

Main article

# The impact of group formation in a cooperative learning environment

Joyce van der Laan Smith <sup>a,1</sup>, Roxanne M. Spindle <sup>b,\*</sup>

<sup>a</sup> Virginia State University, Petersburg, VA 23806, USA

<sup>b</sup> Virginia Commonwealth University, P.O. Box 844000, Richmond, VA 23284-4000, USA

---

## Abstract

Research indicates that cooperative learning (CL) has the potential to increase accounting student satisfaction without impairing the learning of technical material (Lancaster, K., & Strand, C. (2001). Using the team-learning model in a managerial accounting class: An experiment in cooperative learning. *Issues in Accounting Education*, 16(4), 549–567). This study investigates whether instructor-formed heterogeneous groups produce a more effective CL environment than student self-selected groups by measuring individual academic performance and perceptions. Results indicate the presence of a treatment interaction, implying that the best group composition may not be the same for all students. In some circumstances, higher ability students had statistically higher performance in more homogeneous groups. Lower ability students did better (not significant) in heterogeneous groupings. The majority of students gave high ratings to the impact of CL on learning and development of team skills. Students were only mildly interested in increasing the time spent in groups, indicating that they value both traditional teaching methods and CL.

© 2007 Elsevier Ltd. All rights reserved.

*Keywords:* Cooperative education; Accounting education; Group formation; Treatment interaction

---

## 1. Introduction

Cooperative learning (CL) is an instructional environment in which students work in groups with “shared goals and shared responsibility for the learning of other group

---

\* Corresponding author. Tel.: +1 804 828 7113; fax: +1 804 828 8884.

E-mail addresses: [Jsmith@vsu.edu](mailto:Jsmith@vsu.edu) (J. van der Laan Smith), [rspindle@vcu.edu](mailto:rspindle@vcu.edu) (R.M. Spindle).

<sup>1</sup> Tel.: +1 804 524 6995; fax: +1 804 524 5400.

members” (Ravenscroft, Buckless, McCombs, & Zuckerman, 1995, p. 98). The objectives of CL are generally recognized as the improvement of interpersonal skills, content knowledge, and higher-level thinking ability (Ravenscroft, Buckless, & Zuckerman, 1997). Extensive theory, research, and practice support the use of CL in all levels of education (Johnson, Johnson, & Stanne, 2000). Within accounting education, studies find that the learning of technical accounting material is at least as good under various implementations of CL as under traditional methods (Hwang, Lui, & Tong, 2005; Lancaster & Strand, 2001; Ravenscroft, 1997). Accounting studies also indicate that CL<sup>2</sup> has the potential to increase student satisfaction and interactions.

Given the potential benefits and the increasing interest in CL it is important to understand how alternative implementations of CL affect its stated objectives. One alternative implementation that has received very little attention in the education literature is the effectiveness of alternative group formation methods. This lack of research is surprising given that the emphasis in CL is on group interactions and activities.

Within accounting education, it is generally held that effective CL requires heterogeneous groups (Cottell & Mills, 1993; Ravenscroft et al., 1995). A variety of methods are used to form heterogeneous groups. The objective of the group formation method is often not clearly stated or linked to one of the objectives of CL. Since CL is based on small group processes, it is important to consider the effect, if any, of group formation method on CL effectiveness.

This study investigates whether instructor-formed heterogeneous groups produce a more effective CL environment for individual learning than student self-selected groups. This study also investigates the effect of team formation on student perceptions of the overall CL experience as compared to other classroom activities. There is limited accounting research on the effect of group formation on team performance (Clinton & Kohlmeyer, 2005; Collins & Onwuegbuzie, 2000) or individual performance (Swanson, Gross, & Kramer, 1998).

The current study differs from prior research in two primary aspects. First, it focuses on the effect of team formation on individual student performance, as opposed to team performance. Second, it directly measures student perceptions of the effectiveness of CL as opposed to relying on general questions taken from standard end-of-semester evaluations.

The remainder of this paper is organized as follows: prior research and development of hypotheses are presented first, the research method is presented next, followed by the results and analysis, and discussion and conclusion. The final section presents limitations and suggestions for future research.

## **2. Prior research and development of hypotheses**

### *2.1. Prior research*

Accounting educators have used a wide range of methods to form heterogeneous groups. Ravenscroft et al. (1997) report waiting until after the first exam to form formal groups based on scores from that exam; prior to the first exam ad hoc groups are used.

---

<sup>2</sup> See Norman, Rose, and Lehmann (2004) for an overview of CL research within the business and accounting education disciplines.

Lancaster and Strand (2001) report using a learning style questionnaire and “value lines” and “corners”, two methods originally suggested by Kagan (1989). These methods use students’ verbal responses to value related questions to determine the groupings. McConnell and Sasse (1998) assigned students to groups for the purpose of achieving diversity, using learning style, GPA, and work experience. Hite (1996) grouped students by ability, defined as GPA and scores on preceding exams.

While the objective of the group formation method was not explicitly stated in these studies, Cottell and Mills (1993, p. 44) report “most faculty members using cooperative learning form long-term heterogeneous groups composed of high, low, and middle achievers of both genders and various ethnic and cultural backgrounds and ages. This mix increases the likelihood that the group will be able to solve the given problems and helps to build the mutual respect and cross-cultural interpersonal skills so important to the modern workplace.” This statement indicates that the primary objectives for heterogeneous teams within higher education are the development of interpersonal communication skills and team effectiveness that will translate into improved job performance. However, improved individual learning is also a goal of CL which may not require the same group dynamics.

One focus of team performance is team learning. The outcomes of team learning are detection of changes in the environment and customer requirements, improving members’ collective understanding of a situation, and discovering unexpected consequences of previous team actions. Learning behavior in this context includes seeking feedback, sharing information, asking for help, talking about errors, and experimenting (Edmonson, 1999).

Collins and Onwuegbuzie (2000) examined whether high-ability CL groups produced a higher quality group project than lower ability level groups in a college level educational research methodology course. They found that groups with higher ability level students had better group performance than groups of lower ability level students. They also found some evidence that group ability level heterogeneity was related to the group’s performance because the more heterogeneous groups performed better than the more homogeneous groups. The Collins and Onwuegbuzie (2000) study implies that instructors wishing to maximize group performance should maximize group ability heterogeneity.

It is important not to confuse team performance with indications of individual learning. Slavin (1983) notes that, in many cases, the group outperforms individuals even in the absence of cooperative behavior because the most capable group members provide the answer and all members are assigned the group score. Consistent with this finding, Clinton and Kohlmeyer (2005) found that there were no differences in team performance on group quizzes, between self-selected and assigned groups, in two sections of an accounting course.

An alternative motivation for the use of teams is to provide the opportunity for students to help each other learn (peer tutoring). Peer tutoring is one of the primary reasons for the use of heterogeneous groups at the pre-college level (Kagan, 1989). Peer tutoring addresses two of the key objectives of CL related to individual learning, increasing content knowledge and higher-level thinking ability.

Webb (1991) provides a review of nine studies, at the pre-college level, that found that different ability level students may interact differently in alternative group formations. When all ability students are combined, there is some indication that the high and low

ability students form a teacher/student relationship, which increases verbal interactions. However, the medium ability level students in these groups engaged in less group interaction implying that their learning opportunities decreased.

In homogeneous groups, the medium ability level students gave and received more explanations and showed higher achievement than comparable ability students in mixed ability groups. Low ability, homogeneous groups interacted but tended to fail to provide relevant explanations. Finally, verbal interaction declined when high-ability students were formed into homogeneous groups. Two findings with implications for this study may be derived from Webb's results. First, the results indicate that group composition impacts individual learning. Secondly, Webb's results indicate that ability level provides a complex interaction with the peer tutoring aspect of cooperative learning, implying the presence of an aptitude-treatment interaction (ATI).

Swanson et al. (1998) investigate the impact of study group formation methods on individual performance in an undergraduate introductory accounting course. Students could either form their own group, be randomly assigned to a heterogeneous group (method not defined), or work independently. The study groups met outside of class. Swanson et al. concluded that "students in study groups formed voluntarily perform better than those in groups formed by the instructor or those who chose not to be in any study group" (Swanson et al., 1998, p. 10). However, the students in the self-selected groups were the students with higher GPAs.

The findings from prior research indicate that team composition and formation may impact team performance (Clinton & Kohlmeyer, 2005) and individual learning (Swanson et al., 1998; Webb, 1991) differently. However, there is little research linking the impact of alternative team formation methods to individual learning in accounting at the post-secondary level. This study seeks to help fill this void by comparing the impact of instructor-formed heterogeneous groups and student self-selected groups on students' individual academic performance and perceptions when CL group work is formally incorporated within the classroom instruction.

## *2.2. Hypotheses*

The two primary objectives typically provided for using groups within a CL environment are the improvement of content knowledge and interpersonal skills (Ravenscroft et al., 1995). This study investigates how the formation of groups impacts the effectiveness of CL, within the accounting curricula, in terms of individual learning. The first two hypotheses are related to the objective of improvement of content knowledge, specifically as it relates to the performance and perception of the individual student. Hypotheses one and two are:

H1. Instructor-formed heterogeneous groups produce a more effective CL environment for individual learning than student self-selected groups, as measured by individual academic performance.

H2. Students' perceptions of the impact of their CL experience on individual learning differ between the instructor-formed heterogeneous groups and student self-selected groups.

Students' perceptions of CL provide an understanding of their comfort level with the group process and its interaction with other teaching methods. This provides information useful in

understanding the overall impact of CL on the learning environment. In addition, direct feedback from the students can be used to refine and further develop the CL experience. The final hypothesis addresses this perspective:

H3. Students' preferences for the amount of time spent in the classroom on CL activities differ between the students in the instructor-formed heterogeneous groups and the student self-selected groups.

### **3. Research method**

#### *3.1. Experimental design*

This quasi-experiment was conducted in the second semester of a two-semester introduction to accounting course required of all business students at a large urban university located in the Southeast. The class is typically taken at the sophomore level. This experiment, conducted in a field setting with two accounting classes, used a posttest-only design with non-equivalent groups.

Individual academic performance, used in the tests of H1, was measured using individual exam results and total points earned in the course. To improve internal validity, the grade in the prior accounting course (prior grade) was used as a proxy for a pretest measure since it is an operationally similar measurement of academic performance as the posttest measurement. Additionally, since the first accounting class is a prerequisite for the second accounting course, the prior grade provides a current measure of potential ability in an accounting class. Prior studies in accounting education indicate that prior academic performance is related to student performance in introductory accounting courses (for example, see [Doran, Bouillon, & Smith, 1991](#); [Eskew & Faley, 1988](#)).

Students' perceptions of the group experience, used in the tests of H2 and H3 were collected with an anonymous end-of-semester survey scaled using a 5-point Likert design anchored with "strongly disagree" and "strongly agree". Analysis of covariance (ANCOVA) was used to analyze results. Academic performance and students' perceptions were the dependent variables, group formation method the independent variable, and gender and ethnic background (extraneous variables) the covariates.

Both classes were taught in the same semester using the same syllabus and schedule by a professor with 10 years experience. For the five years prior to this experiment, the professor used a partial CL environment using variations of the student team approach employed in this experiment. The instructor explained the format of the individual/group quizzes and the interdependence of the students' grades during the first class. Information on the group work grade component was also provided in the syllabus (see [Appendix 1](#)).

All students completed data sheets (presented in [Appendix 2](#)) during the first class. This information was used to form heterogeneous groups for one of the classes. A coin toss was used to determine which class received this treatment.

The heterogeneous groups were formed using grades for the first semester as the primary sorting criterion. The A and B students were divided among the groups by gender. The C and D students were then assigned to groups based on gender, anticipated major, ethnic background, and hours of employment per week. The additional criteria were used to increase the diversity of the groups and to provide some degree of random design. The

formation of heterogeneous groups based on this process is consistent with other methods used in CL (Cottell & Mills, 1993).

Time was provided for group members to introduce themselves and to exchange contact information. All groups then participated in two team building exercises, Lego building and puzzle completion, for extra credit points. Each exercise required the team members to work together to complete their task in competition with the other groups. In the Lego building exercise, the teams that built the tallest free standing structures received extra credit. In the puzzle completion exercise, each team was given a different puzzle with the same level of difficulty. The first teams to complete their puzzle received extra credit. To promote team interaction, some puzzle pieces were distributed amongst the other teams' puzzles. The students had to figure this out for themselves.

Each class met 28 times and took 20 quizzes given at the end of the class period. The content and format of the quizzes were the same for both classes. The numbers within the problems were changed to prevent the sharing of solutions between classes. Time spent in groups ranged from 10 to 20 min for each quiz. Each student completed a quiz on an individual basis and turned in an answer. The group then took the same quiz and turned in one answer. Solutions were provided at the end of the class session.

### *3.2. Subjects<sup>3</sup>*

There were 90 students initially enrolled in the two classes (45 in each). Eighty-five students completed the final exam and 80 completed the anonymous end-of-semester survey.

Field studies that use the classroom as a laboratory are, by definition, concerned with non-equivalent groups. Even in experiments that use random assignment, respondents can differ on factors that correlate with the dependent variable. However, in a quasi-experiment one must expect differences to exist. These differences can affect internal validity because respondents with certain characteristics may be more affected by the treatment or the characteristic may be related to the outcome being measured.

Therefore, information was collected on variables that might affect individual learning or group dynamics: age, gender, credit hours being attempted, prior grade, hours currently working, intended major, and ethnicity. The end-of-semester survey was used to measure the students' attitudes toward their group experience to evaluate whether group dynamics might provide an alternative explanation for the results.

Descriptive statistics are reported in Tables 1 and 2. As shown in Panel A of Table 1, there were no statistically significant differences at the  $p < .05$  level in age, hours worked, credit hours attempted, or grade in the first accounting course (prior grade). There were statistically significant differences ( $p < .05$ ) in gender and ethnic background (Table 1, panel B).

Table 1 also presents a frequency analysis on ethnic background, intended major, and gender by class. A chi-square test of independence was conducted on group formation method (formation) and intended major; formation and ethnic background; and formation and gender. No significant relationship was discernable between formation

---

<sup>3</sup> Approval for the use of human subjects in this experiment was received from the VCU Office of Research Subjects Protection (IRB# 02623). This experiment complied with all appropriate federal regulations.

Table 1  
Demographics

	Group <sup>a</sup>	Mean	Std. deviation	t-Value	df	Significance (2-sided)
<i>Panel A: Age, work experience, hours worked, registered credit hours, and prior accounting grade – means and t-test</i>						
Age	Self-selected	20.63	1.39	-1.678	78	.097
	Instructor	21.51	3.03			
Hours worked	Self-selected	16.60	13.28	-.546	77	.587
	Instructor	18.31	14.50			
Credit hours	Self-selected	13.61	1.90	.159	78	.874
	Instructor	13.53	2.781			
Prior grade	Self-selected	2.37	1.07	.420	79	.676
	Instructor	2.26	1.27			
	Self-selected groups (%)		Instructor groups (%)	Pearson chi-square value/df		Significance (2-sided)
<i>Panel B: Ethnicity, declared major, and gender – frequency analysis and chi-square test</i>						
Ethnicity <sup>a</sup>				11.79/5		.038
White	44		59			
Black	34		10			
Hispanic	7		0			
Asian	5		18			
Other	7		8			
Undisclosed	3		5			
	100		100			
Intended major <sup>b</sup>				7.83/8		.45
Accounting	22		13			
Other	78		87			
	100		100			
Gender <sup>c</sup>				5.15/1		.023
Male	36		61			
Female	64		39			
	100		100			

<sup>a</sup> Self-selected groups were those where the students chose their team mates. Instructor-formed groups were those where the instructor placed the students in teams based primarily on ability level.

<sup>b</sup> Based on end-of-semester student survey, *N* = 41 self-selected; *N* = 39 instructor formed.

<sup>c</sup> Based on student records, *N* = 44 self-selected; *N* = 41 instructor formed.

and intended major (Pearson chi-square value = 7.83, *df* = 8, *p* = .45). However, a statistically discernable relationship existed between formation and ethnic background (Pearson chi-square value = 11.79, *df* = 5, *p* = .038) and formation and gender (Pearson chi-square value = 5.15, *df* = 1, *p* = .023). These variables were controlled for in subsequent analysis.

Table 2 presents responses to a set of questions designed to determine the students' general comfort level with their groups because differences in group functioning could explain differences in learning unrelated to peer tutoring. Descriptive statistics and results of the analysis of covariance (using gender and ethnicity as covariates) are reported in Table 2. All student responses were positive (three of the four means were above 4 on a 5-point scale). While the self-selected groups' responses were consistently higher, there were no statistically significant (*p* < .05) differences in comfort level between the two formation methods.

Table 2  
Comfort level with group

Questions	Means <sup>a</sup> (Std. deviations)		F-values	Significance
	Self-selected ( <i>N</i> = 41)	Instructor ( <i>N</i> = 39)		
I was comfortable asking questions in my group	4.55 (.85) ( <i>N</i> = 40)	4.33 (.77)	1.279	.262
I believe that my efforts contributed to the group's success	4.32 (.72)	4.13 (.67)	2.262	.137
I believe that I made a unique contribution to the group	4.22 (.72)	4.18 (.68)	.466	.497
My success in this course has been at least partially determined by the success of my group	4.05 (.86)	3.85 (1.09)	.839	.363

This table reports means (standard deviations) and the results of the analysis of variance. The main effects (*F*-values) of the dependent variable (group formation method), along with the respective significance level, are presented. Gender and ethnic background are treated as fixed factors in the tests of between-subjects effects.

<sup>a</sup> Based on a 5-point Likert scale anchored with "strongly disagree" as 1 and "strongly agree" as 5.

Students were also asked if they knew the people in their groups prior to group formation. None of the students reported having close relationships with group members prior to formation. Therefore, students were able to form working relationships under both methods of group formation that were not the result of acknowledged pre-existing personal ties.

## 4. Results

### 4.1. Tests of H1

To test H1, heterogeneous groups produce a more effective learning environment than student self-selected groups as measured by individual academic performance, analysis of covariance was conducted with gender, ethnicity, and prior grade as covariates. A separate ANCOVA was conducted for each of the measures of academic performance; exam 1, exam 2, exam 3, final exam, and total semester points. No significant differences ( $p < .05$ ) in individual academic performance at the class level were found. Thus, H1 was not supported.

However, prior grade was significant ( $p < .05$ ) in each of the models and there was indication of an interaction ( $p < .10$ ) between formation method and prior grade in the exam 2 model suggesting the presence of an ATI. Therefore, the students were sorted by prior grade and the analysis described previously was repeated. The mean and standard deviations for exam scores and total points by formation method and prior grade are presented in Table 3. Tests of significance by prior grade, using ANOVA with gender and ethnic background treated as fixed factors, are presented in Table 4.

There are small numbers in each cell, so the results must be evaluated with care. The findings indicate that A students performed well under both group formation methods, but the self-selected groups had higher (statistically significant at  $p < .05$ ) scores on exam 2. The B students in the self-selected group consistently outperformed the instructor-formed groups. The differences were significant for exam 2 ( $p < .05$ ) and the final

Table 3  
Mean exam scores by formation method and by grade in first accounting course (prior grade)

	All students	Prior grade			
		A	B	C	D/F
<i>Exam 1 (130 pts.)</i>					
Self-selected	86 (19)	96 (19)	91 (15)	79 (24)	85 (15)
Instructor	88 (20)	104 (19)	80 (14)	88 (17)	82 (23)
<i>Exam 2 (130 pts.)</i>					
Self-selected	87 (22)	112 (22)	91 (16)	78 (22)	78 (19)
Instructor	87 (20)	103 (15)	71 (23)	87 (13)	83 (20)
<i>Exam 3 (130 pts.)</i>					
Self-selected	103 (20)	119 (13)	105 (17)	101 (25)	93 (14)
Instructor	104 (18)	118 (11)	96 (21)	107 (11)	97 (19)
<i>Final (310 pts.)</i>					
Self-selected	231 (44)	270 (26)	245 (43)	219 (39)	207 (40)
Instructor	234 (45)	271 (30)	218 (57)	228 (52)	218 (33)
<i>Total points</i>					
Self-selected	858 (130)	992 (66)	893 (93)	835 (87)	768 (165)
Instructor	847 (123)	975 (61)	750 (153)	837 (133)	814 (72)
<i>Number<sup>a</sup></i>					
Self-selected	44	7	13	13	10
Instructor	41	9	7	9	13

The standard deviation is presented in parentheses below the mean score.

<sup>a</sup> One student in the self-selected class and three students in the instructor-formed class did not take the first accounting course. Therefore, the individual cells do not sum to the total.

Table 4  
Effect of formation method on individual learning

Dependent variable	'A' students	'B' students	'C' students	'D/F' students
Total semester points	.781	6.220**	.073	.843
Final exam	.020	3.275*	.050	.385
Exam 1	.147	.613	1.096	.001
Exam 2	5.134**	7.527**	.233	.084
Exam 3	.005	.607	.078	1.075

This table reports the results of the analysis of variance conducted on formation method for students at each ability level (prior grade). The main effects ( $F$ -values) of the dependent variables are presented along with their respective significance level. Gender and ethnic background are treated as fixed factors in the tests of between-subjects effects.

\*,\*\* indicate significance at the .10 and .05 level, respectively.

Table 5  
Students' perceptions of the group experience

Questions	Means <sup>a</sup> (Std. deviations)		F-values	Significance
	Self-selected groups <i>N</i> = 41	Instructor groups <i>N</i> = 39		
<i>H2 – impact on individual learning</i>				
I believe that I acquired considerable knowledge from this course	4.51 (.71)	4.21 (.86)	3.145	.080
I feel that the time spent in class in my group was beneficial to my learning the accounting material	4.71 (.51)	4.28 (.79)	5.602	.021

This table reports means (standard deviations) and the results of the analysis of variance. The main effects (*F*-values) of the dependent variable (group formation method), along with the respective significance level, are presented. Gender and ethnic background are treated as fixed factors in the tests of between-subjects effects.

<sup>a</sup> Based on a 5-point Likert scale anchored with “strongly disagree” as 1 and “strongly agree” as a 5.

( $p < .10$ ). The C and D/F students had consistently lower exam grades in the self-selected groups, but the differences were not significant.

Exam 2 covers the most complex material presented in the course: net present value and relevant costs for decision making. The existence of a significant formation effect for higher ability students suggests that the interaction between formation and individual learning is complex and worth further investigation.

#### 4.2. Tests of H2 and H3

H2 and H3 are based on the students' perceptions of the CL group experience. These data were collected with the anonymous end-of-semester survey described previously. Analysis of variance (ANOVA) was conducted to examine the relationship between group formation method and student perceptions.

H2 states that student perceptions of the impact of the CL group experience on individual learning will differ between the self-selected and instructor-formed groups. Two questions related to this hypothesis were included in the survey. The means, standard deviations, and results of the ANOVA are presented in Table 5. The results indicate that student attitudes toward the role the groups played in individual learning did differ significantly ( $p < .05$ ) supporting H2. Students in the self-selected groups gave significantly higher ratings to the benefit of the group work in learning accounting material than students in the instructor-formed groups (4.71, self-selected, and 4.28, instructor-formed,  $p < .05$ ). Student responses to “acquired considerable knowledge” were also higher in the self-selected groups (4.51 versus 4.4) although the difference was not statistically significant at the  $p < .05$  level.

The objective of H3, students' preferences for the amount of classroom time devoted to CL activities will differ, was to obtain direct evidence of students' attitudes about CL. In prior CL studies, student preferences were derived from standard end-of-semester surveys. In contrast, questions included in the end-of-semester survey in this study specifically asked the students if they would have preferred to use the time allocated for group work to other classroom activities. Means, standard deviations, and results of the

Table 6  
Students' perceptions of the role of group work in course content (H4)

Questions	Means <sup>a</sup> (Std. deviations)		F-values	Significance
	Self-selected groups <i>N</i> = 41	Instructor groups <i>N</i> = 39		
I feel that my group participation was a positive experience	4.61(.67)	4.38 (.78)	.989	.323
I would have preferred to spend more class time working in my group	3.56 (1.03)	3.36 (1.11)	.759	.387
I would have preferred to spend the time in class allocated to group work:				
to cover homework problems	2.63 (1.18)	2.95 (1.23)	2.643	.108
on taking individual quizzes	2.24 (1.18)	2.10 (1.27)	.055	.816
on formal lecture on the accounting material	1.95 (1.05)	1.85 (1.11)	.129	.720

This table reports means (standard deviations) and the results of the analysis of variance. The main effects (*F*-values) of the dependent variable (group formation method), along with the respective significance level, are presented. Gender and ethnic background are treated as fixed factors in the tests of between-subjects effects.

<sup>a</sup> Based on a 5-point Likert scale anchored with "strongly disagree" as 1 and "strongly agree" as 5.

ANOVA are presented in Table 6. There were no significant differences ( $p < .05$ ) in responses, therefore, H3 is not supported. Students in both groups felt that their group participation was a positive experience (4.61, self-selected, versus 4.38, instructor-formed). The mean responses in both groups to a preference for spending the group time on covering homework problems, individual quizzes, or formal lecture were below 3.0. However, when asked if they would have preferred to spend more time on group activities, students indicated that they do find value in the more traditional classroom activities (mean responses were approximately 3.5).

## 5. Discussion

The general perception, at the college level, has been that heterogeneous groups are necessary for effective CL (Cottell & Mills, 1993). Considerable effort has been devoted to ensuring that groups are heterogeneous, especially with respect to ability level, to encourage interaction between higher and lower achievers. In this study, the impact of heterogeneity on individual learning was specifically examined (H1) through a quasi-experiment. Also examined is the effect of group formation method on student perceptions of how the group work affected individual learning (H2) and the role of the group in how class time is used (H3).

The results of this study revealed that instructor-formed heterogeneous groups are not a necessary condition for effective CL when the focus is on individual learning. In contrast, the results of the ability level analysis, ATI, indicate that allowing self-selected groups may increase the effectiveness of CL in terms of individual learning for higher performing students in some learning contexts.

While clearly tentative, these results do suggest that when looking only at improved individual academic performance, the preferred team composition and formation method for college students enrolled in an accounting class may not be the same for students having different levels of ability. This is similar to the results found in Webb

(1991) for pre-college level students in math courses. These results also imply that prior findings of no conclusive academic learning benefits from CL versus traditional learning (Lancaster & Strand, 2001; Ravenscroft et al., 1997) may be partially due to the level of the analysis. Re-evaluating the results of previous CL studies using ATI methodologies may provide stronger evidence on the effectiveness of the CL pedagogical method.

The findings from the analysis of student perceptions indicate that the self-selected groups were more conducive to individual learning than the instructor-formed groups. No statistically significant differences were found in students' preferences for the amount of time spent in class on CL activities between students in the self-selected groups and instructor-formed groups. The nature of the anonymous survey did not permit analysis of students' perceptions by ability level.

The data collected in this study do not permit reaching a definitive conclusion as to why higher performing students showed some sensitivity to the formation method. What the results do suggest is that the group formation method need not involve a complicated and time consuming mechanism when the class is diverse and not composed of groups of friends. This may encourage instructors who are interested in implementing CL but are unwilling to engage in a complicated group formation process.

Overall, students were satisfied with the format of the classes and preferred the CL group format to spending more time using traditional learning formats. This finding supports prior accounting pedagogical research that found increased student satisfaction and interactions under various implementations of CL as compared to traditional methods (Lancaster & Strand, 2001; Ravenscroft, 1997). These benefits appear to accrue under both the self-selected and instructor-formed group formation methods.

## **6. Limitations and areas for future research**

The results of this study are subject to several limitations. First, using grades received in a prior course provides only an approximation of student ability. At this university, the two semesters of introductory accounting are not integrated. The first semester focuses on external financial reporting. The second semester tends to rely more on critical thinking skills. This study used a grade measure as both a covariate (prior grade) and dependent variable (exam scores and total semester points), thus, any weaknesses and limitations arising from the use of grades will be present in this study.

Second, this experiment took place in a field setting using non-equivalent groups. As in all experiments with this design, individual differences in students, that were not identified or measured, may have contributed to the treatment effect. Additionally, given that the setting for this experiment was a large urban university with an ethnically diverse environment, the student self-selected groups may have been more diverse than they might be at other institutions. Instructors at smaller universities or those teaching upper division classes where friendships may be more prevalent and interests more homogeneous may observe different group dynamics and results than those observed.

Third, the dynamics of five-person groups may be significantly different from those of groups with three or four students. Therefore, the tentative conclusion about the different impact of grouping on middle ability level students may not apply to smaller groups.

Finally, there are many mediating factors that influence a student's achievement that have not been measured in this study. Students were provided with annotated solutions to quizzes and encouraged to correct missed exam questions for extra credit. The extent to which students took advantage of these activities and exerted other efforts to learn outside of the CL activity has not been captured. Formal laboratory experiments are needed to help control for some of mediating factors that have not been measured in this study. Randomly assigning subjects and controlling for effort expended outside of the group interaction would make it easier to determine the extent to which group interactions promote individual learning.

Future research in this area could look at the impact of different CL activities on academic performance and interpersonal skills. This study was limited to cooperative quizzes. Group projects, within-group review of homework problems, and presentations during class time would be alternative ways to expand the study of the impact of CL on academic performance and interpersonal skills development. Finally, as discussed previously, the results of this study imply re-evaluating the results of previous CL studies using ATI methodologies which may provide stronger evidence on the effectiveness of the CL pedagogical method.

### **Appendix 1. Grade component of group activity**

*QUIZZES:* After the first week, quizzes will be given in almost every class in which a test is not given. Two hundred and forty points have been assigned to quizzes. I will use the homework material assigned for that class (whether or not covered in class) to determine the content of each quiz.

You will be assigned to a quiz group of 4–5 students. Each member of the group will individually complete the quiz and hand it in for grading. Each group will retake the quiz as a single unit. The 240 points assigned to quizzes will be allocated equally between three grading components: the individual quiz, the group quiz, and peer evaluation points assigned to group members by other members of the group.

If you are not in class you will receive a zero for the individual component of the quiz. If you attend the classes before and after a missed quiz, you will receive the group grade as if you were in class. Otherwise you will receive a zero for the group component of the quiz.

*Group maintenance* To ensure that every team member is an active participant, each member of a team will evaluate the performance of other team members. This will be done four times during the semester. A copy of an evaluation form is included with the syllabus. I will average the evaluations to determine the score for each person. Note that raters must differentiate some in their ratings and cannot just divide by the number of members in the group.

*Example:* Each peer evaluation is worth 20 points. If your group has five members, then you must allocate 80 points among the other four members of your group. At least one must receive a 21, so at least one member must receive a 19.

**Appendix 2. Data summary sheet**

- 1. Name:
- 2. Grade in Acct 203
- 3. When did you take Acct 203?
- 4. Intended Major
- 5. Employer
- 6. Job description
- 7. Hours worked per week
- 8. Phone number

9. What effect do revenue and expenses have on owners' equity?

	<u>Revenue</u>	<u>Expenses</u>
A.	Increase	Increase
B.	Increase	Decrease
C.	Decrease	Decrease
D.	Decrease	Increase

10. The underlying principle that makes adjusting entries necessary is the

- A. Matching principle
- B. Objectivity principle
- C. Cost principle
- D. Consistency principle

11. Company records show the following information for the most recent fiscal year

Supplies Inventory, Beginning Balance	\$ 5,000
Supplies purchased during the year	\$17,500
Supplies Inventory, Ending Balance	\$ 4,900

What is the correct amount of supplies expense for the year?

- A. \$ 7,600
- B. \$17,400
- C. \$17,600
- D. \$27,400

12. The principal purpose for recording depreciation is to

- A. Provide a funding for financing replacement of depreciable assets
- B. Show conservative figures for financial statements
- C. Provide a deduction for income tax purposes
- D. Systematically allocate the asset's cost over its useful life

\*\*\*\*\*

**To answer the remaining questions, use a nine point scale. A 1 indicates that you completely disagree with the statement. A 5 indicates that you neither agree nor disagree. A 9 indicates that you completely agree.**

13. I feel that I am good at problem solving.

-----DISAGREE----- -----AGREE-----
1    2    3    4    5    6    7    8    9

14. My word processing and other computer skills are good.

-----DISAGREE----- -----AGREE-----
1    2    3    4    5    6    7    8    9

15. I feel that my leadership skills are strong.

-----DISAGREE----- -----AGREE-----
1    2    3    4    5    6    7    8    9

## References

- Clinton, B. D., & Kohlmeier, J. M. III, (2005). The effects of group quizzes on performance and motivation to learn: Two experiments in cooperative learning. *Journal of Accounting Education*, 23(2), 96–116.
- Collins, K. M., & Onwuegbuzie, A. J. (2000). Treatment by aptitude interactions as a mediator of group performance in research methodology courses. In *Paper presented at the annual meeting of the Mid-South Educational Research Association 28th*, Bowling Green KY, November 17–19, 2000, ED 448 203.
- Cottell, P., & Mills, B. (1993). Cooperative learning structures in the instruction of accounting. *Issues in Accounting Education*, 8(1), 40–58.
- Doran, B. M., Bouillon, M. L., & Smith, C. G. (1991). Determinants of student performance in Accounting Principles I and II. *Issues in Accounting Education*(Spring), 74–84.
- Edmonson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–352.
- Eskew, R. K., & Faley, R. H. (1988). Some determinants of student performance in the first college-level financial accounting course. *The Accounting Review*(January), 137–146.
- Hite, P. (1996). A treatment study of the effectiveness of group exams in an individual income tax class. *Issues in Accounting Education*, 11(1), 61–75.
- Hwang, N. R., Lui, G., & Tong, Y. J. W. (2005). An empirical test of cooperative learning in a passive learning environment. *Issues in Accounting Education*, 20, 151–165.
- Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. Working paper, University of Minnesota.
- Kagan, S. (1989). *Cooperative learning: Resources for teachers*. San Juan Capistrano, CA: Resources for Teachers.
- Lancaster, K., & Strand, C. (2001). Using the team-learning model in a managerial accounting class: An experiment in cooperative learning. *Issues in Accounting Education*, 16(4), 549–567.
- McConnell, C. A., & Sasse, C. M. (1998). Modifying a federal income tax course to include team and independent learning objectives. *Advances in Accounting Education*, 1, 43–68.
- Norman, C. S., Rose, A. M., & Lehmann, C. M. (2004). Cooperative learning; resources from the business disciplines. *Journal of Accounting Education*, 22, 1–28.
- Ravenscroft, S., Buckless, F., McCombs, G., & Zuckerman, G. (1995). Incentives in student learning: An experiment in cooperative group learning. *Issues in Accounting Education*, 10(1), 97–109.
- Ravenscroft, S., Buckless, F., & Zuckerman, G. (1997). Student team learning-replication and extension. *Accounting Education: A Journal of Theory, Practice and Research*, 2(2), 151–172.
- Slavin, R. (1983). *Student team learning: An overview and practical guide*. Washington, DC: National Education Association.
- Swanson, Z. L., Gross, N. J., & Kramer, T. (1998). Alternative modes of study group formation and student examination performance. *Accounting Educators' Journal*, X(2), 1–11.
- Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, 22(5), 366–389.