THE EFFECT OF GROUP PROJECTS ON CONTENT-RELATED LEARNING

Donald R. Bacon
University of Denver

Business schools often assign student group projects to enhance student learning of course content and to build teamwork skills. However, the characteristics of effective collaborative learning tasks, including group goals and individual accountability, are often not found in student group projects assigned in business classes. The current research found that content learning was actually inhibited by the use of a group project. The results indicate that the students who completed a project in groups learned less of the project-related content than did students who completed a shortened version of the project individually. The characteristics of business school group projects, peer-learning projects, and group projects in the workplace are compared and contrasted. Implications for program and course design are discussed.

Keywords: *collaborative learning; peer-learning; social loafing; group projects; assessment of learning outcomes; Rasch measurement*

Student group projects are a ubiquitous feature of business education for a variety of reasons. In Herman, Keldsen, and Miller's (2001) survey of business school deans and MBA program directors, improving learning was rated as the most important reason for using teams, followed by the practical experience that students gain from working in teams. Although these administrators rated practical considerations, such as facilitating grading, as being

Author's Note: The insightful comments on earlier drafts from Kathy Green, Kim Stewart, and two anonymous reviewers are gratefully acknowledged. I also greatly appreciate the financial support received from the Daniels College of Business. Please address all correspondence to Don Bacon, Department of Marketing, Daniels College of Business, University of Denver, 2101 S. University Blvd. (Rm. 495), Denver, CO 80208; e-mail: dbacon@du.edu

JOURNAL OF MANAGEMENT EDUCATION, Vol. 29 No. 2, April 2005 248-267
DOI: 10.1177/1052562904263729

much less important, other researchers have suggested that these considerations, including increased class sizes, play a significant role in the decision to assign group projects (Bolton, 1999; Kimber, 1996).

The widespread use of group projects has led to extensive research on their use. Much of this research has focused on the factors that affect team performance, such as cohesion (Deeter-Schmelz, Kennedy, & Ramsey, 2002; Gosenpud & Washbush, 1991), individual ability (Bacon, Stewart, & Stewart-Belle, 1998), methods of assignment of students to teams (Bacon, Stewart, & Anderson, 2001), and other factors that the teacher may control such as team size, the use of peer evaluations, the duration of the team project, and the use of team-building exercises (Bacon, Stewart, & Silver, 1999; Bolton, 1999; Bowen, 1998). Yet one of the most frequently cited reasons for using team projects, the fundamental assumption that group projects improve student learning, is underresearched in the business education literature. At the same time, extensive research has been conducted in the broader education literature on the effectiveness of learning in groups. Generally, this research found that group learning improves learning outcomes but only when the group tasks involve group goals and individual accountability (Slavin, 1988).

The current paper examines the issue of group learning effectiveness in traditional business school projects by first contrasting the recommendations of the group learning literature with common practice in business schools. The article then empirically compares the learning outcomes of two student samples: one sample in which students completed a marketing project working in pairs, and a sample in which students completed a shortened, but otherwise nearly identical project working alone.

Background Literature

The group learning literature is so extensive that several substreams exist, each with a slightly different tradition. This literature will be covered first, followed by a discussion of common business school practices and how these practices often deviate from the recommendations of the group learning literature.

COOPERATIVE LEARNING, COLLABORATIVE LEARNING, PEER LEARNING

In the group learning literature, cooperative learning is closely related to collaborative learning, so much so that not all writers agree on the difference. Cooperative learning is often used to describe learning environments with somewhat structured tasks, while collaborative learning typically describes learning environments with somewhat unstructured tasks (Strijbos, 2000). The collaborative learning tradition is more likely to embrace the notion that knowledge is subjective and jointly created through the interaction of the teacher and the learners (Mallinger, 1998; Shaw, Fisher, & Southey, 1999). Following McKeachie (1994), the general term *peer learning* will be used here to include cooperative and collaborative learning, as these involve an "interdependence of group members in working toward a common goal" (p. 143).

Slavin (1988), a leading theorist and researcher in collaborative learning, identified two conditions as being essential for peer-learning success: positive interdependence, which he also called group goals, and individual accountability. In many peer-learning environments, not only does the group produce a group product, such as a paper, but in addition the learning of each individual is measured later by an exam or a written assignment. Slavin's view is supported by many researchers in the group learning field (see reviews by Kimber, 1996; Strijbos, 2000).

Michaelsen's popular "team learning" method (Michaelsen & Black, 1994; see also Hernandez, 2002) provides one innovative example of how the essential conditions of positive interdependence and individual accountability can be implemented in a group activity. In this method, students take a quiz individually, then discuss the quiz in groups, and then take the quiz over again as a group. The students' grades are then determined by a weighted average of their individual quiz score and their group quiz score. When operationalized in this manner, the learners are each assessed individually (individual accountability), and they also depend on each other while they contribute to each other's learning (positive interdependence). In the jigsaw technique, another method of peer learning that has been applied in business education (Mesch, 1991), each member of a student team becomes an expert in one area and then is responsible for teaching the rest of the team. The team then takes a group quiz, which leads to positive interdependence. However, in Mesch's approach, individual accountability is accomplished mainly through peer pressure, and the weak individual accountability in this method can lead to some free riding (Mesch, 1991).

A simpler version of the jigsaw technique is a variation of the learning cell (Goldschmid, 1971). In learning cell designs, each peer-learning group comprises only two people. Thus, the learning cell "jigsaw" has only two pieces, such that each student learns one half the material and then teaches one other student, who, in turn, teaches the first student the other one half of the material. A series of studies demonstrating the beneficial effects on learning outcomes of this and other variations on these two-member learning cells are reviewed in McKeachie, Pintrich, Lin, Smith, and Sharma (1990).

Although there is clear evidence that students learn well in well-designed peer-learning environments (Kimber, 1996; McKeachie et al., 1990; Slavin, 1988), there is less clarity about the exact cause of these effects. Bargh and Schul (1980) found that peer learning is beneficial because students study differently when they anticipate having to teach other students. Webb and Grib (1967) suggested that students may feel more comfortable asking questions and expressing their own opinions in peer-learning environments. Students may also feel more mutual support and stimulation, thus increasing their motivation (Mallinger, 1998; Schomberg, 1986; Williams, Beard, & Rymer, 1991). The experience of explaining content to peers also provides an opportunity for elaboration (McKeachie et al., 1990). The student-as-teacher must extend his or her own understanding beyond superficial memorization to comprehension (Bloom, 1956) and meaningful learning (Ausubel, 2000). Especially for business students, group learning is a form of experiential learning, where they can learn about team functioning from their own experiences (Kahn, 1995; Williams et al., 1991).

The effect of group size on peer learning is an interesting and somewhat unresolved question. The literature on social loafing, which indicates that individuals expend less effort as group size increases, suggests that larger groups would be less appropriate for peer learning than smaller groups. For example, in Ringelmann's classic rope-pulling experiments (never published but described in Latané, Williams, & Harkins, 1979), blindfolded individuals pulled with less effort when they believed that others were pulling with them, and the more people they believed were pulling, the less effort they expended. Therefore, the social-loafing effect should lead to less effort, and therefore less learning, with larger groups.

This question of group size can also be addressed by closely examining the hypothesized causal mechanisms that underlie peer learning mentioned above. Webb and Grib (1967) and McKeachie et al. (1990) seem to support the notion that the learning is caused by the group dialogue. Yet, as group size increases, each person has less opportunity to participate (a form of what Diehl & Stroebe, 1987, call "production blocking"), thus reducing the potential benefit of participation. In larger groups, participants may also feel increased anxiety about speaking in front of the group, and this anxiety may inhibit their participation as well. Thus, the group-dialogue benefits of peer learning likely decrease with increasing group size. Bargh and Schul's (1980) experiments offer compelling evidence that the greatest benefit of peer learning comes from the preparation to teach others, not the act of teaching others. In their first experiment, students who studied material thinking they would have to teach another student performed better on a test than students who studied the same material without thinking they would have to teach another student. In a second experiment, all students first learned some new material, and then some students actually taught other students while others did not. The students who taught another student performed no better on a postteaching assessment than students who did not teach another student (although the length of the teaching experience was limited). If studying as preparation for teaching leads to more learning than studying for one's own learning, then the benefit of peer learning would come mainly from the individual preparation, and not the group experience, and group size may not have any impact on learning.

Larger groups offer one benefit not found in smaller groups; they generally have greater group resources. A larger group is more likely to have at least one member who understands some difficult aspect of the course material and can communicate that understanding to the others. However, whether groups effectively use these potential resources is questionable. Larger groups experience larger coordination losses, which refer to the inefficiencies involved in coordinating the efforts, time schedules, comments, and so on of many individuals. As coordination losses rise exponentially with group size (Latané et al., 1979), we would expect larger groups to be less efficient in many respects, including learning. Some support for this contention is seen in the management education literature. Bacon et al. (1998) found that teams of two generated projects of higher quality than students working alone, however teams of three or more generated projects of no higher quality than teams of two. Furthermore, in an examination of student perceptions of their team experiences, Bacon et al. (1999) found that although students felt they learned more course content from their best team experiences, team size was not a significant predictor of whether a team experience would later be considered one of their best or one of their worst. In summary, although there is no definitive support for any particular group size in peer-learning experiences, there is strong theoretical and empirical evidence that a group size of two is adequate to achieve the benefits of peer learning, as long as the learning task involves positive interdependence and individual accountability.

TRADITIONAL BUSINESS STUDENT GROUP PROJECTS

In business classes, the typical student group project differs in some important ways from the prototypical peer-learning task. On business projects, students jointly produce a "deliverable" of some sort, such as a written report or group presentation. Each member of the group then often receives the same reward, typically the same grade. Thus, a group goal exists, however often there is little if any formal individual accountability.

Social loafing is commonly observed in situations with group rewards and lack of individual accountability (Comer, 1995). The presence of social loafers, or free riders, in student groups is commonly observed and widely researched (Bacon et al., 1999; Strong & Anderson, 1990). When students free ride in groups, we would expect that the group task would do little if anything to enhance their learning of project-related content. To overcome the free-rider problem, some teachers integrate end-of-the-semester peer evaluations to adjust individual grades based on contribution to the group. Strong and Anderson (1990), however, suggested that peer evaluations may not work because they allow students to penalize free riders while avoiding confrontation, and direct confrontation is the best way to offer the peer pressure and guidance necessary to reduce free-riding behavior. Bacon et al. (1999) found empirical support for the hypothesis that peer evaluations can be counterproductive. Their study asked students to describe their best and worst team experiences on a number of measures, including the presence of social loafing and the use of peer evaluations. They found that social loafing and the use of peer evaluations were directly related to worst team experiences. These observations led Bacon et al. to recommend against using traditional (anonymous, end-of-course) peer evaluations and instead simply allowing teams to fire members (see also Kahn, 1995).

Even with more effective peer evaluations, perhaps done face-to-face during the course, or with the option to fire members, business school projects often possess a higher degree of specialization of labor than do peer-learning tasks, going so far as to limit the scope of what students will learn. In these more functional, cooperative student groups, the students divide up the work equitably and then dutifully perform their assigned tasks. One student may perform the library research, for example, while another student creates PowerPoint slides, and another conducts the financial analysis. One group member might be assigned the task of combining all the individual contributions into a coherent project. Faculty generally encourage this approach to group management, as it mimics the approach to group projects in the business world. The type of workplace team that most closely resembles a well-run student project team is what Meyerson, Weick, and Kramer (1996) described as a “temporary group.” These groups typically include employees who were selected by a contractor for group membership based on the expectation that each would offer unique skills to the team. The task each member completes is often conducted independently of the other group members, and the group members are not expected to learn the entire set of “jigsaw pieces,” as they are in peer-learning tasks.

With the temporary group approach to student group work, students likely enjoy the opportunity to specialize, building the skills that interest them and avoiding the tasks that do not. Under the best conditions, this specialization of labor enables the students to reduce their overall workload, adds to the efficiency of the team, and creates a synergy that leads to higher group performance. As long as a high-quality, low-cost group output is the main goal, as it is in the workplace, this best-case student group scenario works fine.

However, the main objective in business education is education, not necessarily the production of high-quality group output. Therefore, the highly functional group described above may be dysfunctional in an educational setting. The specialization of labor may amount to little more than "collaborative loafing" (Bacon et al., 1998) wherein students divide up the work in a manner that allows them to apply the skills they are already comfortable with and thereby find the path of least resistance to a satisfactory grade. Bacon et al. (1998) hypothesized that the tendency to parcel out the workload as finely as possible but as much as necessary to meet the instructor's expectations for the project explains why, among the teams they studied, no increase in project quality was observed as the team size increased. Thus, student group projects may enable students to feel they have accomplished something significant, as the whole will generally be greater than any one individual's effort, and to feel that they have gained something close to the hallowed "real-world experience." The feelings of accomplishment evident in the students may conceal from the instructor the fact that each student has actually learned very little new material or new analysis tools and instead only rehearsed previous learnings.

Table 1 shows a summary of the characteristics of peer learning, the typical business school group project, and the typical workplace group project as discussed in this review. The table shows that student projects are somewhat "stuck in the middle" and as such may neither produce much content learning nor accurately emulate the real world. The ideal peer-learning environment has individual and group rewards. Project groups in the workplace also have individual and group rewards. By contributing effectively to a workplace team, the members build their own reputations and social capital among their team members, which can be leveraged later for further success or advancement. In addition, as a member of a successful workplace team, those outside the team are likely to think more highly of all the team members, further contributing to the reputations of team members.

As mentioned earlier, however, student projects typically have only group rewards. In terms of process, the peer-learning environment emphasizes students learning as a group, while the school project and business environments emphasize some initial coordination to divide up the tasks; however,

TABLE 1
A Comparison of Group Experiences

Group Characteristic	*Peer-learning Group*	*Business Student Project Group*	*Traditional Real-World Project Group*
Reward level	Individual and group	Primarily group	Individual and group
Process	Learning through preparation and teaching	Coordinate independent efforts	Coordinate independent efforts
Organizational goal	Learning content	Learning content, building team skills	Achieving strong project performance

most of the work is done independently. This process works well for the business project, because the organizational goals will be accomplished if the project is performed well and as efficiently as possible, regardless of how much anyone learns. With school projects, however, the organizational goal (the school's goal) is student learning, just as it is in peer learning, and not necessarily project performance. Thus, instructors as project designers find themselves in an ineffective middle ground. The more that business school projects imitate real-world projects, the more they move away from the ideal format for peer learning, and the lower is the likelihood that content-learning objectives will be achieved. It should also be noted that goal incongruence may further limit the effectiveness of school projects. Students may see their objective as "satisfying the boss to get a reward, like in the real world," whereas the instructors' objectives will generally be related to learning outcomes.

The presence of social loafing and specialization of labor means that not all the students will perform all the tasks, and so most students will miss some learning opportunities. However, will this loss of learning opportunities be more than compensated for by the benefits of peer learning? Do business students learn more when they work in student groups because of the opportunity to teach and learn among themselves, or do they learn less because of the opportunity for free riding or collaborative loafing? Because of the lack of individual accountability generally present in student group projects, we suspect that peer learning does not occur to a significant degree and that instead student learning of project-related content is diminished by working in groups. Therefore, the major hypothesis in the present research is:

Hypothesis 1: Business students will learn more project-related content when working on a project alone than they will when working in groups.

Of course, a great deal of learning in groups may be in an outcomes domain that is separate from the course content. In group projects, students ideally learn the team functioning skills and leadership skills that will enable them to achieve greater performance on workplace teams. Unfortunately, there is little evidence that this type of learning occurs in the typical student team project. For example, Bacon et al. (1999) found that although students indicated they learned about group functioning from at least some group experiences, no significant relationship was found between the quality of student team functioning and how far team members had progressed in an MBA program. The contradiction between what students said they learned and how they actually behaved on teams would caution against the use of self-reports of learning effectiveness as they relate to group learning (e.g., Mundell & Pennarola, 1999). Others (Bolton, 1999; Bowen, 1998) suggested that more direct teaching of team skills, and more feedback regarding team skills, must be integrated in the typical student project to achieve meaningful progress in team skill development. However, instruction in team skills rarely occurs in the context of student group projects (Bolton, 1999).

Method

The current study utilized an experimental design to determine whether students learn more when completing a project on a group or individual basis. Participants were undergraduate business students in seven sections of a consumer behavior course that were taught by the same instructor over seven academic quarters. During the first five quarters, a group project was assigned wherein students worked in pairs. In the last two quarters, the students were required to complete the project alone. The students from the first five quarters comprise a "group-project group" and the students from the last two quarters comprise an "individual-project group." The literature reviewed previously provides evidence that a group size of two is adequate to achieve peer learning, as long as the learning task is appropriately designed.

Measures of learning outcomes were routinely collected on a midterm exam about a week after the project was submitted. By identifying which questions on the exam covered project-related content and which covered non-project-related content, it is possible to test for differences in project-related content while controlling for extraneous error due to differences in non-project-related content. Rasch scaling is used to facilitate the comparison of two different forms of the midterm exam.

THE PROJECT

In the first five quarters, the (group) project required students to interview two people (participants) who had recently made a major purchase and to prepare a written analysis of the decision-making processes their participants used to make the purchase decision. Because there were two members on each team, students were encouraged to interview one subject each. Students were required to follow a detailed project description and to submit a paper limited to 12 pages in length. The project description specified five sections: introduction, problem recognition, information search, alternative evaluation, and postpurchase processes. The project description also specified concepts and tools to be used in each section, such as drawing a perceptual map, applying one of five decision rules, or analyzing the extent to which each of four different factors affected postpurchase dissonance. A 23-item scoring sheet was included in the project description so that students could see exactly how their projects would be graded. The detailed project description made the distinction between project-related content (material specifically mentioned in the project description) and non-project-related content (other material covered in the course) straightforward. The students were asked to integrate the relevant content from each of the interviews in each section, and form a conclusion and recommendation for each section, thus encouraging peer learning.

As one can see from this project description, the content-learning outcomes for the project included several levels of Bloom’s (1956) taxonomy. For example, to complete the decision-rule aspect of the assignment, the students first must understand each of the decision rules and know how they differ, reflecting learning outcomes at the lowest level of Bloom’s hierarchy, knowledge and comprehension. The ability to apply the appropriate decision rule in a novel situation reflects a higher level of content learning, the application level of Bloom’s taxonomy.

In the last two quarters, students were only required to interview one person, however the outline remained essentially unchanged, and the page limit was dropped to 6 pages. Thus, the total effort required was cut approximately in half, while the required content remained the same.

SAMPLE

The samples studied here comprise traditional undergraduate students who took the same marketing course (Consumer Behavior) in the Winter 2000, Spring 2000, Winter 2001, Spring 2001, Fall 2001, Winter 2002, and Spring 2002 quarters. All sections of the course were taught by the same instructor, who had taught this course for more than 10 years by the start of the Winter 2000 quarter. Thus, the content, structure, and delivery of the course was very consistent across quarters. Most (56%) of these 233 students were 3rd-year business majors, while 14% were 2nd-year students, and 30% were 4th-year students. Just more than one half (53%) of the sample were women.

INSTRUMENTS

The critical measures in the current study are drawn from the midterm exam given in the class routinely 1 week after the project due date. The instructor identified two groups of questions on the exam: those that covered project-related content (PRC) and those that covered non-project-related content (NPRC). These two groups of questions formed subtests (T_{PRC}, T_{NPRC}) that can then be calibrated to form Rasch measures (M_{PRC}, M_{NPRC}). In the measurement paradigm that most readers are familiar with, a student's score on a test is often computed by simply summing the number of correct responses. This paradigm is more formally known as classical test theory (CTT). Rasch measures are analogous to test scores in CTT but are derived differently (see the Appendix for additional explanation of Rasch measurement). The key dependent variable in this study is M_{PRC}, the measure of learning of project-related content.

Two versions of the midterm exam were used over the seven quarters of study. During the first four quarters, the midterm contained 43 multiple-choice and 13 short-answer questions. The format of the exam was changed to include only multiple-choice questions (65 items) for the last three quarters. The format change was made in response to research that has shown that multiple-choice tests can be as reliable and valid as short-answer tests and can be even more reliable on a reliability-per-hour basis (Bacon, 2003). Some of the short-answer questions were rewritten as multiple-choice questions in the process of revising the exam. The second version of the exam contains 41 items that also appeared on the first version. One of the major advantages of using Rasch scaling in the current research is the ability to "equate" the two test forms used here, mapping each test score into the same underlying latent scale.

The Rasch measures were formed using BIGSTEPS (Wright & Linacre, 1998). The "simple procedure" (Bond & Fox, 2001, p. 54) for test equating was used because of the similarity of the test versions and substantial overlap of test items. The M_{PRC} and M_{NPRC} measures were found to have acceptable reliability (.61 and .78, respectively). Furthermore, the item infit statistics (information weighted fit statistics) for the items constituting each measure were found to be consistent with the rules of thumb presented by Wright, Linacre, Gustafson and Martin-Löf (1994) (M_{PRC} infit M = 1.0, SD = .12, range = .78 to 1.21; M_{NPRC} infit M = 1.01, SD = .12, range = .74 to 1.34). Only 1.7% of the students achieved a perfect score on the M_{PRC}, and none scored perfectly on the M_{NPRC}, meaning nearly all the students could be measured by the items. The person fit statistics showed wider than expected variance (M_{PRC} infit M = 1.0, SD = .31, range = .34 to 2.22; M_{NPRC} infit M = 1.07, SD = .37, range = .46 to 3.14). In this research, it is assumed that all the student responses, although perhaps reflecting some overconfidence, guessing, or gaps in knowledge, are within the bounds of normal student behaviors. Therefore, all students, including those with fit scores outside the statistically expected norms, were retained in the current study.

ANALYSIS

The differences in project-related content-learning outcomes (M_{PRC}) across the two experimental groups were compared in an ANCOVA analysis. Using M_{NPRC} as a covariate controls for any extraneous differences in student learning across experimental conditions and also reduces the error variance (Reichardt, 1979; Wildt & Ahtola, 1978). Bloom (1976) offered considerable evidence that academic achievement is strongly predicted by earlier academic achievement (correlations of .70 or so are commonly observed, see p. 68). In the current research, the measure of non-project-related-content learning (M_{NPRC}) reflects the student's current academic achievement level. This measure was entered first in the model as a covariate because it has the strongest theoretical and empirical support. The "treatment" variable of interest here, group versus individual project, was then entered as a main effect in the ANCOVA. Finally, gender and a gender-by-treatment interaction was entered in the model and tested in an exploratory fashion. Previous research has suggested that women may be more motivated students than men (Bacon & Novotny, 2002), which may affect the degree to which they benefit from project experiences. Shaw et al. (1999) advocated controlling for the effects of individual differences, including gender, to increase the rigor of studies of educational effectiveness.

Results

The results of the ANCOVA analysis are shown in Table 2. As shown in the table, the covariate, M_{NPRC}, had the strongest effect on project-related content outcomes. The simple correlation between M_{NPRC} and M_{PRC} was .66 (n = 233, $p < .001$). The main effect of group- versus individual-project assignment was found to be significant ($p = .002$), supporting Hypothesis 1, that students learned more project-related content when working individually than they did when working in groups. An examination of cell means determined that the average increase in PRC learning relative to NPRC learning that occurred when the project was changed from group completed to individually completed was .38 standard deviation units. This can be interpreted as a moderate effect size according to Cohen (1992).

The exploratory examination of gender as a main effect in project learning is also shown to be significant in Table 2 ($p = .024$). The interaction between individual versus group and gender was not found to be significant ($p = .764$). An examination of the cell means revealed that the average PRC – NPRC difference for women was .12 while the average PRC – NPRC difference for men was only .05. Thus, the women learned project-related content better than the men did regardless of whether they completed the project on their own or in a group.[1]

FOLLOW-UP RESEARCH

One quarter after completing the data collection for this research, the instructor had students complete the project working individually. After the project was submitted, students completed a short survey to determine student opinions about the project redesign. Of the students who responded (27 of 32), 44% believed they would have learned less if they had completed it with someone else, 52% believed they would have learned about the same, and 4% believed they would have learned more working with someone else. Thus, these student perceptions of learning are consistent with the study's primary findings. In terms of enjoyment, two thirds (67%) of the students indicated they would have preferred working alone to working with someone else, and over three fourths (78%) said it would have been more work to complete the project in pairs. After reflecting on these results, the students were given the option of completing the second project assigned in the class, designed to be an extension of the first, as a 9-page individual project or a 15-page group project (up to 4 students per team). Nearly three fourths of the class (23 of 32) chose to work individually. In an MBA class given similar options, the entire class (20 students) chose to work individually. Thus, among the projects studied here, it appears as though most students learn more, enjoy themselves more, and work less when working on a solo project compared to a team project of proportional length.

TABLE 2
ANCOVA for Project-Related Content Learning

Source of Variance	*Sum of Squares*	df	F	p *Value*
Covariate				
Non-project-related content	75.01	1	184.09	< .001
Main effects				
Individual vs. group	3.99	1	9.80	.002
Gender	2.09	1	5.13	.024
Interaction				
(Individual vs. Group) × (Gender)	.04	1	.09	.764
Model	81.14	4	49.78	< .001
Residual	92.90	228		
Total	174.04	232		

Discussion and Conclusion

The results of the current study indicate that students learned more project-related content when they worked alone rather than when they worked as a member of a group. Although the use of an experimental design for the current study gives some assurance of a causal relation between learning and working alone versus working as a member of a group, the design did not offer a detailed explanation of the causal process. Anecdotal evidence and the follow-up study offer some support for each of three possible processes suggested in the literature review. First, there was some evidence of social loafing. Casual conversations with students indicated that some students did indeed slack off when working in pairs, and this free riding, so commonly observed in student teams, likely reduced the learning of some students.

Specialization of labor was the second factor that likely reduced the amount of learning in groups. Although students working in pairs were encouraged to interview one student each, a few conversations with students about their group process revealed this process was not always followed (e.g., “He did the interviews and I did the writing,” or “She did the first three sections, and I did the last two”). Thus, this project, similar to many business school group projects, may well have offered a way for students to “specialize,” employing a subset of skills in each group task and neglecting the development of additional skills. For some projects, these skills may be those the students entered the program with (e.g., Excel skills, or presentation skills). Thus, these students may progress through a degree program without adding much to their personal skill set. The use of teams in business education may explain why some students who complete an MBA appear to be missing some basic business analysis tools, such as the ability to calculate the net present value of a stream of cash payments (Pfeffer & Fong, 2002). In designing a team-based learning task, the instructor should carefully consider which learning outcomes should be achieved by all students, and which need only be achieved by some students.

The issue of which outcomes should be achieved by all and which should be achieved by some should be explicitly addressed at the program level as well. When students choose for themselves what to learn and what not to learn, teachers in higher-level courses may be frustrated by the gaps in student knowledge. Many students will lack prerequisite skills because of their self-guided specialization (cf., cognitive entry behaviors; Bloom, 1976) and so will have great difficulty developing the new, more advanced skills required in higher level classes.

The third factor that likely reduced the learning in groups in the current study was coordination losses. This effect was reflected in the follow-up studies wherein students rated a group project as being more work, even though the work was theoretically no more than the individual project. The logistical challenges in finding a time to meet and integrating disparate parts of a project raise the cost of group work. To find a time for the group to meet, group members invariably must compromise, giving up other activities, making special trips, or making other sacrifices that they would not have to make if they worked on their own. To the extent that the time spent meeting with the group requires giving up some other valuable opportunity, a group project involves higher opportunity costs than does a solo project even if the time-on-task is exactly the same. Given the necessity of integration and compromise inherent in group projects, the time-on-task may actually be higher in group work, other things being equal. The presence of these coordination losses and increased opportunity costs reduce the learning efficiency of group projects. Although increased individual accountability could be used to ameliorate social loafing and specialization of labor, increased individual accountability would not be expected to improve coordination losses. One area for future study is the degree to which the allocation of class time for student group-project work enhances the learning of project-related content.

The fact that gender had a small but significant main effect suggests that women learn more from projects (individual or group) than men do. Project work may be a more effective learning task for women at this research site because the women are generally higher in academic motivation, as measured by achievement striving (Bacon & Novotny, 2002). This finding suggests that project-based learning may generally be more effective for more highly motivated students. The effect of gender or motivation on project learning represents an important issue for future research.

The observation that NPRC learning outcomes explained 43% of the variance in PRC learning outcomes, while the effect of working on the project in a group or as an individual explained only an additional 2% of the variance, can be seen two ways. One might conclude that for only a 2% increase in explained variance, the project may make so little difference that it is not worth doing in any format. We should consider, however, that the large explained variance because of NPRC learning outcomes likely reflects cognitive entry behaviors over which the teacher has little control; that is, students enter a class with substantial differences in educational background, prior subject interest, and general academic motivation, and the teacher may not be able to profoundly affect these characteristics. However, of the variance that remains, doing the project as an individual, as opposed to as a member of a group, did have some effect. Here, the project explained 4% of the variance remaining after controlling for NPRC (from Table 2, 3.99/[174.04 – 75.01]). Viewed in this way, perhaps the teacher should be happy with such meager gains and continue to use the project.

It is important to note that the negative impact that group projects have on learning may be greater than that observed here for at least two reasons. First, in the course studied here, all the course content and the material on the exam was discussed in class or available in the textbook. The project was designed to enhance learning by getting students to review and strengthen their knowledge while applying it in a real-world context, however the project was not the sole source of learning the material. In a course where the project is the sole source of learning for a significant amount of course content, it is expected that the group-project versus individual-project learning differences will be greater. Second, in the current study, the group project was specifically designed to be completed as a pair. By its design, the project encouraged each student to be an equal partner on the team. If the project had not been so easily divided, or if the project team were larger, it seems more likely that students would free ride or specialize, making the difference between group-project learning and individual-project learning even greater. Furthermore, if the benefits of peer learning stem mainly from the differences in studying when students expect to teach others (Bargh & Schul, 1980), then a larger group size would lead to a finer division of labor, narrowing the scope of the material that each student learns in depth, and thus further reducing the average amount learned as group size increases. A similar argument is made by McKeachie (1990) in support of learning cells, which have a team size no greater than two.

In conclusion, the current study found evidence that students learn more course content through individual projects than through group projects. This finding directly contradicts what many feel is the main purpose for using student groups—to enhance learning. Management course designers should make explicit decisions about whether a group project is intended to enhance content learning (in which case they should imitate peer-learning models) or enhance learning about teams (in which case they should include team exercises). It is not obvious from the current research that projects that are designed to closely model workplace experiences accomplish either learning objective. Through more careful group-project design, and the redesign of ineffective group projects as individual projects, business schools may be able to facilitate more learning while improving student satisfaction.

Appendix
Rasch Measurement

This research uses Rasch scaling, one of the item-response theory (IRT) models (a one-parameter IRT model). Rasch, building on earlier contributions by measurement researchers such as Thurstone and Thorndike from the 1920s, developed the fundamental mathematics of what is now called Rasch measurement in the 1950s and 1960s (Wright & Stone, 1979, for a more complete history). Because of the computational intensity of Rasch measurement, the tool did not gain wider acceptance until recent decades when faster computers and more efficient algorithms made Rasch measurement a more practical tool. IRT is now the tool of choice for many educational measurement researchers and is very commonly used in journals such as *Educational and Psychological Measurement* and *Journal of Educational Measurement*. The popularity and scholarly interest in this tool has led some journals to focus exclusively on Rasch and IRT, including the *Journal of Outcomes Assessment*.

ADVANTAGES OF RASCH MEASUREMENT

Rasch analysis was used in the current study for two reasons. First, Rasch analysis creates interval-level measures, as opposed to ordinal-level measures typically obtained with classical test theory (CTT) (see, for example, Nunnally, 1978). Rasch obtains interval-level measurement because it explicitly and independently estimates the difficulty of each question and the ability of each person. To offer a simple example, consider a test with only three questions, two of which are very easy and one is very hard. Applying CTT, we would compute an assessment of each student's knowledge by summing the number of correct answers. Thus, students could score a 1, 2, or 3. Note, however, that because two of the questions are very easy, scores of 1 and 2 reflect fairly similar levels of knowledge, while a score of 3 reflects a much higher level of knowledge. Thus, this summed scale is not truly an interval measure of knowledge. Rasch measurement generates estimates of the difficulty of each question and uses these to estimate the ability of each person. Although the mathematics behind the Rasch model are beyond the intended scope of this article (interested readers might start with Baker, 2001, or Wright & Stone, 1979), the method produces assessments of knowledge or ability that are interval scaled, measured in units called "logits."

The second reason for using Rasch analysis in this research is the ease with which Rasch can equate different tests. One way to think of IRT is that each item response is used to estimate a person's ability, not the sum of item responses (Baker, 2001, p. 6). Thus, a test with three items would be expected to generate about the same estimate of a person's ability (measured in logits) as a test with 300 items, although the longer test would generate estimates with much lower errors and higher reliability. This means that a person's ability can be measured on the same latent-interval scale with different versions of a test, as long as the versions can be meaningfully linked because they share some items or some people took both versions. The ability to compare learning outcomes from one class to the next, even if the test versions have changed slightly, offers a tremendous benefit to teacher-researchers who study their own classes and wish to be able to compare learning outcomes over time periods covering different evolutions of the assessment instruments.

Note

1. The results using Rasch measurement were compared with the results that might have been obtained using CTT. The Rasch measures generally correlated highly with the CTT measures (around .98), as would be expected because Rasch measures are a nonlinear transformation of CTT measures. In statistical tests, the CTT measures showed differences in the same direction; however, the differences did not achieve statistical significance. The lack of significance is mainly due to the fact that the use of CTT measures limited test equating, and thus fewer items or persons can be used in the analysis, reducing the statistical power.

References

Ausubel, D. E. (2000). *The acquisition and retention of knowledge: A cognitive view.* Norwell, MA: Kluwer Academic.

Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short-answer questions in a marketing context. *Journal of Marketing Education*, *25*, 31-36.

Bacon, D. R., & Novotny, J. (2002). Exploring achievement striving as a moderator of the grading leniency effect. *Journal of Marketing Education*, *24*, 4-14.

Bacon, D. R., Stewart, K. A., & Anderson, E. S. (2001). Methods of assigning players to teams: A review and novel approach. *Simulation and Gaming*, *32*, 6-17.

Bacon, D. R., Stewart, K. A., & Silver, W. S. (1999). Lessons from the best and worst student team experiences: How a teacher can make the difference. *Journal of Management Education*, *23*, 467-488.

Bacon, D. R., Stewart, K. A., & Stewart-Belle, S. (1998). Exploring predictors of student team project performance. *Journal of Marketing Education*, *20*, 63-71.

Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation. Retrieved from http://ericae.net/irt/baker/

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology*, *72*, 593-604.

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook I: Cognitive domain*. New York: Longman, Green.

Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.

Bolton, M. K. (1999). The role of coaching in student teams: A "just-in-time" approach to learning. *Journal of Management Education*, *23*, 233-250.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.

Bowen, D. D. (1998). Team frames: The multiple realities of the team. *Journal of Management Education*, *22*, 95-103.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Comer, D. R. (1995). A model of social loafing in real work groups. *Human Relations*, *48*, 647-667.

Deeter-Schmelz, D. R., Kennedy, K. N., & Ramsey, R. P. (2002). Enriching our understanding of student team effectiveness. *Journal of Marketing Education*, *24*, 114-124.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, *53*(3), 497-509.

Goldschmid, M. L. (1971). The learning cell: An instructional innovation. *Learning and Development*, *2*, 1-6.

Gosenpud, J. J., & Washbush, J. B. (1991). Predicting simulation performance: Differences between individuals and groups. In W. J. Wheatley & J. Gosenpud (Eds.), *Developments in business simulation and experiential exercises* (pp. 44-48). Stillwater, OK: Association for Business Simulation and Experiential Learning.

Herman, E. D., Keldsen, S. Z., & Miller, J. G. (2001). *States of development in collaborative learning: Administrative perspectives on teaming in MBA programs*. Boston: Boston University School of Management Center for Team Learning.

Hernandez, S. A. (2002). Team learning in a marketing principles course: Cooperative structures that facilitate active learning and higher level thinking. *Journal of Marketing Education*, *24*, 73-85.

Kahn, W. (1995). Group process checkpoints for team learning in the classroom. *Journal of Policy Analysis and Management*, *14*, 310-326.

Kimber, D. (1996). Collaborative learning in management education: Issues, benefits, problems and solutions: A literature review. *ultiBASE*. Retrieved from http://ultibase.rmit.edu.au/Articles/june96/kimbe1.htm

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*, 822-832.

Mallinger, M. (1998). Maintaining control in the classroom by giving up control. *Journal of Management Education*, *22*, 472-483.

McKeachie, W. J. (1990). Research on college teaching: The historical background. *Journal of Educational Psychology*, *82*, 189-200.

McKeachie, W. J. (1994). *Teaching tips: Strategies, research, and theory for college and university teachers* (9th ed.). Lexington, MA: D. C. Heath.

McKeachie, W. J., Pintrich, P. R., Lin, Y., Smith, D. A. F., & Sharma, R. (1990). *Teaching and learning in the college classroom: A review of the research literature* (2nd ed.). Ann Arbor, MI: National Center for Research to Improve Post Secondary Teaching and Learning.

Mesch, D. J. (1991). The jigsaw technique: A way to establish individual accountability in group work. *Journal of Management Education*, *15*, 355-358.

Meyerson, D., Weick, K. E., & Kramer, R. M. (1996). Swift trust and temporary groups. In R. E. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 166-195). Thousand Oaks, CA: Sage.

Michaelsen, L. K., & Black, R. H. (1994). Building learning teams: The key to harnessing the power of small groups in higher education. In S. Kadel & J. Keehner (Eds.), *Collaborative learning: A sourcebook for higher education* (Vol. 2, pp. 65-81). State College, PA: National Center for Teaching Learning and Assessment.

Mundell, B., & Pennarola, F. (1999). Shifting paradigms in management education: What happens when we take groups seriously? *Journal of Management Education*, *23*, 663-683.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Pfeffer, J., & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. *Academy of Management Learning and Education*, *1*, 78-95.

Reichardt, C. S. (1979). The statistical analysis of data from nonequivalent group designs. In T. D. Cook & D. T. Campbell (Eds.), *Quasi-experimentation: Design & analysis issues for field settings* (pp. 147-205). Boston: Houghton Mifflin.

Schomberg, S.F. (1986, April). *Involving high ability students in learning groups*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Shaw, J. B., Fisher, C. D., & Southey, G. N. (1999). Evaluating organizational behavior teaching innovations: More rigorous designs, more relevant criteria, and an example. *Journal of Management Education*, *23*, 509-536.

Slavin, R. E. (1988). Cooperative learning and student achievement. *Educational Leadership*, *46*, 31-33.

Strijbos, J. W. (2000). A classification model for group-based learning. *European Journal of Open and Distance Learning*. Retrieved from http://kurs.nks.no/eurodl/eurodlen/index.html

Strong, J. T., & Anderson, R. E. (1990). Free-riding in group projects: Control mechanisms and preliminary data. *Journal of Marketing Education*, *12*, 61-67.

Webb, N. J., & Grib, T. F. (1967). *Teaching process as a learning experience: The experimental use of student-led groups* (Final Report, HE-000-882). Washington, DC: Department of Health, Education, and Welfare.

Wildt, A. R., & Ahtola, O. T. (1978). *Analysis of covariance* (Sage University Paper series on Quantitative Applications in the Social Sciences 07-012). Beverly Hills, CA; London: Sage.

Williams, D. L., Beard, J. D., & Rymer, J. (1991). Team projects: Achieving their full potential. *Journal of Marketing Education*, *13*, 45-53.

Wright, B. D., & Linacre, J. M. (1998). *BIGSTEPS computer program for Rasch measurement* (version 2.82), Chicago: MESA Press. Available from www.winsteps.com/bigsteps.htm

Wright, B. D., Linacre, J. M., Gustafson, J. E., & Martin-Löf, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*, 370.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.