

Cooperative learning in secondary mathematics: A quantitative review

Harsh Suri
La Trobe University

This paper is a quantitative review of research on cooperative learning in secondary mathematics. Results from the analysis demonstrate that cooperative learning has an overall positive effect in the cognitive domain as well as the social and affective domain. The paper also identifies several potential moderators of the effects and makes implications for further research reviews in the field.

Introduction

During the past few decades there has been a growing interest among researchers in cooperative learning and it is being accepted as an important teaching methodology (Cohen, 1994; Lou et al., 1995). International associations, such as The International Association for the Study of Cooperation in Education (IASCE) have been formed to provide a forum for researchers and educationists to share ideas about cooperation in education (Slavin, 1985).

The use of cooperative learning in mathematics classrooms is a related area which has also attracted considerable interest (Davidson & Kroll, 1991). Curriculum developers and educational agencies such as Board of Studies and National Council of Teachers of Mathematics (NCTM) explicitly recommend the use of cooperative learning strategies in secondary mathematics (Board of Studies, 1995; NCTM, 1991).

Despite an overwhelming acceptance of cooperative learning among researchers and educational organisations, the strategy is not used frequently in schools (Bossert, 1988; Stebler & Reusser, 1996). In addition, a closer look at the literature shows that cooperative learning is more often used in subject areas like English than in mathematics. To some extent this discrepancy can be attributed to teachers' beliefs about cooperative learning; often teachers are unsure of the cognitive and metacognitive productivity of cooperative learning even if they believe in its productivity in affective and social domains. The awareness that cooperation has positive effects in affective and social domains is not enough to motivate teachers to use this technique in their class rooms. They are more willing to use the strategy if they feel confident that it will promote student achievement and be effective in cognitive domain (Stebler & Reusser, 1996).

The research literature on cooperative learning is not easy to interpret and analyse, as studies conducted in this field have a diversity of aims, subject matter, topics, grade levels and study designs. Several attempts have been made to synthesise this overwhelming volume of literature (e.g. Bossert, 1988; Johnson, Maruyama, Johnson, Nelson & Skon, 1981; Slavin, 1990). However, often these research reviews have been very broad in scope and hence may not have provided teachers with the specific information they need for restructuring their classrooms. A need therefore arises to synthesise that subset of literature which may interest secondary mathematics teachers.

This paper reports the results of a pilot study for a comprehensive review of research on cooperative learning in secondary mathematics. It is a synthesis of the quantitative data which could be converted to an effect size (ES). The paper also examines the pros and cons of several methodological approaches.

Method

Operational definition

In this study, cooperative learning is used in a broad sense. It refers to learning that takes place when two or more students work in small groups and it involves some form of resource sharing, reward sharing or both among the group members. The definition includes what is referred to as peer tutoring, collaborative learning and small group work. It encompasses both heterogenous grouping as well as homogenous grouping.

Collecting the empirical studies

In accordance with Slavin's (1986) best-evidence synthesis method, this study used well-justified *a priori* inclusion criteria for empirical studies. It included all the studies which: (a) examined the effects of cooperative learning in mathematics on secondary (grade levels 7-12) students; and (b) were conducted on normal progress students even if from different socio-economic backgrounds or with varying academic potential such as low achievers or high achievers.

The synthesis excluded the empirical studies which: (a) were conducted on special need students such as academically handicapped students or gifted students; (b) were conducted in special educational programs, such as remedial classes, community college, adult education program or university preparatory courses; (c) examined a form of cooperative learning which did not agree with the operational definition; (d) were not reported in English; (e) focussed on other factors and in which cooperative learning was very incidental; (f) focussed only on teachers' attitudes to cooperative learning and reported no measure of students' beliefs or achievement variations; (g) were reported before 1981.

Using this selection criteria, electronic searches were conducted on ERIC (Education Resources Information Center), AUSTRUM (Australian Social Science, Law and Education Database on CD-ROM), PsycLIT and APA (Psychology and related subjects by American Psychological Association), Maths Sci (Mathematics and Mathematically related research by American Mathematical Society), SSI (Social Sciences Index), SSCI (Social Sciences Citation Index). Owing to financial constraints, no attempts were made to find the studies from the DAI (Dissertation Abstract International). Consequently, 29 studies were selected for the analysis.

Data collection and analysis

Each selected study was closely examined to fill in the relevant information in a coding form and to compute the Effect Size (ES). Slavin's (1986) formulae were used to compute the ES with a slight modification; in each denominator, the control group standard deviation was replaced by the pooled standard deviation. The extent to which the effects were pooled within a study or across studies was guided by Slavin's (1986) best-evidence synthesis method. Within each sub category, the mean of the appropriately weighted effect sizes (as recommended by Hedges & Olkins, 1985) rather than the median ES (as recommended by Slavin, 1986) was taken as the pooled ES.

Study outcomes within each category were analysed for homogeneity to determine if a single ES was a good representation of the studies (Hedges & Olkins, 1985; Johnson, 1989). Significant within group inconsistencies were observed in each sub category. Hence, the pooled ES was not taken as the conclusive result. An outlier diagnosis was performed to account for the within group heterogeneity. If the largest outlier had major methodological difference from the other studies, it was eliminated from further analysis.

Outlier diagnosis was followed by categorical model testing. This involved dividing the findings into sub categories on the basis of a study quality and testing these sub categories for within group and between group homogeneity to quantify the extent to which the study quality moderated the effect. In general, DSTAT, a software for meta-analysis, was used for computing and averaging the effects and for carrying out the outlier diagnosis and the categorical model testing (Johnson, 1989).

Results and Discussion

Overall effect of cooperative learning

This pilot study quantitatively synthesised the primary research literature on the effects of cooperative learning in secondary mathematics. Appendix A lists the findings extracted from different studies. The mean sample size weighted ES (\bar{d}) in the cognitive domain was +0.63, a value close to the values reported by Johnson et al. (1981), Slavin (1990) and Lou et al. (1995). Cooperative learning was also found to promote the retention capacity of students (\bar{d} = +0.50). In general, most of the previous broadly based reviews agreed that cooperative learning enhances the student achievement. This

study adds another specific dimension to the sweeping generalisation. It demonstrates that cooperative learning has overall positive effects on achievement in secondary mathematics.

In the social and affective domain, the effect was measured on variables such as learning goal, intrinsic motivation, performance goal, extrinsic motivation, efficacy, social satisfaction, mathematics value and mathematics cost. The mean sample size weighted ES ($d^+ = +0.25$) was significantly positive. This concurs with the findings of Johnson & Johnson (1985), Slavin (1990), Davidson & Kroll (1991) and Lou et al. (1996).

Further, cooperative learning method is also preferred by students. An analysis of findings from the Learning Preference Style for Students (LPSS) administered to Sydney students revealed that in general, students prefer cooperative rather than competitive or individualistic goal structure ($d^+ = +0.29$). This result confirms similar assertions made by Barnes & Owens (1992).

Potential moderators of the effect

The study also investigated potential moderator variables of the ES. The presence or absence of equivalence between experimental and control conditions, duration of the study, form of cooperative learning method used, source of the outcome, and grade level of the students were the qualities which significantly moderated the ES.

Equivalence of experimental - control conditions: The mean weighted effect on achievement for the findings from studies in which both the experimental and control groups were subjected to equivalent conditions ($d^+ = 0.76$) was significantly higher than those for which the conditions were non equivalent ($d^+ = 0.55$), $Q_B(1) = 4.04$, $p < .05$. This is in contrast with the finding of Lou et al. (1995). They found that studies which subjected the experimental and control group students to similar treatments yielded lower average ES than those which differed in their treatments. Hence they inferred that to a large extent the positive effect on achievement could be attributed to the non equivalence of the experimental and the control groups, a claim which is challenged by these results.

To account for this difference in results, the aims, methods and scope of Lou et al. (1995) were examined. A closer inspection revealed that they had studied the effects of "within-class grouping" rather than cooperative learning. They also appear to have included some studies where the small learning groups were not cooperative. In addition, their review was not restricted to secondary mathematics only. These marked differences in the aim and scope of the previous review could have been responsible for the difference in findings.

Duration of the study: In the cognitive domain, the mean weighted ES for the findings from studies which lasted up to 20 weeks ($d^+ = 0.75$) was significantly higher than those which lasted from 21 to 40 weeks ($d^+ = 0.55$), $Q_B(1) = 3.87$, $p < .05$. This is in agreement with Johnson et al. (1981) who concluded that studies of shorter duration produced relatively higher effects on achievement. Perhaps studies of shorter duration show higher ES due to the novelty of the cooperative learning technique. This Hawthorn effect wears out with time in the studies of longer duration. In the social and affective domain, the effect of study duration could not be examined on account of the limited number of findings.

Source of outcome: The source of outcome significantly moderated the effect on achievement, $Q_B(2) = 15.50$, $p < .05$. The mean sample size weighted ES for the findings based on researcher made tests ($d^+ = 0.77$) was significantly higher than those based on teacher made ($d^+ = 0.53$) or standardised ($d^+ = 0.12$) tests. One possible explanation could be that larger researcher bias was introduced in studies which used researcher made testing instruments. Effect of source of outcome in the affective domain could not be studied as all the findings were based on researcher made tests.

Form of cooperative learning: The form of cooperative learning used for the study also significantly moderated the effect on achievement, $Q_B(4) = 19.29$, $p < .05$. The mean sample size weighted ES for the findings using Slavin's Student Team Learning (STL) methods ($d^+ = +0.79$) was significantly higher than those using AGO ($d^+ = 0.+68$), MTC ($d^+ = +0.49$), the structural approach ($d^+ = +0.24$) or 'Groups of four' method ($d^+ = -0.06$). This concurs with the assertion of Newmann and Thompson

(1987) and Slavin (1990) that Student Team Learning methods are more effective than other cooperative learning methods.

Slavin (1990) suggested that his Student Team Learning methods were more effective because they use individual accountability and team rewards. In order to test Slavin's hypothesis, the findings were divided in two categories: those which used team rewards and individual accountability and those which did not. No significant difference was found in the mean sample size weighted ES of the two categories, $Q_B(1) = 1.40$, $p > .05$. This supports the case made in chapter 2 that the higher effects of the STL methods could be attributed to several factors other than team rewards and individual accountability. For example, learning while playing could make learning more interesting. Again, it is hard to say that all the studies in the latter category did not use any trace of team rewards: the teachers might have verbally recognised the group's collective efforts (Bossert, 1988). Also, Slavin's STL methods such as STAD, TAI are older and more commonly used than some of the new methods such as MTC and the scripted cooperative learning. It is possible that both researchers and teachers were more comfortable with using the STL methods.

Grade level: The mean sample size weighted ES for grade 11 students ($d^+ = +0.55$) was significantly higher than that of grade 7 students ($d^+ = +0.11$) on LPSS. This cumulative effect was computed by pooling the 8 findings extracted from a study conducted by Barnes & Owens (1982). The original study specified the direction of the effect but did not report its magnitude. In the present thesis, the primary data was reanalysed to answer the research hypothesis. This in turn added another dimension to the results of the original study. It not only confirmed the direction of the effect but also computed its magnitude.

Other variables: Two other variables, grouping criteria and the number of students in each cooperative group, tended to moderate the effect although not at 95% level of confidence. Cooperative learning tended to be more effective when each cooperative group was composed of heterogenous ability groups rather friendship groups. Likewise, the effect was higher when four students worked in each group as against dyadic groups. Several other variables which were examined included the year of reporting, publication status, student ability, ethnic breakdown, socio economic background, location, and the topic of instruction. Many of these qualities demonstrated variation in within group effect of different sub categories although the difference was not at 95% significance level. Perhaps more potential moderators could be diagnosed by increasing the number of studies included in the analysis.

Limitation of the Study

The small number of empirical findings included in the analysis appeared to be a limitation in the diagnosis of many potential moderators of the effects. In a meta-analysis, the chances of detecting significant heterogeneity between the sub categories tends to increase with an increase in the size of each sub category (Slavin, 1986). In this analysis, several sub categories showed between group heterogeneity even if not at 95% significance level. It appears that increasing the number of empirical studies could result in the finding of some more mediating variables of the effect.

Implications for further Research

The overall effect of cooperative learning was significantly positive in the cognitive domain as well as the social and affective domain, even though the individual empirical findings were significantly inconsistent. The overall productivity of cooperative learning suggests that the area is worthy of future research. The inconsistencies between individual empirical findings caution against the generalisation that cooperative learning is effective in all learning situations in secondary mathematics. Future research should focus on identifying the specific factors which hinder or enhance the effects.

The study has worthwhile implications for methodological decision making in related future research. It confirms several strengths and limitations of the meta-analytic procedures. First, the meta-analytic procedures bring all study findings to a common metric. This practice gives appropriate weight to each finding by largely eliminating the

effect of subjective interpretation of the original result in an empirical study, thereby reducing the chances of Type I error. For example, Owens & Barnes (1982), in their paper, concluded that female students demonstrated significantly higher preference for cooperative learning style when compared to their male counterparts. In this thesis, a reanalysis of their primary data, confirmed the tendency towards the effect but the effect was found to be insignificant at 95% confidence level, thus challenging the claims made by authors of the original study.

At times, a reviewer finds empirical studies in a related area which contain the data with a potential of addressing the reviewer's research hypothesis, even though they do not directly do so. Meta-analytic procedures enable a reanalysis of the original data to address the reviewer's research hypothesis. For instance, Owens & Barnes (1982) concluded that grade 11 students when compared to grade 7 students demonstrate higher preference for cooperative as well as competitive learning modes. Using meta analytic procedures, the data was reanalysed to address an aim of this study which was to quantify the difference in relative preferences of grade 11 and grade 7 students for the two learning styles.

The meta-analytic procedures described in this study appeared to comprehensively cover most cases of computing and averaging the effects in a variety of empirical findings, with varied study designs and reported statistical data. When the literature consists of a large number of studies, the meta-analytic procedures would appear, on the basis of the analysis carried out for this study, to be sufficiently robust to objectively identify the overall trends. The meta-analytic procedures also provide a method of identifying variables that mediate the overall trends.

In general, the software DSTAT appeared to be sufficiently powerful for computing the effect sizes from a wide range of statistical data. It was also found to be efficient during the analysis stage which involved computation of the standard deviations, the pooled effect sizes and identification of the potential moderators of the effect. It appeared to have no operational problems.

Despite the obvious advantages of the meta-analytic procedures in analysing quantitative data, reviewers on cooperative learning are recommended not to rely solely on these procedures. A significant number of studies in cooperative learning are qualitative studies, case studies on two or three students, or studies on interaction analysis of students working in small groups which fall beyond the scope of meta-analytic procedures. In future reviews, following the principles of best-evidence synthesis, a meta-analysis of quantitative data should be supplemented with a rich literature review of the data not covered in the scope of a meta-analysis.

Conclusions

Cooperative learning has an overall positive effect in the cognitive domain as well as the social and affective domain in secondary mathematics. The presence or absence of equivalence between experimental and control conditions, duration of the study, form of cooperative learning method used, source of the outcome, and grade level of the students were the qualities which significantly moderated the ES. The small number of empirical findings included in the analysis appeared to be a limitation of this thesis.

Results of this study indicate that there is still scope for further research into cooperative learning in secondary mathematics classrooms. On the basis of the analysis carried out in this thesis, meta-analytic procedures are found to have several strengths. They bring all the findings on a common metric which reduces the effect of subjective interpretations of the primary data. Meta-analytic procedures make provision for a reanalysis of the primary data in an empirical study to change the focus of the original study. This, at times, enables a reviewer to include even those studies from related areas that do not directly address the reviewer's research hypothesis.

The meta-analytic procedures described in this study appeared to comprehensively cover most cases of computing and averaging the effects in a variety of empirical findings, with varied study designs and reported statistical data. These procedures appear sufficiently robust to objectively identify the overall trends and the moderating variables, even for a large number of studies. The DSTAT software used in this study appeared to meet most of the computing requirements of a quantitative review. It is recommended that

in future reviews, following the principles of best-evidence synthesis, a meta-analysis of quantitative data should be supplemented with a rich literature review of the data outside the scope of a meta-analysis.

References

- Barnes, J. & Owens, L. (1992). *Learning Preference Scales: Handbook and test master set*. Hawthorn, VIC: The Australian Council for Educational Research.
- Berg, K.F. (1993). *Structured cooperative learning and achievement in a high school mathematics class*. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA. (ED 364408)
- Board of Studies (1995). *Mathematics Curriculum and Standards Framework*. Melbourne, Board of Studies.
- Bossert, S.T. (1988). Cooperative activities in the classroom. In E.Z. Rothkopf (Ed.), *Review of research in education*, 15, 225-250. Washington, DC: American Educational Research Association.
- Cohen, E.G. (1994). Restructuring the Classroom: Conditions for productive small groups. *Review of Educational Research*, 64 (1), 1-35.
- Davidson, N. & Kroll, D.L. (1991). An overview of research on cooperative learning related to mathematics. *Journal for Research in Mathematics Education*, 22 (5), 362-365.
- Duren, P. E. & Cherrington, A. (1992). The effects of cooperative group work versus independent practice on the learning of some problem-solving strategies. *School Science and Mathematics*, 92 (2), 80-83.
- Hedges, L.V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Johnson, B.T. (1989). *DSTAT: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Lawrence Erlbaum.
- Johnson, D.W. & Johnson, R.T. (1985). The internal dynamics of cooperative learning groups. In R.E. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 103-124). NY: Plenum
- Johnson, D.W., Maruyama, G., Johnson, R., Nelson, D. & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structure on achievement: A meta-analysis. *Psychological Bulletin*, 89 (1), 47-62.
- Johnson, L.C. (1984). *The effects of the Groups of Four program on Sstudent achievement*. Paper presented at the Annual Meeting of the Sixty-Ninth American Educational Research Association, Chicago, IL. (ED 254399)
- Leighton, M.S., Slavin, R.E. & Davidson, N. (1989). *Achievement effects of individual, small group and cooperative learning strategies on math problem-solving*. Published by the Centre for Research on Elementary and Middle Schools, Baltimore, MD. (ED 332862)
- Lou, Y., Abrami, P.C., Spence, J.C., Poulsen, C., Chambers, B. & d'Apollnia, S. (1995). *Within-class grouping*. Portions of this paper were presented at the annual meetings of the American Educational Research Association in San Francisco (April, 1995) and New York (April, 1996). Montreal: Concordia University.
- NCTM (1991). *Professional standards for teaching mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Newmann, F.M. & Thompson, J.A. (1987). *Effects of cooperative learning on achievement in secondary schools: A summary of research*. Madison, WI: University of Wisconsin-Madison, National Ceter on Effective Secondary Schools.
- Nichols, J.D. & Hall, N. (1995). *The effects of cooperative learning on student achievement and motivation in a high school geometry class*. Paper presented at the annual meeting of the American Educational Research Association, San Franciso, CA. (ED 387341)
- Nichols, J.D. & Miller, R.B. (1994). Cooperative learning and student motivation, *Contemporary Educational Psychology*, 19 (2), 167-178.

- Owens, L. & Straton, R. (1980). The development of a cooperative, competitive and individualised learning preference scale for students. *British Journal of Educational Psychology*, 1980, 50, 147-161.
- Reid, J. (1992). *The effects of cooperative learning with intergroup competition on the math achievement of seventh grade students*. (ED 355106)
- Slavin, R.E. (1985). Preface. In R.E. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn*. NY: Plenum.
- Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, 15 (9), 5-11.
- Slavin, R.E. (1990). *Cooperative learning: Theory, research and practice*. Needham Heights, US: Allyn and Bacon.
- Stacey, K. (1992). Mathematical problem solving in groups: Are two heads better than one?, *Journal of Mathematical Behavior*, 11 (3), 261-275.
- Stebler, R. & Reusser, K. (1996). *Purpose and benefits of peer collaboration: What teachers believe and students perceive*. Paper presented at the Annual meeting of the American Educational Research Association, NY.
- Terwel, J., Herfs, P.G.P., Mertens, E.H.M. & Perrenet, J. (1994). Co-operative learning and adaptive instruction in a mathematics curriculum. *Journal of Curriculum Studies*, 26 (2), 217-233.
- Thomas, M. & Sherman, L.W. (1986). *Individualistic and cooperative goal structures in high school mathematics achievement*. Paper presented at the Annual Meeting of the Social Science and Mathematics Association, Lexington, KY. (ED 277566)
- Townsend, M.A.R. & Hicks, L. (1995). *Classroom goal structures, social satisfaction and the perceived value of academic tasks*. Paper presented at the Annual Meeting and Exhibit of the American Educational Research Association, San Francisco, CA. (ED 387762)

Appendix A: Samples used for extracting the effect sizes

Finding	G	D	Design	O	d+
Reid (1992)	7	40	Coop Vs Trad	C	+0.49
Duren & Cherrington (1992)	7-8	4	Coop Vs Trad to promote retention Post-test 3 months after instruction	C	+0.50
Johnson (1984A)	8	35	Coop (Experimental 1 gp) Vs Con (M)	C	-0.11
Johnson (1984B)	8	35	Coop (Experimental 1 gp) Vs Con (F)	C	-0.02
Terwel et al. (1994)	7-11	40	Coop Vs Trad	C	+0.68
Berg (1993)	11	8	Coop Vs Trad (Treatment period)	C	+0.24
Stacey (1992)	9	-	Administered a timed test to students in experimental (Coop)and Con (Ind)	C	-0.04
Leighton, Slavin & Davidson (1989A)	7	5	Team practice, a Coop with no reward sharing (T B), Vs Con (T A)	C	+0.84
Leighton, Slavin & Davidson (1989B)	7	5	STL, a Coop method with reward sharing (T C), Vs Con (T A)	C	+0.85
Thomas & Sherman (1986)	10	5	Coop Vs Ind	C	
Nichols & Hall (1995A)	10	9	Coop (T gp 1) Vs Trad (T gp 2) [during first 9 weeks]	C	+0.48
Nichols & Hall (1995B)	10	9	Coop (T gp 2) Vs Trad (T gp 1) [during next 9 weeks]	C	+0.59
Nichols & Hall (1995C)	10	9	Coop (T gp 1) Vs Trad (T gp 2) [during first 9 weeks]	A	+0.48
Nichols & Hall (1995D)	10	9	Coop(T gp 2) Vs Trad (T gp 1) [during next 9 weeks]	A	+0.59
Nichols & Miller (1994A)	11-12	18	Coop Vs Trad	C	+0.45
Nichols & Miller (1994B)	11-12	18	Coop Vs Trad	A	+0.51
Townsend & Hicks (1995A)	8	—	Coop Vs Trad (M)	A	-0.00
Townsend & Hicks (1995B)	7-8	—	Coop Vs Trad (F)	A	+0.24
Owens & Barnes (1982A)	7	—	Coop Vs Con (Comp & Ind) (M from school A only)	A	+0.36
Owens & Barnes (1982B)	7	—	Coop Vs Con (Comp & Ind) (F from school A only)	A	+0.15
Owens & Barnes (1982C)	11	—	Coop Vs Con (Comp & Ind) (M from school A only)	A	+0.61
Owens & Barnes (1982D)	11	—	Coop Vs Con (Comp & Ind) (F from school A only)	A	+0.92
Owens & Barnes (1982E)	7	—	Coop Vs Con (Comp & Ind) (M from school B only)	A	-0.15
Owens & Barnes (1982F)	7	—	Coop Vs Con (Comp & Ind) (F from school B only)	A	+0.20
Owens & Barnes (1982G)	11	—	Coop Vs Con (Comp & Ind) (M from school B only)	A	+0.60
Owens & Barnes (1982H)	11	—	Coop Vs Con (Comp & Ind) (F from school B only)	A	+0.24

Symbols used: A: Affective; C: Cognitive; Comp: Competitive; Con: Control group; Coop: Cooperative learning; D: Duration in weeks; F: Female sample only; G: Grade level; gp: group; Ind: Individualised; M: Male sample only; O: Outcome feature; Trad: Traditional learning; T: Treatment.