

Collaborative Team Testing To Support Individual Learning: Can Teamwork Motivate Learning?

Leigh Stelzer, Seton Hall University, USA
Joan Coll-Reilly, Seton Hall University, USA

ABSTRACT

A challenge in the contemporary classroom is that many students do not prepare for class. While technology is a boon in the classroom it can often be an apathy-fostering distraction. To encourage greater student preparation a course was designed with 4 quizzes to be taken first as individuals and then as members of predetermined teams. We reasoned that students knowing they would be quizzed on the chapter material would be motivated to prepare for the quiz on the assigned day. It was hypothesized that team spirit would lead students to higher levels of performance. We reasoned that students, believing other members of the team were depending on them, would be motivated to prepare for the quiz. While a majority showed no improvement, a minority who expressed team spirit did show gains.

Keywords: Teams; team power; business school students; classroom distraction; technology in the classroom; Team Learning; cohort learning; learning motivation

INTRODUCTION

The power of the team comes from the potential to combine resources, specialize and divide the work, and build motivating camaraderie. As a result teams will perform better than the individual members of the team acting alone. Studies support the proposition that teams can indeed outperform their individual members (Weldon & Weingart, 1993; West, 2004).

Managers use teams in the belief that workers acting in concert can achieve more than they can when acting alone (Alie, Beam, & Carey, 1998; Gerben, Van Der Vegt, Bunderson, & Oosterhof, 2006; Glassop, 2002; Shaw & Schneier, 1995). Although the strength of teams often derives from division of labor, there is the potential for cross learning, teaching and mentoring. Teams also socialize recruits and increase their skill levels. Members of teams teach and learn from each other (Haller, Gallagher, Weldon, & Felder, 2000).

Teaming also has the potential to create social bonds that support the task (Ammeter & Dukerich, 2002, Steinbrink & Jones, 1993). Group cohesiveness that supports prosocial values can build up over time (Tan & Tan, 2008). Members of the team develop felt responsibility to support one another's efforts. Ideally, teams exhibit mutual commitment and accountability (Bryant & Albring, 2006).

The problem in the contemporary classroom is that often students are disinclined to read assigned material in preparation for class (Editorial Board, 2008). When and if students study, many wait until the last minute when a test is imminent. Additionally, students often complain that the tests cover too much material. Finally, the students sitting behind their laptop screens have a tendency to lose track of the professor's presentation of the material. In sum, student apathy or a larger problem of social disconnectedness are hindering learning and teaching (Baker, Comer, & Martinak, 2008; Jassawalla, Sashittal, & Malshe, 2009).

Technology has become a constant distraction in today's classroom. Gina Trapani (2005), a writer on technology, quotes an article in the New York Times. "In the era of e-mail, instant messaging, Googling, e-commerce and iTunes, potential distractions while seated at a computer are not only ever-present but very enticing... there is a universe of diversions to buy, hear, watch and forward, which makes focusing on a task all the more challenging." Crawford (2005) has applied the term "borgs" to describe people who exhibit these behaviors. Commenting on the apathy associated with student borgs, Dave Taylor (2003), a widely published expert and teacher on internet and business issues, expresses his chagrin with student lack of attention and concludes that faculty can lecture to students but they can't make them learn.

These considerations led us to think about using teams to shake up the classroom for improved performance. A pilot study was initiated to see if the power of the team can be harnessed to motivate individual students to improve their performance and increase their learning (Metheny & Metheny, 1997; Nowak, Miller, & Washburn, 1996; Steinbrink & Jones, 1993). Students would work in teams to complete quizzes that assess their learning. Will working in teams lead students to increase their individual performance and learn more than they would learn individually?

RESEARCH DESIGN

A course was designed to include four individual chapter quizzes each of which was followed immediately by a team quiz with identical questions (Nzeogwu, 1997). The quizzing plan was premised on a belief that students could learn better in small bites, a chapter at a time. Knowing they would be quizzed on the chapter material, students would be motivated to read the chapter and prepare for the quiz on the assigned day. Additionally, believing other members of the team were depending on them, students would be motivated to prepare for the quiz. Team spirit, group pressures or feelings of shared responsibility would stimulate their individual motivation (LaFasto, 1989; Tan & Tan, 2008). Students who experienced that they were the weak contributors to team success would increase their preparation and over time increase their contribution to the team. Clark (Nowak et. al., 1996) argued that team testing stimulates students' sense of responsibility. Finally, students who initially did not learn the material on their own by individual study would learn from other students and, thus, take advantage of the pooling of knowledge (Nzeogwu, 1997). We tested the following hypotheses:

Hypotheses

Hypothesis 1a: Team quiz scores will exceed the quiz scores of individual team members.

Hypothesis 1b: The team quiz score average will exceed the averaged quiz scores of individual team members.

Hypothesis 2a: Individual quiz scores will improve over time.

Hypothesis 2b: Individual quiz scores on every later quiz will improve over every earlier quiz.

Hypothesis 2c: Individual quiz scores on every subsequent quiz will improve on the immediate preceding quiz.

Hypothesis 2c: Individual quiz scores on the final quiz will improve on the first quiz.

Hypothesis 2d: Individual quiz scores on the final quiz will improve on the average of the first three quizzes.

Data Collection

Six teams of 5 students each were formed in the first days of the semester on the basis of alphabetical order of student last names. This selection procedure, though not a systematic random selection, is quite close. There was no reason to believe the teams were anything but heterogeneous on variables such as knowledge, intelligence and motivation. Many of the students in the teams were strangers to each other. Some students in the class had prior contact with each other from being in other classes but the selection process did not affect the likelihood of them being on the same team.

The students took four quizzes worth a total of about 35% of the course grade during a mid-semester two month period. The quiz was the first order of business on the assigned dates. After the thirty students were seated, the professor posted the individual quiz on Blackboard. 1 Students were allowed to take the quiz only once. Blackboard withdraws the quiz when it is completed or when time expires.

We have no reason to believe that any quiz was more difficult than another. Ideally, for the purpose of this study we would prefer an objective basis for saying that the quizzes were the same difficulty level. Mean grades and median grades for the quizzes differed, as did the first and third quartiles. The following are the means, medians and first and third quartiles respectively; Q1= (53.3%, 54.5%, 42.0%, 62.5%); Q2= (58.9%, 63.6%, 45.5%, 72,7%); Q3=(51.1% , 54.2%, 41.7%, 58.3%); Q4= (50.0%, 50.0%, 30.0%, 62.5%). Grades were noticeably higher on the 2nd quiz and lowest on the 4th quiz. When we look at the results of the team test however, the third quiz garnered the highest average and the 4th quiz was 2nd highest: Means = 65%, 72%, 86%, and 74%.

Efforts were made to deter collaboration among students. Blackboard allows random ordering of questions for every use of the quiz. Another option allows only one question to appear at a time. The instructor walked around the room to reduce the students' inclination to share answers, to look at others' screens, or to consult notes. The number of questions on the quizzes varied between 10 and 12. Most students completed the quiz within 10 to 12 minutes.

When they were done with the individual quiz, students moved to be in their teams and took a posted team version with the same questions as the individual quiz. The order of the questions was again randomized. As teams came together, they identified which student's computer was being used for the team quiz as the designated team computer. Again, students were monitored to deter cheating. Several times students not on the designated team computer had to be reminded to close their laptops. There surely was some cheating amidst the cacophony of team work. However, would it be too cynical to suggest that as likely as not these students who opened their laptops were checking their IMs or email rather than trolling for correct quiz answers? After all, they had been "off the grid" for at least 10 minutes.

The analysis reports on the results of individual quizzes, team quizzes and an on-line follow-up student questionnaire. The data are organized as follows:

- A. Individual quiz results. N=116 quizzes. (4 absences.)
 - 1. Scores are the number of answers correct.
 - 2. Percents are the percent of answers correct.
- B. Individual average. N= 30. Based on average of 4 individual quizzes for each student (3 quizzes for the 4 students absent 1 time)
- C. Team individual quiz average. N=24. Based on averaged results of the individual scores of the 5 team members for each quiz; 6 teams taking 4 quizzes.
- D. Team quiz score. N= 24 quizzes; 6 teams taking 4 quizzes.

FINDINGS

Hypothesis 1

The first hypothesis says that teams will do better than the individuals. There were 116 possible comparisons of individuals' quiz grades with their team quiz scores. The team scored more points than the individuals on the team in 101 (87%) instances. There were only four cases of an individual doing better than her/his team.

How much better did the teams do? When the team quiz scores were compared with the four-quiz averages of the students on their individual quizzes, the differences ranged from .4, essentially no gain to the student from teaming, to 5 points, a gain of 50%. When team scores on the quizzes are compared to the team individual quiz average on the comparable quiz, the average gain over the 24 comparisons was 2.4 points. When the gains for each team were averaged – the difference between the team individual quiz average and the team quiz score - the range was 1.7 to 3.7 points. Thus the students in one team gained on average 3.7 points per quiz by taking the team quiz.

From an alternative perspective of percent correct answers, the average gain from teaming over the course of four quizzes was 22%. Table 1 illustrates that the gains to each team from working in teams for the combined four quizzes ranged from 16% to 35%. On the first quiz, team members working together scored on average 30% better than their five member average. The gain from teamwork fell to 16% on the second and third quizzes but rose to 27.6% on the fourth quiz.

Table 1
Team Gains (Percentage) From Taking The Team Quiz

Team	Q1 Gain	Q2 Gain	Q3 Gain	Q4 Gain	Avg. Gain
Team 1	17.1%	10.9%	13.3%	22.5%	16.0%
Team 2	41.4%	10.9%	16.7%	36.0%	26.2%
Team 3	49.6%	36.4%	21.7%	34.0%	35.4%
Team 4	19.7%	7.3%	25.0%	16.0%	17.0%
Team 5	37.1%	10.9%	13.3%	36.0%	24.3%
Team 6	24.5%	27.3%	3.3%	25.0%	20.0%
Average, All	30.2%	30.2%	16.1%	15.6%	27.6%

A brief illustration of the actual gains from teaming is in order. The point value of the quizzes varied between 10 and 12 points. On the first quiz worth 10 points, one team’s individual quiz average was 50% and its team quiz score was 100%. On the second quiz one team went from a team individual quiz average of 55% to 91% on the team quiz, from a failing grade to an A.

Team quizzing worked for the best as well as the worst individually performing students though the gain from teaming benefitted the worst performers more. When we compared the average number of points scored for each individual with his/her team score, not one student outperformed the team over the run of 4 quizzes. When we compared the students’ individual quiz averages with their team quiz score averages, students with the best individual averages on every team gained points from teamwork. While the best performing individual student gained on average less than a point from teaming, other highest scoring individual team members averaged gains as high as 1.7 points from teamwork. Worst scoring individuals gained much more from teamwork. The average gain of the worst performing students ranged from 2.3 to 5.4 points. Table 2 illustrates the percentage gains from teaming for the best and worst performing students on all six teams. The worst students on each team averaged a gain of 32.4% while the best students averaged a gain of 11.6%

Table 2
Difference between Students’ Grades (Four Quiz Average Percent Correct) and Team Grade
For Worst and Best Performing Students on All Six Teams

	Gain to Worst for Each Team	Gain to Best For Each Team
	22.5%	17.9%
	23.9%	14.1%
	26.1%	11.4%
	35.0%	10.6%
	36.8%	8.9%
	50.3%	6.9%
Average Gain	32.4%	11.6%

When we did a t test to test the significance of the difference of taking the test as individuals in comparison to taking the test as a team, the differences were significant at .05 for every one of the six teams. Working as a team to answer quiz questions made a significant difference for every quiz. An examination of Table 3’s comparison of the mean scores illustrates the consistently huge gains to the teams from taking a team quiz.

Hypothesis 2

The second hypothesis says that individual team members will improve over time their performance on the quizzes. Confirmation would be for students to score higher grades on later individual quizzes. The average percent correct on the first quiz was 52 percent. The average increased to 59 percent on the second quiz but fell back to 51 and 50 percent on the third and fourth quizzes respectively.

While there does not appear to be any improvement in the aggregated performance of individual students, perhaps there is individual improvement. Did individual students improve their quiz performance over time and over

the range of 4 quizzes? There are 6 possible comparisons for each student where we can look for a pattern of improvement: $Q2 > Q1$, $Q3 > Q2$, $Q4 > Q3$, $Q3 > Q1$, $Q4 > Q1$, $Q4 > Q2$. There are a possible 168 comparisons. (With 30 students there are 180 opportunities to improve performance. Each of the 4 absences eliminates 3 comparisons per student, 12 all together.) In total there were 74 (44%) instances when students improved their performance and 90 backslides when student performance decreased over time. (There were four instances of no change. Total 168.) Hypothesis 2a is not confirmed. Students did not improve over every earlier quiz.

Table 3
Team Quiz Scores Exceed the Averaged Quiz Scores of Team Members, Four Quizzes

	Mean	Variance	Observations	Pearson Correlation	df	t Stat	P(T<=t) one-tail	P(T<=t) two-tail
Team1 Ind Q Avg.	0.49	0.002	4	0.73	3	-6.322	0.004	0.008
Team1 Quiz	0.65	0.005	4					
Team2 Ind Q Avg.	0.45	0.014	4	0.05	3	-3.562	0.019	0.037
Team2 Quiz	0.72	0.009	4					
Team3 Ind Q Avg.	0.51	0.001	4	0.1	3	-6.186	0.004	0.008
Team3 Quiz	0.86	0.013	4					
Team4 Ind Q Avg.	0.57	0.004	4	0.33	3	-4.556	0.01	0.019
Team 4 Quiz	0.74	0.004	4					
Team5 Ind Q Avg.	0.58	0.013	4	0.61	3	-3.441	0.021	0.041
Team5 Quiz	0.83	0.032	4					
Team6 Ind Q Avg.	0.58	0.003	4	0.6	3	-3.578	0.019	0.037
Team6 Quiz	0.78	0.018	4					
Significant at .05								

Monotonic increases in quiz scores may be too high a standard to expect of students. Also, it penalizes early success on the quizzes. Indeed, only one student showed a monotonically increasing score over the 4 quizzes and thus had 6 increases when scores were compared. Hypothesis 2b confines the test to three sequential comparisons, $Q2 > Q1$, $Q3 > Q2$, $Q4 > Q3$.

Using the percentage of answers correct to compare quiz results, Table 4 does show that students improved their performance between Quiz 1 and Quiz 2. Eighteen (64%) of the 28 comparisons (one student was absent for the first quiz and one was absent for the second) showed improvement. From then, however, it was mostly downhill, in line with the aggregated percentages. Only 9 (31%) students improved from Quiz 2 to Quiz 3. Fourteen (50%) improved from Quiz 3 to Quiz 4. Thus, while there was some improvement from quiz to quiz among individuals, forty-four (52%) of the possible 85 comparisons (5 missing) failed to show improvement. Hypothesis 2b is not confirmed.

Table 4
Proportion of Students Who Improved Their Grade Compared to Those Who Regressed Between Sequential Quizzes, Between First and Last and Between Average of First Three and Last Quiz

	Improved	Regressed	No Change	Total
Q2>Q1	64.3%	35.7%	0.0%	28
Q3>Q2	31.0%	72.4%	0.0%	29
Q4>Q3	50.0%	46.4%	3.6%	28
Q4>Q1	33.3%	63.0%	3.7%	27
Q4>Avg. 1st Three Quizzes	37.0%	63.0%	0.0%	27

Our thinking was that social pressure for performance would build up over time. If so, improved performance should come more at the end than at the beginning. When we looked at the aggregate results we saw just the reverse, improvement on Quiz 2 and regressing on Quiz 3 and Quiz 4. Quiz 4 had the lowest average percentage. If social pressure was being exerted or felt over a period of time, we might expect to see a surge in individual performance on the last test. Looking at individual performance, if we just compare the first quiz with the

last quiz (27 cases, 2 missed last quiz and 1 missed the first quiz) there are only 9 (33%) improvements. If we compare the last quiz against the average for the first three quizzes, only ten students (37%) improved against their average. Again, there is no sign of improved performance, no surge at the end that might suggest that students felt social pressure to perform. Hypotheses 2c and 2d are not confirmed.

Table 5
Comparison of Individual Quiz Scores (Percentage Correct) Between Sequential Quizzes,
Between First and Last Quizzes and Between Average of First Three and Last Quiz

	Variables	Mean Quiz Grade	Variance	Observations	Pearson Correlation	Hypothesized Mean Difference	df	t Stat	P(T<=t) one-tail	P(T<=t) two-tail
Q2>Q1	Quiz 1	0.52	0.020	28	0.1643	0	27	-1.796	0.042	0.084
	Quiz 2	0.59	0.023	28						
Q3>Q2	Quiz 2	0.59	0.022	29	-0.101	0	28	1.7978	0.042	0.083
	Quiz 3	0.51	0.030	29						
Q4>Q3	Quiz 3	0.51	0.032	28	0.2976	0	27	0.2008	0.421	0.842
	Quiz 4	0.50	0.047	28						
Q4>Q1	Quiz 1	0.54	0.022	27	0.1037	0	26	0.6555	0.259	0.518
	Quiz 4	0.51	0.047	27						
Q2>Q1	Significant at the .05 level, 2nd quiz is higher mean.									
Q3>Q2	Significant at the .05 level, but mean score is down, not up									
Q4>Q3	Not significant. Mean score goes down, not up									
Q4>Q1	Not significant. Mean score goes down, not up									

Finally, we did a t test to test for changes from earlier to later quizzes (Table 5). We thought team spirit would develop and get stronger over time and lead students to do better over time. Only the comparison of Q1 and2 was in the hypothesized direction and showed significance at the 05. level. All the other comparisons were in the wrong direction. Student performance regressed after Quiz 2. The t test confirms rejection of Hypothesis 2. There was no improvement over time.

It might be argued that there is a ceiling effect on improvement. A student cannot improve on 100%. But as we see grades were far from 100%. The median individual quiz grades for the first three quizzes were 53%, 64% and 54% respectively. There was a lot of room for improvement. And indeed, more often than not the worst performer improved. There were no repeat worst performances from Quiz 1 to Quiz 2. However, there were a total of six repeats subsequently.

In summary there were some students who raised the bar but there were more examples of backsliding than improving. Neither at the aggregate nor at the individual level is there evidence that teaming encouraged a burst of individual improved performance or learning. To the contrary, students may have engaged in social loafing.

Social loafing is doing less than your best while relying on others to pick up the slack (Brandyberry & Bakke, 2006; Latane, Williams, & Harkins, 1979; Tan & Tan, 2008). Free riding, Brandyberry & Bakke (2006) write “is the attempt to benefit from team membership without bearing a proportional share of the costs. These individuals believe their efforts won't be identified or are dispensable.” Recently Jassawalla et al. (2009) argued that social loafing is more complex than a conscious choice to free ride. They reported that their student subjects identified the non-performer as a student who, “does not care and is socially disconnected from the team members”.

An indicator of social loafing, disconnectedness, or free riding could be scoring lowest among one's teammates on the individual quiz. Sixteen students had the lowest score on their team one or more times. Eighteen had the lowest score if we count ties. Of course, someone had to score lowest. Theoretically twenty-four different students could have scored lowest (more if we count ties). Six students scored lowest twice, five of them twice in a row (7 if we include ties for worst). Only one scored lowest more than twice, again if we include ties. Ten scored lowest one time if we include ties. Only eleven students never had a lowest score (including ties for lowest).

Another indicator of social loafing may be a comparison of the quiz grades of the 1st and 3rd quartiles. The grade of the 1st quartile dropped to 30% on the last quiz while the grade of the 3rd quartile was comparable to the Q1 and higher than Q3. The score on the team quiz was second highest. This looks like the classic free rider example of a few students dropping out and not pulling their weight while the best push on.

Low scoring was not consistent. No student scored lowest on both Quiz 1 and Quiz 2. From Quiz 2 to Quiz 3, there were 4 of 6 instances where lowest scorers were the same (includes one tie). From Quiz 3 to Quiz 4, there were 3 of 6 instances where the low scorer stayed lowest (includes one tie). Bottom students, however, did not always stay on the bottom. Ten students had both a highest and a lowest score, counting ties.

The volatility and inconsistency in these numbers suggest the possibility of considerable social loafing of the free-riding choice variety. From six to eighteen students at a minimum may have engaged in social loafing from the perspective that they scored low on a quiz. Additionally, there could have been any number of students who loafed but still scored better than the worst. The inconsistent performance does not appear to support the idea that the poor performers were disconnected in any permanent sense, although this may be the case for a few. Rather, poor performance may have been an occasional opportunistic choice reflecting poor time management or study skills. A student who considered himself a “strong teammate” expressed admiration for his team because “Whenever someone was slacking off, there was always someone else there to fill his place” (Student, 2009).

Inconsistent performance may also be explained by the operation of the “sucker effect.” The sucker effect “occurs when team members scale back their efforts to match those of the free riders” Brandyberry & Bakke (2006). Tolerance for filling in for others may be finite (Jones, 1984). That the mean score (59%) and the 3rd quartile score (72.7%) peaked on the second quiz may reflect reduced tolerance for the free riders on the subsequent quizzes.

WHAT DID STUDENTS SAY?

Twenty one of 30 students voluntarily responded to an on-line questionnaire at the end of the term asking them about their preparation and interaction with regard to the quizzes. See appendix available from authors. To see if there was any difference between those who answered the survey and those who did not, we compared the students’ midterm and final exam results. The midterm average for both groups was 7.2. The final test averages were 7.8 and 7.5. These results suggest that the ability level of those who took the survey was similar to that of the students who did not respond.

The purpose of the quizzing was to encourage greater preparation than was customary. When asked if they prepared, only 7 (33%) students said they prepared for all the quizzes. Twelve (57%) said they prepared for some. Two students said they prepared a little.

The goal had been for students to prepare by reading the relevant chapter in advance of the quiz. When asked how they prepared, only one student responded that he/she always read the chapter, while 3 said they often read the chapter. This left ten (47%) who said they sometimes read the chapter and 7 (33%) who never read the chapter. If they did not read the relevant chapter, perhaps they scanned it? Eight (38%) of the twenty-one often or always scanned the chapter and 10 said they sometimes scanned the chapter.

Students had other support material available, including PowerPoint slides and lecture notes on Blackboard. There was also a text website that contained a lecture, a quiz, a vocabulary list and other study aids. All of the students said they accessed the PowerPoint slides to some degree, 9 always and 7 often. Six (28%) always or often studied the internet quiz. In an open ended follow-up question 4 said they read the lecture notes posted on Blackboard.

In addition to encouraging greater preparation, the quiz design sought to invoke peer pressure to get more study preparation. Seven (33%) of twenty-one said they did not feel any peer pressure. Of the 14 (66%) who did feel pressure, 9 said sometimes, while only 3 said often and 2 said always. Thus, peer pressure was a continuing or major factor for a minority (5 = 24%) of students.

The students who said they often or always felt peer pressure to prepare did better on the quizzes. Their average quiz grade was 61% compared to 50% for those students who sometimes felt pressure and 53% for those who denied feeling pressure. The difference was amplified on the fourth quiz. Averages were 72%, 44% and 49% respectively.

Perhaps pressure was internalized? We asked the students if they cared if they let down their teammates or if they were pretty much indifferent (Tan & Tan 2008). Only 4 (19%) said they cared a lot. Fourteen (66%) said they cared “somewhat” and 3 said they did not care. Caring and feeling peer pressure were highly correlated at $r=.55$. Further, students who cared scored higher on the quizzes, with an average quiz grade of 57% compared to 52% for those who said they somewhat cared and 53% for those who did not care.

Thus, to the extent that students developed or responded to team camaraderie – they felt peer pressure and they cared about the team- the course design worked. The problem is that so few had team spirit. Given that few students appear to have felt strong social pressures nor to have cared a lot about letting down their teammates, maybe the testing procedure really encouraged free riding or, at best, division of labor. Asked directly if they were a “social loafer”, 13 said never but 7 (33%) said sometimes they were a loafer. Only one said he was often a loafer. When asked if there were one or two members of their team that they could rely on to prepare, only one student said that there was no one he could rely on. While in some sense this is a rousing endorsement of teaming, it may also suggest that the process actually reduced felt responsibility to prepare (Tan & Tan, 2008).

We did not expect students to immediately experience peer pressure to prepare for quizzes, but we thought cohesion, team spirit, and felt responsibility would build over time to support greater preparation. We asked if students saw any change over time in preparation. Twenty (95%) said the same students prepared each time. Respondents also saw some change. Thirteen (62%) said that some students prepared more over time. However, eleven (52%) students said there were students who prepared less over time.

Finally, did the quizzing help learning? Asked, did you experience any synergies, sixteen (76%) responded that they thought the group dynamics and discussion led to greater knowledge. Five respondents said they did not experience any synergies.

CONCLUSION

In the task of taking the quizzes, teams did far better than individuals. The differences between the team individual averages and the team quiz score were consistently significant. Hypothesis 1 was confirmed. Additionally, both the best and the worst individual performers were advantaged by working in teams, though the gains to the students who performed worst on the individual tests were much greater.

Our hypothesis that individual members would improve their performance over time proved false. When we examined the total number of changes from quiz to quiz, less than a majority were improvements. Comparing Quiz 4 to Quiz 1 only 9 students improved. Comparing Quiz 4 to their average for the first three quizzes, only ten students improved. These comparisons along with the t test failed to confirm Hypothesis 2 or any of its sub-hypotheses.

The great degree of volatility in the individual scores may be indicators of free riding, social loafing and the sucker effect. Nineteen of the thirty students had a lowest score and eleven had both a lowest and highest score. The uneven performance of so many leads one to wonder if some students engaged in occasional free riding as a form of division of labor, relying that others would pick up the slack. The low grades of the first quartile on the last quiz in comparison to the grade of the third quartile suggests social loafing. Further, it is reasonable to assume that some students would be more likely to engage in social loafing when they see others loafing at their expense (Jones 1984; Jones & George, 1998) though Liden et al. (2004) found the opposite.

It seems reasonable to conclude that there was a mix of social loafing and team spirit as students went from quiz to quiz. Social loafing may have been part of a consciously adopted mixed strategy on the part of individual team members or a form of division of labor. Individual students prepared for some quizzes and relied on others to prepare for other quizzes. Free riding may have increased as the semester progressed and work assignments accumulated. The sucker effect may also have played a role.

It appears that the social disconnectedness and apathy that we sought to overcome by having students work in teams was, at most, minimally affected by taking quizzes in teams. The survey results indicate that students who said they felt peer pressure and those who said they cared about the team did better on the quizzes than those who did not. Thus, those students who were affected appear to have performed as hypothesized. To the degree students took on prosocial values of felt responsibility, cohesiveness and caring about others there was superior performance (Tan & Tan, 2008). Their performance in combination with the survey response that students thought they learned as a result of the teams should be taken as a marker of some success.

We wonder if teaming in the standard class can generate sufficient social pressures to motivate students to prepare more for classes and quizzes. First, as with all class performances students must perceive a relationship between success on the quiz and course and some goal they wish to achieve. One class and a quiz or two are small episodes in their busy lives. Second, social bonds develop over time through continuous interaction. Given that the externally designated class teams are small episodes, will enough students identify with their teams and care what other members think to the point of not wanting to disappoint them? Even if students are motivated and care about their teams, circumstances often intervene in students' abilities to prepare. Students seldom have total control of their schedules. Finally, division of labor and relying on others while a form of shared responsibility may undermine the goal of learning.

If instructors are going to try to use the power of the team to promote learning they will first have to overcome these structural problems. A group of students does not make a team. Team work values need to be addressed head on. Instructors can teach team skills, promote the power of true teamwork and work to increase the penalties for free riding. Teamwork can be made a larger proportion of the grade to give it more salience. Confidence to use teams may increase knowing that while some students got greater advantage from teaming, no student was disadvantaged. Increasing the cohesion among students in the larger context of the school might also help. If students moved through the business school from class to class as cohorts there could be more opportunities to build prosocial values.

This was a pilot effort to see if teaming made a difference in student learning. There are limitations to this study. The analysis reflects data from one class in one semester. We cannot assert with certainty that the quizzes had exactly the same difficulty level. While it is in no way conclusive, the study does suggest that if instructors are going to try to harness the power of the team to overcome apathy and improve individual learning, they must first address the issue of creating viable teams. Teams with prosocial values could lead to learning improvements. Future studies of performance in the context of aggressive efforts to build team spirit could show results that eluded this course design effort.

ENDNOTE

1. Blackboard is a course delivery system that permits an instructor to post teaching materials and tests on the web for the exclusive use of students in the designated class.

AUTHOR INFORMATION

Dr. Coll-Reilly is an Organizational Behavior Psychologist and Professor of the Department of Management, at the Stillman School of Business, Seton Hall University. Her experience is in the fields of management, culture and business, stress management, and team building.

Dr. Coll-Reilly teaches both the undergraduate and graduate courses in organizational behavior and cultural issues in management.

Dr. Leigh Stelzer is an Associate Professor of Management and a former Chairman of the Department of Management at the Stillman School of Business, Seton Hall University. With the integration with his article "Learning and Teaching Management on the Web: What Do We Know?" (2003) <http://www.infoagepub.com/www/products/product3/wankelcombo.pdf> Additional research addressed issues of the match of technology to pedagogy and employment of technology for higher-order critical thinking skills.

REFERENCES

1. Alie, R. E., Beam, H. H., & Carey, T.A. (1998). The use of teams in an undergraduate management program. *Journal of Management Education*, 22(6): 707- 720.
2. Ammeter, A. P. & Dukerich, J. M. (2002). Leadership, team building, and team member characteristics in high performance project teams. *Engineering Management Journal*, 14 (4): 3-11.
3. Baker, S., Comer, D., & Martinak, M. (2008). All I'm askin' is for a little respect: How can we promote civility in our classrooms? *Organization Management Journal*, 5(2): 65-80.
4. Brandyberry, A., & Bakke, S. (2006). Mitigating Negative Behaviors in Student Project Teams: An Information Technology Solution. *Journal of Information Systems Education*, 17(2): 195-209.
5. Bryant, S. M. & Albring, S. M. (2006). Effective Team Building: Guidance for Accounting Educators. *Issues in Accounting Education*, 21(3): 241-26.
6. Crawford, W. (2005). The coming of the borgs. *EContent*, 28(5): 43.
7. Editorial Board. (Nov. 2008). Why our education is falling behind; lazy students. Minnesota Daily. Retrieved on June 12, 2009, from <http://www.mndaily.com/blogs/unfit-print/2008/11/11/why-our-education-falling-behind-lazy-students>.
8. Gerben S., Van Der Vegt, J., Bunderson, S., & Oosterhof, A. (2006). Expertness diversity and interpersonal helping in teams: Why those who need the most help end up getting the least. *Academy of Management Journal*, 49(5): 877-893.
9. Glassop, L. I. (Feb., 2002) The organizational benefit of teams. *Human Relations*, 55(2): 225-250.
10. Haller, C. H., Gallagher, V. J., Weldon, T. L. & Felder, R. M. (Jul 2000). Dynamics of peer education in cooperative learning workgroups. *Journal of Engineering Education*, 89(3): 285-294.
11. Jassawalla, A., Sashittal, H., & Malshe, A. (2009). Students' perceptions of social loafing: Its antecedents and consequences in undergraduate business classroom teams. *Academy of Management Learning & Education*, 8(1): 42-54.
12. Jones, G. R. (1984). Task visibility, free riding, and shirking: Explaining the effect of structure and technology on employee behavior. *The Academy of Management Review*, (9)4: 684-696.
13. Jones, G. R. & George, J. M. (1998). The experience and evolution of trust: Implications for cooperation and teamwork. *The Academy of Management Review*, (23)3: 531-547.
14. LaFasto F. M. J. (1989). *Teamwork: what must go right, what can go wrong*. Sage.
15. Latane, B., Williams, K., & Harkins, S. (1979). Many hands make light work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37: 822-832.
16. Liden, R., Wayne, S. J., Jaworski, R. A., & Bennett, N. (2004). Social loafing: A field investigation. *Journal of Management*, 30(2): 285-311.
17. Metheny, D., & Metheny, W. (1997). Enriching technical courses with learning teams. *College Teaching*, 45(1): 32-35.
18. Nowak, L I, Miller, S. W., & Washburn, J. H. (1996). Team testing increases performance. *Journal of Education for Business*, 71(5): 253-256
19. Nzeogwu, O. (1997). Collaborative learning: group examinations. Retrieved May 18, 2009 from http://66.102.1.104/scholar?hl=en&lr=&q=cache:PxCj55d1IG_OJ:www.sbaer.uca.edu/research/TemporarilyDisabed----wdsi/1997/pdf/136.pdf+Nzeogwu,+Okeleke++
20. Steinbrink, J. E., & Jones, R. M. (1993). Cooperative test-review teams improve student achievement. *The Clearing House*, 66: 307311.
21. Shaw, D., & Schneier, C. (1995). Team measurement and rewards: How some companies are getting it right. HR. *Human Resource Planning*. New York: 1995. 18(3): 34-50.
22. Student. (2009). Private communication.
23. Tan, H., & Tan, M. (2008). Organizational Citizenship Behavior and Social Loafing: The Role of Personality, Motives, and Contextual Factors. *The Journal of Psychology*, 142(1): 89-108.
24. Taylor, D. (May 8, 2003) Lazy students, a rant. The Business Blog at Intuitive.com. Retrieved May 31, 2009, from http://www.intuitive.com/blog/lazy_students_a_rant.html.
25. Trapani, G. (Feb. 11, 2005) On computer distraction. Lifehacker.com. Retrieved May 31, 2009, from <http://lifehacker.com/software/productivity/on-computer-distraction-032865.php>.
26. Weldon, E., & Weingart, L.R. (1993). Group goals and group performance. *British Journal of Social Psychology*, 32: 307-334.
27. West, M. A. (2004). *Effective teamwork, practical lessons from organizational research*. Wiley-Blackwell.