Accounting Educators' Journal
Volume X, Number 2

Alternative Modes of Study Group Formation and Student Examination Performance

Zane L. Swanson
Emporia State University

Norma Jean Gross
Thomas Kramer
Saint Louis University

Abstract

The nature of group activities is an important part of the workplace and consequently the educational process that prepares students for employment. This study's analyses compare individual student exam performance according to alternative means of study group formation. Several analyses of variance designs examine group formation effects on student performance while controlling for cumulative grade point averages, gender and particular tests. The findings indicate that student performance is higher in groups which are formed by student choice as contrasted with student performance in groups assigned from a pool, or the performance of those students who elect not to be in a group. The results indicate that groups may help weaker students.

Data Availability: The data will be made available upon request to the authors, but students will not be identified in any way. Supply a 3.5" disk.

Introduction

One of an instructor's decisions is whether or not to have group activities as part of a course. Groups have been used for numerous purposes including exam study preparation and collaborative in-class learning. In general, groups can facilitate a more thorough knowledge of accounting subject matter. There are many advantages of employing groups (Gruber and Weitman 1962, Beach 1960 and 1968 and Webb and Grib 1967) (e.g., improving weaker students' performance), but also some pitfalls (Parry 1990) (e.g., students may not be able to discern appropriate roles). Even if an instructor does not purposely form groups, some students will gather together on an informal basis. If the instructor's decision is to use groups, then the next issue is the determination of group members. This paper examines alternative means of forming groups and analyzes the association of individual exam performance with the mode of group formation. This study's experimental design permitted students: (1) to form a

group of their own choosing; (2) to join a pool that the instructor used to assign group membership; or (3) to remain unaffiliated with a group. The stated group function was to assist each other in exam preparation, but students were encouraged to work together at other times as well. The level of the resources devoted to the group was at the students' discretion.

The student test scores are compared on a basis of the set of individual tests for the semester. The independent variables are the cumulative grade point average (GPA), recitation section and mode of group formation. The cell means show that individual performance is significantly (α=.05) higher in the self-selected groups than the students from pooled groups or students who elect not to be in a group. There is no significant difference at α=.05 between the scores of the students in the pooled groups or those individuals who elected not to be in a group.

The remainder of the paper is organized in three sections. The next section reviews the motivation for group activities in accounting education and discusses a theoretical framework of factors that influence student performance. The following section covers the experiment. The final section gives conclusions and discusses implications.

Motivations for Group Activities and a Performance Framework

There are two major reasons for examining group behavior in accounting education activities. First, in the 1990's various reform groups and committees have issued statements that indicate group activities should be part of the accounting educational experience. Second, group activities have the potential to increase student performance and instructors should act to maximize student learning.

Group Activities in Accounting Education

During the 1990s, accounting education reform discussions (Williams 1993) make repeated references to group activities. The printed versions of the reform positions state that actual workplaces often use groups to handle complex tasks. The argument is *ipso facto* that accounting education should also place students in team situations (e.g., audit teams). An increasing body of collaborative learning research (Caldwell, Weishar and Glezen 1996 and Hite 1996) is generally supportive of the team approach. The design of classroom experiences is also being modified to reflect collaborative learning values (Peek, Winking and Peek 1995). This issue also relates to the core purpose of education which is to provide students with knowledge and experiences to be successful in a particular career. Alternatively, arguments are made in faculty meetings that the educational process serves as a gatekeeper for entry into professional careers. If one holds this proposition to be true, then the educational process needs to insure that group activities are part of the hurdles that students must pass. In general, the reform movement has not addressed, or only superficially mentioned implementation issues. Also, the collaborative learning literature has not yet analyzed the acculturization process for its impact on the effectiveness of group learning. The current paper addresses a key implementation problem of group

formation mode with respect to study groups.

Performance Framework

Instructors should be very interested in getting their students to achieve the maximum scores possible, if for no other reason than enlightened self-interest. This paper makes the proposition that group study activities should improve student performance and then tests the hypothesis. In order to test this proposition, it is appropriate to identify a suitable framework to examine the hypothesis.

This study will use the Lawler and Porter (1967) model which has previously been used by Parry (1990) in analysis of MBA students taking financial accounting[1]. Parry's findings were that students in assigned groups did not fare better than students who were not in groups. The current study differs because students are given the choice of: (1) forming their own groups, (2) being assigned a group from a pool or (3) electing not to join a group. The current study also differs because it uses undergraduate introductory financial accounting students.

The Lawler and Porter (1967) model (See Figure 1) proposes that individual performance is a function of ability, effort and role perception which are very basic irrefutable factors. The effort is driven by the value of the reward and the probability that the reward depends on the effort. In the current study, ability is measured by cumulative GPA. The GPA has been shown to be indicative of a business student's ability to perform in accounting classes (Eckel and Johnson 1983, Hicks and Richardson 1984, Ingram and Petersen 1987 and Doran et. al 1991).

The effort that a student makes is presumed to be a multiplicative function of the value of the reward and the probability that the reward depends upon the effort. Every instructor has witnessed evidence that undergraduate students are driven by the desire to attain the highest cumulative grade point possible because they know that their future job opportunities depend principally on the grades.

In this particular study, students were given two extra percentage points if every member of the group achieved a "C" average or better. In this way, the design made a link between the effort and the reward. The selection of two extra points as a reward was arbitrary. The reward had to be sufficient to motivate the student (i.e., potentially raise a student to a higher semester grade level). The relatively low reward does work against any finding and makes the experiment conservative in nature. If others raise the reward level, then a proportionately higher performance should be expected within the context of diminishing returns and a reduced motivation on other learning objectives.

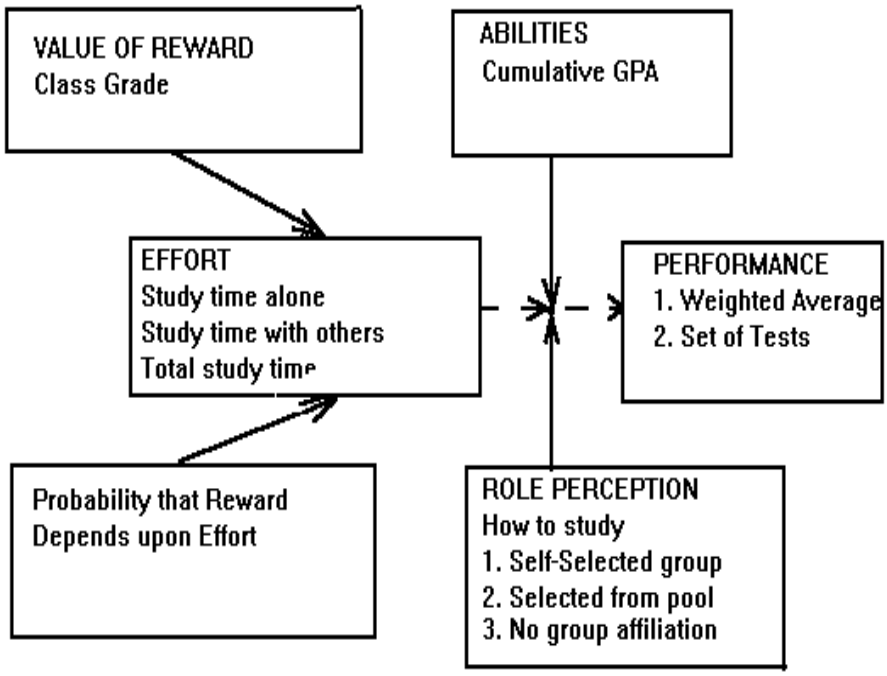
In this particular analysis, the role perceptions are likely to be driven by levels of knowledge about fellow students. It must be noted that students could discover a "social slacker" in their group during the semester. The discovery of a "social slacker" is less likely in the groups that formed by choice because the likelihood is higher that students have more prior knowledge about each other. On the other hand, groups that are assigned from a pool are more likely to have a "social slacker" problem (i.e., someone who wants the extra two points

---

1Ott et al. (1988) and Ott (1988) adapted the Lawler and Porter model for educational purposes.

but doesn't want to work for them). At the minimum, the students from the pooled group will have to invest personal capital and time in the discovery phase of interactions. The students who elect not to be in a group may have decided that the potential two points aren't worth the time spent in the group, or may be self-conscious about group activities. The individuals who aren't in groups obviously will receive no input from others, but don't have to expend any time in a group activity. The next section formulates an experiment that compares performance according to the previously stated three alternative modes of studying for examinations.

Figure 1
Lawler and Porter Theoretical Model



Empirical Analyses

Description of Experiment

The subjects were students in an undergraduate introductory financial accounting class during Spring of 1996. These students were generally sophomores with a few rare freshmen with advanced standing. There was also one staff member who took the class for credit. The instruction delivery was a three-times-a-week lecture and a once-a-week recitation section. In the recitation sections, the teaching assistants go over the homework with students.

The same instructor taught two back-to-back lecture times. Because the contiguous lectures should be similar, it

should not be necessary to control for this factor. In fact, when this feature was included, its significance varied greatly (α=.01 in Table 2's model and α=.76 in Table 3's model) for the same dependent variable. This behavior suggests multicollinearity. Thus, the appropriate remedy was to drop the lecture-time variable and this step was taken here. Note, the main conclusions were the same with or without this effect.

On the first day of class, the instructor explained the study group formation options of: 1) choosing a group of four members by the first exam, (2) submitting their names to a pool by the first week that the instructor used to assign groups, or (3) electing not to participate in a group. The instructor read a protocol concerning the experiment as required by school institutional review board procedures in which the stated primary function of the groups was a study support mechanism for exams.

The pool students were randomly assigned with the aim of forming heterogeneous groups. To this purpose, it was initially decided to rebalance groups which had more than two accounting majors, but this situation also did not arise. Groups were encouraged to work on in-class exercises together, but were not required to do so. The reward for good group performance was two percentage points which were added to their final semester average. The school grading system does have B+ and C+ categories which may provide more motivation impact than the two points might indicate on a superficial basis. Every member had to have a 70 percent overall semester average (C grade) in order for the group to receive the two-percentage-point reward.

The course had four tests. There were two in-class one-hour quizzes worth 75 points each, a two-hour midterm for 150 points and a two-hour final for 150 points. The first quiz was given five weeks into the semester. The midterm covered material for the first half of the semester. The second quiz was given three weeks after the midterm and covered only material subsequent to the midterm. The final exam only covered material after the midterm. Not all students took every exam because of individual problems (e.g., sickness).

All exams had the same format. There were declarative knowledge definitions and multiple-choice questions which constitute 70 percent of the potential raw score. There were two questions for 20 percent which mirrored homework numerical problems and require an integrative approach. One question for 10 percent required interpretation of actual company financial statements. Table 1 Panel A has descriptive statistics about the four exams and the group formation modes.

The four exam averages range from 73.38 to 80.39 before any group reward[2]. A seven point difference was worthy of further investigation because the quizzes and tests do cover different material. Quiz one covers the accounting cycle. Exam one (midterm) tests financing activities in addition to the accounting cycle. Investing and some operating activities are tested in quiz two. The final test (exam two) covers investing and operating

[2].All subsequent score metrics refer to performance before the two point reward was added to the end of semester weighted grade average.

activities and cash flow statements. According to the cell means in Panel B's linear model of Table 1, there is a significant (α=.05) difference between the first half of the semester results and the second half. Therefore, the exam is included as a covariate in one of the analyses in order to control for exam design differences, even though the conclusions are the same without this covariate.

Table 1: Test Descriptive Information

Panel A: Sample Characteristics

Variable Tests	N	Mean	Std Dev	Minimum	Maximum
Quiz 1	95	79.83	13.26	38.00	98.00
Exam 1	94	80.39	12.24	46.50	100.00
Quiz 2	91	73.38	12.03	48.00	99.00
Exam 2	89	76.15	14.42	47.00	100.00
Groups-Overall Statistics					
Choice	167	79.43	11.71	52.00	99.50
None	96	76.83	14.23	38.00	100.00
Pool	106	75.05	14.31	46.00	100.00

Panel B: ANOVA Specific Test Effect

Dependent Variable: Test Score

Source	df	Sum of Squares	Mean Square	F Statistic	P Level
Test	3	3005.047	1001.682	5.92	.0006

R^2=.05

Duncan's Multiple Range Test for variable: Test Score Alpha=0.05

Means with the same letter are not significantly different.

Duncan Grouping	Mean	Obs.	EXAM
A	80.390	94	Test 1
B A	79.832	95	Quiz 1
B C	76.151	89	Test 2
C	73.385	91	Quiz 2

Panel C: ANOVA Gender Effect

Dependent Variable: Test Score

Source	df	Sum of Squares	Mean Square	F Statistic	P Level
Test	1	3611.443	3611.443	21.67	.0001

R^2=.06

Duncan's Multiple Range Test for variable: Test Score Alpha= 0.05

Means with the same letter are not significantly different.

Duncan Grouping	Mean	Obs.	Gender
A	81.713	131	Female
B	75.175	238	Male

On the top end, the high score in each category is close to the maximum possible. On the bottom end, none of the subject data was zero. These facts indicate that the existence of outlier observations is minimal in any category and that mode choice did not preclude good performance.

The choice group students have a variance descriptive statistic that is lower than either the pooled category subjects or students who did not select a group. This information suggests that the choice group had a better consensus about accounting material.

The academic community and society at large are very interested in gender effects on the educational process. In a similar manner to the previous test analysis, the scores are examined for gender differences. The results of an ANOVA are in Panel C of Table 1. Females scored significantly higher than males in this sample. As a consequence of this gender significance, gender is included as a covariant in the analyses, even though the percentage of participation in groups by gender does not differ appreciably from the male/female ratio of the entire class.

Research Design and Results

The research design purpose is to test whether the mode of group formation associates with test scores. The analysis used the Scientific Analysis System (SAS) Proc GLM ANOVA test procedure and examined the mode of group cell means. Two different ANOVA analyses are made of the data according to the construction of the dependent performance variable.

In the first analysis, the performance measure is the individual's test scoring. There is a maximum of four observations from each student. A repeated measure design is used. The results are in Table 2. Consistent with Eskew and Faley (1988), GPA explains a significant (α=.01) part of the variance. Particular exams do have an effect consistent with the findings of Table 1. MODE, the mode of group formation, is significant at the .05 level. Given the relatively low reward for group activities, this finding suggests that study groups are useful. Panel B of Table 2 has the cell means for the group selection mode. There is a significant (α=.05) difference between students' scores who were in groups formed by choice and both of the other two group formation modes (the student in groups formed from a pool and student who elected not to be in a group.)

One of the initial stated group purposes was to help the weaker students. Therefore, the independent variables are organized by: categorical values for GPA, mode of group formation and an interaction variable of group formation and the inverse of GPA. The reason for the GPA inverse is the need to measure the effect on the weaker students who should be helped relatively more by group activities. The categorical groupings are appropriate to control for different levels of student knowledge. Teachers express the perception that stronger students' scores will not be helped by group activities and it is worthwhile to examine this issue. The results in Table 3 are consistent with Doran, Bouillon and Smith (1991) that categorical groupings of GPA have a significant (α=.01) affect on performance. The interaction between the mode of group formation and the GPA inverse is significant (α=.01). Thus, weaker students appear to be helped by groups even after controlling for gender

(α=.01) and the level of a student's knowledge. The main effect of the group formation is also significant (α=.01) in this particular research design. The group formation cell means show a similar but less significant finding than Table 2 because of the previous approach has more power.

Table 2: Repeated Measure ANOVA Analysis by Exam (OBS=369)

Panel A

Dependent Variable: Test Score

Source	DF	Sum of Squares	Mean Square	F Statistic	P Level
GPA	1	23405.33	23405.33	351.29	.0001
GENDER	1	276.49	276.49	4.15	.0426
MODE	2	463.15	231.57	3.48	.0324
ID(MODE)	90	18930.42	210.34	3.16	.0001
EXAM	3	3833.73	1277.91	19.18	.0001
MODE*EXAM	6	221.96	36.99	.56	.7657

R^2=.73

Panel B

Duncan's Multiple Range Test for variable: Test Score
Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	Mode
A	79.43	167	Choice
B	76.83	96	None
B	75.05	106	Pool

Where
EXAM =Exam(1=quiz 1, 2=test 1, 3=quiz 2, and 4=test 2),
MODE =Mode of group formation(c=by Choice; n=None; p=from Pool),
GPA =Cumulative grade point average, and
ID =Individual student identifier.

Next, the analysis further investigates the proposition that the best students formed groups by choice. To check this proposition, the dependent variable is group formation and the independent variable is grade point average. The results indicate that students with higher GPAs formed into groups by choice. Most professors will say that the finding is not surprising, but these results document the anecdotal beliefs. We discuss the implications in the next section.

Discussion and Conclusions

This study analyzes the proposition that the method of study group formation relates to student performance. The findings indicate that individuals perform better in groups which form of their own accord versus groups formed by assignment or students who elect not to be in a group. In this study, there is no practical difference between student scores in groups formed by pool and student scores of individuals who elect not

Table 3: ANOVA Analysis by Grade Levels (OBS=369)

Panel A

Dependent Variable: Test Score

Source	df	Sum of Squares	Mean Square	F Statistic	P Level
GENDER	1	3611.44	3611.44	31.21	.0001
GPTOP	1	12358.68	12358.68	106.80	.0001
GPMID	1	3061.77	3061.77	26.46	.0001
MODE	2	705.30	352.65	3.05	.0487
GPINV*MODE	3	3393.34	1131.11	9.78	.0001

R^2 =.36

Panel B

Duncan's Multiple Range Test for variable: Test Score
Means with the same letter are not significantly different.

Duncan Grouping		Mean	Obs.	Mode
A		79.43	167	Choice
A	B	76.83	96	None
	B	75.05	106	Pool

where
GPTOP = 1 if cumulative grade point average $\geq$ 3 and 0 otherwise,
GPMID = 1 if cumulative grade point average $\geq$ 2 and < 3 and 0 otherwise,
GPINV = 1 /(Cumulative grade point average), and
MODE = Mode of group formation(c=by Choice; n=None; p=from Pool).

to be in a group. There is a consistency in this particular result with Parry (1990).

One could ask: what is it about groups formed by students on their own that led to better test scores? There are a few possibilities. First, it may be that students who formed their own groups did so with other class members they knew. This familiarity may have made it easier for the students to interact with each other, and thus profit more from the information of other group members. A second possibility is that students who formed their own groups met more often. Students who are already friends may know each other's schedules and/or formed groups with people who did have a compatible schedule. Or they may be more willing to accommodate the needs of their group by changing their routine, such as when they work or watch TV, so the group could meet. If, for these or other reasons, groups who formed on their own were able to meet more often, then the improved performance may be due to simply the increased frequency of meeting and therefore the opportunity to exchange more information. No data was collected to determine how frequently the groups met or about the quality of the interaction within the groups when they did meet. However, the lower variance of the scores in groups formed by choice does suggest that these groups were able to reduce the amount of uncertainty in the decision-making. Further research may need to gather additional information on such variables as the frequency of meeting, the

quality of the interaction of the groups when they meet, the level of participation of each of the members, and the like, to determine the underlying mechanisms that contribute to present findings. Whatever the underlying mechanisms turn out to be, the major finding of this study is that students in study groups formed voluntarily perform better than those in groups formed by the instructor or those who choose not to be in any study group.

Table 4: ANOVA Analysis Prior GPA by Group Formation (OBS=95)

Panel A

Dependent Variable: Grade Point

Source	df	Sum of Squares	Mean Square	F Statistic	P Level
MODE	2	4.019	2.009	5.37	.0062

R^2=.10

Panel B

Duncan's Multiple Range Test for variable: GPA

Means with the same letter are not significantly different.

Duncan Grouping	Mean	Obs.	M
A	3.1659	42	Choice
B	2.7596	25	None
B	2.7449	28	Pool

where

MODE = Mode of group formation (c=by Choice;n=None;p=from Pool).

Given that the mode of group formation does associate with performance, if students self-select groups, then they tend to choose fellow members with higher GPAs. Some students may be shut out of study groups and a consequent issue arises as to the scope of instructor responsibility in providing a successful learning experience. However, when an instructor assigns group membership, there will be inevitable tension between members because the higher performing students will resent the lower performing students occupying time or being social slackers. One conclusion is that if groups are to function effectively, they have to have a set of directed tasks or a well-defined mission from the instructor. A second conclusion is that student groups that are formed from prior acquaintances appear to be more successful, possibly because they are better able to focus on the course objectives. Thus, if groups are to succeed, then the instructor needs to provide appropriate initial guidance on group behavior and social interactions so as to minimize socialization costs.

Our and Parry's (1990) findings indicate no difference between the performance of assigned group members and those individuals who are not in a group. However, our findings show that groups formed by choice have higher individual scores. In general, can we improve the performance of students through group involvement? The answer may be in structuring of initial group activities and training of group members.

References

Beach, L. R. (1960). Sociability and Academic Achievement in Various Types of Learning Situations. ***Journal of Educational Psychology***, 208-212.

_____. 1968. Student Interaction and Learning in Small Self-directed Groups, Final Report. Washington, DC: Department of Health, Education and Welfare.

Caldwell, M., J. Weishar & G.W. Glezen. (1996). The Effect of Cooperative Learning on Student Perceptions of Accounting in the Principles Courses. ***Journal of Accounting Education*, Vol**(No. 1): 17-36.

Doran, B. M., M.L. Bouillon & C. Smith. (1991). Determinants of Student Performance in Accounting Principles I and II. ***Issues in Accounting Education***, (Spring): 74-84.

Eckel, N. and W. A. Johnson. (1983). A Model for Screening and Classifying Potential Accounting Majors. ***Journal of Accounting Education***, (Fall): 57-65.

Eskew, R. & Faley, R. (1988). Some Determinants of Student Performance in the First College-Level Financial Accounting Course. ***The Accounting Review*, Vol 63**(No. 1): 137-147.

Gruber, H.E. & M. Weitman. (1962). Self-Directed Study: Experiments in Higher Education, Report No. 19. Boulder: University of Colorado, Behavioral Research Laboratory.

Hicks, D. W. & F. M. Richardson. (1984). Predicting Early Success in Intermediate Accounting: The Influence of Entry Examinations and GPA. ***Issues in Accounting Education***, (Spring): 61-67.

Hite, P. (1996). An Experimental Study of the Effectiveness of Group Exams in an Individual Income Tax Class. ***Issues in Accounting Education***, (Spring): 61-75.

Ingram, R. W. & R. J. Petersen. (1987). An Evaluation of AICPA Tests for Predicting the Performance of Accounting Majors. ***The Accounting Review***, (January): 215-23.

Lawler, E.E. III & L. Porter. (1967). Antecedent Attitudes of Effective Managerial Performance. ***Organizational Behavior and Human Performance***, 122-42.

Ott, R.L. (1988). Pretest Reviews in Intermediate Accounting: An Empirical Analysis. ***Issues in Accounting Education***, (Fall): 378-87.

_____, D.S. Deeines & D.P. Donnelly. (1988). The Use of a Fundamental Practice Set in Intermediate Accounting. ***Issues in Accounting Education***, (Spring): 131-8.

Parry Jr., R. (1990). The impact of assigned study groups on study effort and examination performance. ***Issues in Accounting Education***, (Fall): 222-239.

Peek, L., C. Winking & G. Peek. (1995). Cooperative Learning Activities: Managerial Accounting. ***Issues in Accounting Education***, (Spring): 111-125.

Webb, N. J. & T.F. Grib. (1967). The Experimental Use of Student-Led Groups, Final Report Washington, DC: Department of Health, Education and Welfare.

Williams D. (1993). Reforming Accounting Education. ***Journal of Accountancy,*** (August): 76-82.