.

The method of cooperative learning, the control conditions, and the direct-conceptual nature of the cooperative learning methods were coded by two or more analysts, all psychology professors, with extensive experience coding and analyzing research on social interdependence. Interrater reliability was calculated using the kappa coefficient (Cohen, 1960). The interrater reliability kappa was 0.82. The occasional differences in coding were discussed and resolved through consensus.

Dependent Variable

The dependent variable was student achievement. Achievement was defined as an outcome measure for some type of performance (standardized and teacher-made tests, grades, quality of performances such as compositions and presentations, quality of products such as reports, and so forth). A variety of experimental settings and tasks were used in the studies yielding effect sizes for the dependent variable of achievement.

Effect Size

The statistical methods and terminology for meta-analysis are from Cohen (1987), Hedges and Olkin (1985), Cooper (1989), Hunter, Schmidt, and Jackson (1982), and Glass, McGaw, and Smith (1981). The effect size *d* was the difference between treatment divided by the pooled standard deviation of the two groups (Cohen, 1989). When means were not given, but significance task results were, the *F, T*, or X^2 was converted to *d* (Cooper, 1989). All effect sizes were adjusted to control for small sample bias (Hedges & Olkin, 1985). Within studies where there were multiple achievement measures, the average effect size was found by averaging the multiple measures to derive one effect size for each treatment contrast. The mean weighted effect size was found by multiplying each independent effect size by the inverse of its variance and then the sum of these products was divided by the sum of the inverses. The resulting weighted mean effect size is referred to as "d+." Confidence intervals (95 percent) were calculated to determine the statistical significance of each weighted mean effect size (Cooper, 1989). Tests for homogeneity of variance (Q_w) of effect sizes were calculated .

Results

Characteristics Of The Studies

A total of 158 studies on specific cooperative learning studies met the criteria for inclusion in this meta-analysis. The characteristics of the studies are found in Table 2. All studies have been conducted since 1970 with 28 percent conducted since 1990. Thirty percent did not randomly assign participants to conditions, 45 percent randomly assigned participants to conditions, and 25 percent randomly assigned groups to conditions. Forty-six percent were conducted in elementary schools, 20 percent were conducted in middle schools, 11 percent were conducted in high schools, and 24 percent were conducted in post-secondary and adult settings. Sixty-six percent of the studies were published in journals. Fifty-two percent lasted for 2 to 29 sessions (a session was defined as 60 minutes or less), and 46 percent lasted for 30 sessions or more. Ninety-four percent of the studies involved mixed gender groups. Four studies were conducted in Southeast Asia, 3 studies were conducted in the Middle East, 3 studies were conducted in Europe, four studies were conducted in Africa, and several of the North American studies contained minority group students.

Table 2: General Characteristics Of Studies Of Cooperative Learning Methods

Characteristic	Number	Percent
1970 – 1979	26	16
1980 – 1989	88	56
1990 – 1999	44	28
No Random Assignment	48	30
Randomly Assigned Subjects	71	45
Randomly Assigned Groups, Subject Unit of Analysis	30	19
Randomly Assigned Groups, Group Unit of Analysis	9	6
Primary (K - 3)	22	14
Intermediate (4 - 6)	43	27
Primary & Intermediate	8	5
Middle School (7 - 9)	32	20
High School (10 - 12)	17	11
Post Secondary	33	21
Adult	4	3
Journal Article	105	66
Book	2	1
Theses	28	18

Technical Report	17	11
Unpublished	6	4
1 Session	3	2
2 - 9 Sessions	38	24
10 - 29 Sessions	45	28
30+ Sessions	72	46
Same Gender Groups	10	6
Mixed Gender Groups	148	94
Total	158	100

Total Number of Reports = 158; Total Number of Studies = 164 (some reports gave results for multiple studies)

Table 3: Meta-Analysis Results For Cooperative Learning Methods

	Average Effect Sizes			Weighted Effect Sizes						
Learning Together	**Effect**	**Sd**	**K**	**Effect**	**SE**	**k**	**Cld 95%**	**fsn**	**Qw**	**pvalue**	**df**
Cooperation vs. Competition	0.82	0.50	25	0.70	0.06	25	±0.12	62	54.23	0.00	24
Cooperative vs. Individual.	1.03	0.69	56	0.91	0.04	56	±0.08	200	188.66	0.00	55
Competitive vs. Individualistic	0.06	0.47	10	0.08	0.10	10	±0.19	0	15.76	0.07	9
TGT											
Intergroup Comp vs. Competition	0.48	0.69	9	0.47	0.05	9	±0.10	12	141.30	0.00	8
Intergroup Comp vs. Individualistic	0.58	0.43	5	0.55	0.11	5	±0.22	9	10.60	0.03	4
Group Investigation											
Cooperation vs. Competition	0.37	1.19	2	0.86	0.14	2	±0.27	7	24.73	0.00	1
Cooperation vs. Individual.	0.62		1	0.62	0.44	1	±0.86	2	0.00		0
Academic Controversy											
Cooperative vs. Competition	0.59	0.44	16	0.61	0.07	16	±0.14	32	36.82	0.00	15
Cooperative vs. Individual.	0.91	0.59	11	0.86	0.10	11	±0.19	36	22.08	0.01	10
Jigsaw											
Cooperation vs. Competition	0.29	0.78	9	0.41	0.05	9	±0.11	9	68.46	0.00	8
Cooperation vs. Individual.	0.13	0.29	5	0.09	0.11	5	±0.21	0	4.86	0.30	4
STAD	**Effect**	**Sd**	**K**	**Effect**	**SE**	**k**	**Cld 95%**	**fsn**	**Qw**	**pvalue**	**Df**
Intergroup Comp vs. Competition	0.51	0.72	15	0.46	0.05	15	±0.09	19	205.32	0.00	14
Intergroup Comp vs. Individ.	0.29	0.71	14	0.28	0.07	14	±0.14	6	53.89	0.00	13
TAI											
Cooperative vs. Competitive	0.25	0.14	7	0.19	0.04	7	±0.09	0	4.92	0.55	6
Cooperative vs. Individual.	0.33	0.26	8	0.19	0.06	8	±0.12	0	11.57	0.12	7
Competitive vs. Individual.	-0.08	0.52	2	-0.32	0.13	2	±0.25	0	5.00	0.03	1
CIRC											
Cooperation vs. Competition	0.18	0.23	7	0.20	0.04	7	±0.07	0	13.43	0.04	6
Cooperation vs. Individual.	0.18	0.00	1	0.18	0.00	1	±0.22	0	0.00		0

Note: sd = standard deviation; k = the number of averaged effect sizes in the meta-analysis; SE = standard error; Cld 95% = the value of the 95% confidence interval around the weighted effect size; fsn = fail safe N (the number of additional studies needed to change the significance of the results below 0.20).

Results For Different Cooperative Learning Methods

The results for the different methods of cooperative learning appear in Table 3. While both the averaged and the weighted effect sizes appear in Table 3, only the averaged effect sizes are discussed here. For Learning Together, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.82 and 1.03 respectively). There tends to be no meaningful difference between competitive and individualistic efforts, effect size = 0.06. For Teams-Games-Tournaments (TGT), cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.48 and 0.58 respectively). For Group Investigation, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.37 and 0.62 respectively). For Academic Controversy, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.59 and 0.91 respectively). For Jigsaw, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.29 and 0.13 respectively). For Student-Teams-Achievement-Divisions, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.51 and 0.29 respectively). For Team Assisted Individualization, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.25 and 0.33 respectively). There tends to be no meaningful difference between competitive and individualistic efforts, effect size = -0.08. Finally, for Cooperative Integrated Reading and Composition, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.18 and 0.18 respectively). The results for the weighted effect sizes were very similar, somewhat lower for Learning Together and TAI and somewhat higher for Group Investigation and Jigsaw. No studies were found for Cooperative Learning Structures. There are studies of Complex Instruction (see Cohen & Lotan, 1997), but the studies compared complex instruction with other group instructional procedures (rather than comparing them with competitive or individualistic instruction) and, therefore, relevant effect-sizes could not be derived.

Table 4: Ranking Of Cooperative Learning Methods

Method	Coop v Comp	n	Method	Coop v Ind	N
LT	0.85	26	LT	1.04	57
AC	0.67	19	AC	0.91	11
STAD	0.51	15	GI	0.62	1
TGT	0.48	9	TGT	0.58	5
GI	0.37	2	TAI	0.33	8
Jigsaw	0.29	9	STAD	0.29	14
TAI	0.25	7	CIRC	0.18	1
CIRC	0.18	7	Jigsaw	0.13	5

The cooperative learning methods may be ranked by the size of the effect they have on achievement and by the number of comparisons available (see Table 4). When the impact of cooperative lessons is compared with competitive learning, Learning Together promotes the

greatest effect, followed by Constructive Controversy, STAD, TGT, Group Investigation, Jigsaw, TAI, and finally CIRC. When the impact of cooperative lessons is compared with individualistic learning, Learning Together promotes the greatest effect, followed by Constructive Controversy, Group Investigation, TGT, TAI, STAD, Jigsaw, and CIRC. There are reasons, however, why these rankings should be suggestive only. The few number of studies conducted on several of the methods makes the effect sizes very tentative. In addition, different measures of achievement were used in the different studies. The confidence educators can have in the effect sizes, furthermore, is inversely related to the number of studies that have been conducted on the method. When the methods are ranked by the number of effects associated with each findings, for the cooperative versus competitive comparison, the ranking of the methods is Learning Together, Academic Controversy, STAD, Jigsaw, TGT, TAI, CIRC, and Group Investigation. For the cooperative versus individualistic comparison, the ranking is Learning Together, Academic Controversy, STAD, TAI, TGT, Group Investigation, and CIRC.

Table 5: Rating Of Direct-Conceptual Nature Of Cooperative Learning Methods

Method	Learn	Initial Use	Maintain	Robust	Adaptability	Total
Learning Together	5	5	5	5	5	25
TGT	3	3	1	2	2	11
Group Investigation	5	5	3	2	2	17
Academic Controversy	5	5	5	4	4	23
Jigsaw	2	2	3	3	3	13
STAD	2	2	1	2	2	9
TAI	2	2	1	1	1	7
CIRC	2	2	1	1	1	7
Complex Instruction	5	5	3	3	3	19
Cooperative Structures	1	1	1	1	5	9

Note: Methods of cooperative learning were evaluated on five dimensions: (a) ease of learning the method, (b) ease of initial use in the classroom, (c) ease of long-term maintenance of use of the method, (d) robustness of the method (applicability to a wide variety of subject areas and grade levels), and (e) ease of method's adapting to changing conditions.

There are five dimensions on which the methods of cooperative learning may be evaluated (see Table 5): (a) ease of learning the method, (b) ease of initial use in the classroom, (c) ease of long-term maintenance of use of the method, and (d) robustness of the method (applicability to a wide variety of subject areas and grade levels), and (e) adaptability of method to changing conditions. Each cooperative learning method may be classified on a five-point scale (easy-moderate-difficult) on these dimensions. When the resulting score is

correlated with the effect-sizes for each method, the results indicate that the more conceptual the cooperative learning method, the higher the achievement of cooperative compared with competitive, $r(197) = 0.32$, $p<0.001$, and individualistic learning, $r(197) = 0.46$, $p<0.001$.

File Drawer Problem

A potential source of bias in reviewing a set of studies may be that only studies that tend to find significant differences are published and available for review. There may be numerous unpublished works that might change the overall findings. Orwin (1983) presented a procedure for determining how many studies would have to be unpublished to change the results found. He makes an assumption that the effect sizes from unretrieved findings are equal to zero (which is very conservative). His statistic then determines how many studies in file drawers with an average effect size of zero would be needed to shift the obtained weighted mean effect size to a criterion level such as 0.20 (which is small as defined by Cohen [1987, p. 25-26]). The results from Table 3 indicate that for TAI, CIRC, the cooperative vs. individualistic comparison for Jigsaw, the competitive vs. individualistic comparison for LT, and the cooperative vs. individualistic comparison for GI, even one more study could significantly change the results. For the other cooperative learning methods, it would take from 2 to 206 additional studies to change the results significantly.

Discussion

Cooperative learning has been around a long time (Johnson, 1970; Johnson & Johnson, 1989, 1999). It will probably never go away due to its rich history of theory, research, and actual use in the classroom. Markedly different theoretical perspectives (social interdependence, cognitive-developmental, and behavioral learning) provide a clear rationale as to why cooperative efforts are essential for maximizing learning and ensuring healthy cognitive and social development as well as many other important instructional outcomes. Hundreds of research studies demonstrate that cooperative efforts result in higher individual achievement than do competitive or individualistic efforts. Educators use cooperative learning throughout North America, Europe, and many other parts of the world. This combination of theory, research, and practice makes cooperative learning one of the most distinguished of all instructional practices.

Knowing that cooperative learning can significantly increase student achievement (compared with competitive and individualistic learning) when properly implemented does not mean, however, that all operationalizations of cooperative learning will be effective or that all operationalizations will be equally effective. Without reviewing the research on the different cooperative learning methods, it is difficult to recommend specific cooperative learning procedures to educators. This meta-analysis, therefore, focuses on four issues: (a) determining how much research has been conducted on cooperative learning methods, (b) determining how many different cooperative learning methods have been evaluated, (c) determining how effective each method evaluated is in maximizing student achievement, and (d) determining the characteristics of the more effective cooperative learning methods.

The first issue was to determine the amount of research that has been conducted on cooperative learning methods. One-hundred-sixty-four studies on specific cooperative learning methods were found. This is a substantial number of studies, especially considering that 28 percent of them have been conducted since 1990 and 100 percent have been conducted since 1970. The studies have been conducted at all levels of schooling (46 percent were conducted in elementary schools, 20 percent were conducted in middle schools, 11 percent were conducted in high schools, and 24 percent were conducted in post-secondary and adult settings) and the majority lasted for considerable time (46 percent lasted for 30 sessions or more, 52 percent lasted for 2 to 29 sessions, and 2 percent of the studies lasted only for one session). Most of the studies used good to excellent methodology (45 percent randomly assigned participants to conditions, 25 percent randomly assigned groups to conditions, and only 30 percent did not randomly assign participants or groups to conditions). The research has been conducted in North America, Europe, the Middle East, Asia, and Africa and has involved minority as well as majority populations. Thus, there is considerable research on specific cooperative learning methods and the research has considerable validity and generalizability. As with the overall research, educators can have a great deal of confidence in the effectiveness of cooperative learning.

The second issue investigated was to determine how many different cooperative learning methods have been evaluated. Of all the numerous ways that cooperative learning is used, only eight methods have been subjected to empirical validation in a way that a relevant effect size could be computed. Of these methods, some have more empirical support than others. The more research studies conducted on any method, the more valid and reliable the results can be expected to be. There are 113 independent effects in the studies on Learning Together and Constructive Controversy, 66 independent effects in the studies on the cooperative learning methods developed at Johns Hopkins University, 12 independent effects in the studies on the Jigsaw Procedure, and 3 independent effects in the studies on the Group Investigation Method. It is somewhat surprising that so few methods have been evaluated. While any teacher may develop a version of cooperative learning that is very effective, without research studies it is unknown whether other teachers can expect reliable results when the method is used. The unevaluated cooperative learning methods, therefore, should be used with some caution. In addition, there is a need for a new generation of researcher-developers who formulate new operationalizations of cooperation for classroom and school use and who subject their formulations to rigorous empirical evaluation.

The third issue investigated is the effectiveness of the different cooperative learning methods researched. There is no reason to expect that all operationalizations of cooperation will be effective. While the largest effect sizes were found for the Learning Together, Constructive Controversy, Teams-Games-Tournaments, and Group Investigation Methods, all of the methods have substantial effect sizes and all of the methods have been found to produce significantly higher achievement than did competitive or individualistic learning. Any teacher should feel quite comfortable using any of these eight cooperative learning methods.

The diversity of the eight cooperative learning methods provides additional validation of the effectiveness of cooperative learning. The methods range from specific procedures (such as Jigsaw and CIRC) to conceptual frameworks educators use to build their own cooperative

lessons (such as Learning Together and Group Investigation) to curriculum packages in which cooperative learning is a central part (such as TAI and STAD), to rather complex procedures that require some sophistication to use (such as Constructive Controversy). That all of these methods are effective in increasing achievement is a tribute to the power of cooperation.

The fourth issue investigated was the characteristics of the different cooperative learning methods. Among the researcher-developers of cooperative learning, there are those who believe that the best way to ensure implementation of cooperative learning is to devise very specific techniques that teachers can learn in a few minutes and apply immediately (direct approach) and those who believe that teachers must learn a conceptual system and use it to adapt current lessons and activities into cooperative ones (conceptual approach). Previous research indicates that direct methods may be easier to learn and implement than are conceptual methods, but once implemented, conceptual methods are more robust and are more frequently maintained over time and easier to adapt to changing conditions and circumstances (Antil, et al., 1998; Berman, 1980; Berman & McLaughlin, 1976; Fullan, 1981; Griffin & Barnes, 1984; Johnson, 1970, 1979; Johnson, Druckman, & Dansereau, 1994; Johnson & Johnson, 1994a, 1994b; Joyce & Showers, 1980, 1982; Smith & Keith, 1982). There is very little research, however, on whether direct or conceptual methods differentially affect achievement and productivity. The results of this meta-analysis indicate that the more conceptual the method of cooperative learning, the greater its impact on student achievement tends to be. This is an important addition to the literature on implementation and institutionalization of innovations. Differences in the way achievement was measured, however, make these findings tentative. Further research is needed to corroborate this finding.

It seems reasonable to hypothesize that the effectiveness of a cooperative learning method will tend to increase the more that cooperation is the foundation on which classroom and school life is based. If cooperative learning is used within a primarily competitive or individualistic school, for example, its effectiveness may be dampened by the overall culture of the school. Two of the cooperative learning methods have been extended to the overall organizational structure of the school. The Learning Together method has been adapted to include faculty interactions as well as student interactions and is known as the Cooperative School (Johnson & Johnson, 1994). School leaders are trained to implement a cooperative structure in colleagial teaching teams, faculty study groups, task forces, site-based decision making, and cooperative faculty meetings. The procedures have been used in elementary, middle, secondary schools and institutions of higher education. The Johns Hopkins cooperative learning methods have been extended into a schoolwide program for elementary schools known as Success For All (Slavin, et al., 1996). The extension of cooperation to the overall school structure is a promising area for future research.

There is no reason to expect the different methods of cooperative learning to be contradictory. All the methods may be used in the same classroom and school. A teacher, for example, may use TAI in math, Learning Together in science and language arts, and Group Investigation in social studies and expect that the different methods will enhance and enrich each other's effectiveness. There is currently, however, no research on the ways in which the different methods of cooperative learning may enhance or interfere with each other's effectiveness.

The current research findings present a promise that if cooperative learning is implemented effectively, the likelihood of positive results is quite high. Results, however, are not guaranteed. The results of this meta-analysis provide evidence that considerable research has been conducted on cooperative learning methods, that eight diverse methods have been researched, all methods have produced higher achievement than competitive and individualistic learning, and the more conceptual approaches to cooperative learning may produce higher achievement than the direct methods. These conclusions are all the stronger due to the diversity of the research on which they are based, ranging from controlled field experimental studies to evaluational case studies.

Despite the amount and diversity of the research, several conclusions about the effectiveness of the cooperative learning methods may be made. First, while future research is needed, conducting research to compare directly the effectiveness of different cooperative learning methods is not very helpful. Studies in which two or more methods of cooperative learning are directly compared are difficult to interpret, especially if they are conducted by a researcher-developer who has a vested interest in one of the methods. It is virtually impossible to implement different methods at exactly the same strength. If one method is strongly implemented and another method is weakly implemented the resulting differences would be due to the strength of the implementation, not the differences between the methods.

Second, the differences in effect sizes for the different cooperative learning methods should be interpreted cautiously. The measures of academic achievement in various studies may not be equivalent. Lower effect sizes, for example, would be expected on standardized tests than on nonstandardized tests. Methods of cooperative learning aimed at lower-level tasks may produce high effect sizes on simple recognition level tests than methods of cooperative learning aimed at higher-level reasoning and critical thinking. Thus, a lower effect size may be due to the type of measure of academic achievement or the match between the method and the dependent measure, not the overall effectiveness of the method.

Third, more research is needed on the various methods. The more studies conducted on a method, the more accurate the effect size may be. Conclusions about methods that have only a few validating studies could be misleading.

Fourth, most of the validating studies on methods of cooperative learning have been conducted by the researcher-developer who originated the method. This introduces potential bias into the results. Ancient Romans advised individuals to ask, "cui bono" (who benefits) and the researcher-developer often has interests at stake that may bias his or her results toward confirming the effectiveness of his or her program. More studies conducted by independent investigators are needed.

Finally, many of the studies conducted on the impact of cooperative learning methods on achievement have methodological shortcomings and, therefore, any differences found could be the result of methodological flaws rather than the cooperative learning method. In the future, researchers should concentrate on conducting highly controlled studies that add to the confidence with which their conclusions will be received.

References – Available Upon Request

Appendix A: Complete List Of Studies Used In Meta-Analysis – Available upon Request