

## **Diversity, Effort, and Cooperation in Team Based Learning**

Molly Espey  
John E. Walker Department of Economics  
Clemson University  
Clemson, SC 29634  
864-656-6401  
[mespey@clemson.edu](mailto:mespey@clemson.edu)

December 2015

## **Diversity, Effort, and Cooperation in Team Based Learning**

### **Abstract**

Student and team performance in sixteen sections of an introductory microeconomic theory course taught using team-based learning are analyzed to determine what measurable characteristics of teams influence outcomes. Team success on in-class quizzes and activities and team cohesiveness are modeled as a function of diversity of class level, gender, and geographic origin, previous academic performance, and team size and class size. Individual performance is estimated as a function of individual characteristics and team characteristics, including team performance and team cohesiveness. Results suggest that team performance is not significantly influenced by team size, class size, or the mix of class levels. The GPA of the top individual on the team is positively correlated with overall team success. Interestingly, team cohesiveness is not significantly impacted by team success. Further individual success is positively and significantly influenced by team cohesiveness but not by overall team performance.

Constructive use of teams in the classroom may enhance student success both in the classroom and in a work environment. The National Association of Colleges and Employers annual survey of employers consistently finds “teamwork skills” and the ability to work with others among the top five qualities employers most desire in employees. Good teamwork requires effective communication, regular interaction, mutual respect and trust. Successful use of teams in the classroom can improve student motivation (Denton 1994; Dommeyer 1986), improve communication skills (Meyer 1994), and increase subject matter knowledge and critical thinking skills (Nichols 2002; Espey 2012).

Unfortunately, students’ most common experience working with others in academic settings is either in group activities with groups changing from activity to activity or in groups assigned for one-time projects, with the work primarily conducted outside class time. Such group activities are vulnerable to the pitfalls of free riding and social loafing and may also encounter problems associated with personality differences and inexperience with conflict resolution. Team-based learning (TBL) is an instructional strategy designed to minimize the influence of many of these potential problems, in part by retaining the same teams throughout the course. Within the context of TBL, groups are transformed into teams, ideally comprised of a small number of people with complementary skills working cooperatively to achieve a common goal and hold themselves mutually accountable (Michaelsen et al 2002). What constitutes “complementary skills”, however, is not always apparent. Team and individual success in the TBL classroom could be enhanced by quantitatively identifying characteristics of teams that contribute to such success. Increased success arguably correlates positively with improved attitudes and ability to work with others in subsequent activities, be they in the classroom or in a work environment.

This study reviews the literature on team composition and team effectiveness, provides an overview of TBL, then presents the analysis of ten years' worth of TBL in introductory microeconomic theory classes to determine specific characteristics of the composition of teams that contribute to team success, team cohesiveness, and individual success.

## **BACKGROUND**

An extensive literature exists on the value of collaborative learning for enhancing individual learning, as well as the relationship between team composition and team success, but there is a paucity of research on the relationship between team composition and individual learning. The literature related to team composition and team effectiveness tends to fall into two general categories: teamwork in organizations and teamwork in specific course activities. Within the organizational behavior literature, team effectiveness is commonly modeled using an input-process-output framework (Cummings 1981; Hackman 1987; Salas et al 1992) where various input factors such as features of the group, its task, and the work environment influence the group interaction process, ultimately influencing the group output (Hackman 1987). In a work setting, individuals are typically assigned to teams based on availability and competencies relevant to the task at hand. Occasionally, members may volunteer to participate and it may also be possible for individuals to opt out of certain team assignments. In the classroom however, all students must participate and will generally face the same task and work environment, at least within a given class. Nonetheless, a variety of individual level, team level, and environmental factors may influence team effectiveness (Hackman 1987; McGrath 1964; Yeatts and Hyten 1998).

Surface level variables are overt demographic characteristics such as age, race, education level, and gender. Demographic heterogeneity may lead to differences in team performance (Bunderson and Sutcliffe 2002; Pelled et al 1999), but meta-analysis has not found a consistent, statistically significant relationship between such heterogeneity and team performance (Webber and Donahue 2001). Deep level variables that may influence team performance include personality factors, values, abilities, and attitudes toward the task at hand and have generally been found to have a more significant impact on team success over time (Bell 2007). However, these characteristics are typically much more difficult to measure than demographic variables and previous empirical results have been inconsistent in terms of which variables matter most (Bell 2007), with only general mental ability (GMA) having been found in two meta-analyses as a strong predictor (Devine and Phillips 2001; Stewart 2006).

The influence of team size has also been investigated with mixed results. Team size is posited to influence the distribution of participation and the nature of group interactions (Dawe 1934; Miller 1951), with smaller teams providing more opportunity to interact, yet other research has found that larger groups perform at higher levels (Campion et al 1993). Cossé et al (1999) found team size to be positively correlated with team performance in a marketing class, but Deeter-Schmelz et al (2011) did not find size to be significant in their study of marketing classes. Class size does not seem to have been studied in terms of its influence on individual team performance in the classroom.

## **TEAM BASED LEARNING**

Team-based learning (TBL) attempts to foster effective group interaction by keeping students in the same group throughout the semester and utilizing collaborative activities in each

class period. In terms of team effectiveness, the TBL literature tends to focus on team size and regular interaction as keys to team success, emphasizing team sizes of 5 to 7 students and daily in-class interaction (Michaelsen et al 2002; Michaelsen and Sweet 2008). In such a context, “teams” become distinct from and more effective than “groups.” Over time, as students begin to trust each other and develop a commitment to the group, the group becomes a team (Michaelsen et al, 2002).

Sweet and Michaelsen (2011) further describe four pieces of “the practical framework of TBL” as proper teams, the readiness assessment process, 4-S application exercises, and student peer evaluations. In addition to being permanent teams of 5 to 7 students, proper teams are strategically formed by the instructor to provide a balance of resources across teams; this may be done by surveying students on the first day of class regarding prior experience relevant to the class, previous coursework, or any other features that may contribute to success in the specific course. The readiness assessment process involves beginning-of-unit readiness assessment tests (RATs) over assigned reading which are taken individually, then again by teams. This holds individuals accountable for doing the assigned readings and teams accountable for working together to ensure students understand the basic concepts well enough to begin applying that knowledge. 4-S application activities are significant, involve specific choices, include the same problem for all students, and require simultaneous reporting. Significant problems engage students in concrete examples so they understand the usefulness of the course concepts. Specific choices require teams to take a position, sometimes also requiring them to support that position with a short rationale for their choice. Forcing all students to confront the same problem enables them to better engage with each other across teams, while simultaneous reporting precludes

teams from simply agreeing with the majority of others, forcing them to decide before knowing what other groups will say.

The final critical piece of TBL is student peer evaluation. In the typical TBL class, student grades are based on both individual work and teamwork. Graded teamwork includes both the team RATs and application exercises, although the number of graded versus ungraded application activities will vary, as will the weight given to each component of the grade. How peer evaluation is implemented across instructors also varies, but it will generally involve both quantitative and qualitative evaluation, be anonymous, and include a mid-semester formative assessment, as well as end-of-semester summative assessment. This peer evaluation will factor into student grades either as a third component or used to weight the team portion of the grade.

TBL is designed to address many of the pitfalls of group work. Free riding and social loafing can be minimized through effective use of mid-semester formative peer evaluations and end-of-semester summative peer evaluations. Personality or cultural differences and inexperience with conflict resolution can be mitigated over time, with team performance generally increasing with hours of interaction for at least 30 hours (Watson et al 1993), close to the length of time students have to work together during a regular course. Nonetheless, differences in academic capabilities, attitudes, and effort, and demographic heterogeneity may still play an important role in influencing team cohesiveness and team performance.

## **DATA**

Team performances of 106 teams, comprised of a total of 642 students in sixteen sections of introductory microeconomic theory taught between 2006 and 2015 at a four-year public university, were analyzed to determine what observable characteristics of teams appear to

influence team success. Surface-level variables available for each student include gender, class level, major, and hometown and state. Deep-level variables include grade point average and end-of-semester peer evaluations for each student. Environmental variables include team size and class size<sup>1</sup>. Table 1 summarizes team level data and Table 2 individual data.

About 39 percent of the students were female, just over a quarter were from out of state, and two-thirds were freshmen or sophomores, as might be expected for an introductory level course. The use of major as an explanatory variable in the regression proved somewhat problematic. In most cases it was not possible to discern whether or not students were economics majors, or intended to be, as prior to completion of the business core such students are listed as pre-business majors, making it impossible to use this information to accurately account for the number of economics majors on each team. Thus any attempt to measure the impact of having more economics majors on a team was not statistically significant; as this was not a reliable measure of majors, it was ultimately omitted from the final analysis.

Diversity indices are created for gender and class level using a Herfindahl index such that diversity is measured as one minus the sum of squared proportions such that a higher value indicates greater diversity. For gender, this index can range from 0 for a team with no gender diversity to 0.5 for a team that is 50 percent male and 50 percent female. The class level diversity index can range from 0 for a team for which all members are at the same class level, to 0.75 for a team evenly balanced in terms of freshman, sophomores, juniors, and seniors. Geographic diversity was accounted for by the percentage of students on the team from out-of-state. While 27 percent of the students overall were from out-of-state, and teams varied from 0 to 83 percent out-of-state students, only eight of the 106 teams had more than half from out-of-

---

<sup>1</sup> Semester and year controls were also tested but were generally found not to be statistically significant.



state, so a geographic diversity index using in-state versus out-of-state measures was not significantly different from diversity measured simply using the percent of students from out-of-state.

Several different measures of grades were used in the analysis: the highest individual's grade point average (GPA) for each team to determine how much influence one person might have on the outcome, the average GPA of all the individuals on the team, and the highest individual RAT score for each team for the model of team RAT scores. Interestingly, the correlation between the best students' GPA and the maximum individual RAT score for each team was only 0.21. For the analysis of individual final exam scores, dummy variables are included for first semester freshmen (nearly twenty percent of the students) and transfer students (six percent overall), as GPA information is not available for these students.

Team grades include beginning-of-unit readiness assessment tests (RATs) and in-class team activities. Teams also worked on numerous ungraded activities daily. Over the course of the semester, teams took five RATs covering basic concepts related to the readings for each unit. These tests were taken first as individuals, then as teams. Each team also completed 12 to 13 graded in-class activities over the course of the semester. The overall team grade was a weighted average of these two components, 25-30% weight given to the RATs and 70-75% weight given to the activities. Individuals' grades were a weighted average of individual activities (homework and tests) and the team grade, weighted by the peer evaluations.

Peer evaluations completed at the end of the semester require students to rate their teammates in terms of their contribution to learning throughout the semester. Each student has points equal to ten times the number of teammates s/he has and must allocated those points to his/her teammates such that at least one score is differentiated, thus not everyone can receive a

ten. Evaluations range from 0 to 20, but typically fall between 7 and 12 inclusively, with 95% of all evaluations falling in this range. The lowest average any individual who completed the class received was 2.3 and the highest was 13.8. Only nineteen students received average peer evaluations below 7 and only fourteen received averages greater than 12. Peer evaluations were used in the analysis of individual grades as a measure of effort and engagement in the class. The negative of the variance in peer evaluations within each team is used to create a team cohesiveness variable, with a lower variance in evaluations assumed to correlate to a more cohesive team.

As TBL guidelines recommend that teams range from five to seven in size, all but two teams in this study fall in that range, with an overall average size of six. In both instances, students withdrew too far into the semester to make team size adjustments, resulting in teams of four. Classes ranged in size from 19 and 60 students in size, with an average class size of 43.

## **EMPIRICAL MODELS**

### **Team Success**

Team success on RATs, in-class team activities, and overall are modeled as a function of surface-level and deep-level variables, as well as team size and class size as follows:

- (1) Team grade = f(class level diversity, gender, geographic diversity, GPA, cohesiveness, team size, class size)

Class level diversity is estimated using the percentage of the team at each level as well as the diversity index. Gender is measured both as the percent of the team that is female as well

as by the gender diversity index. The percentage of each team from out-of-state is used to reflect geographic diversity. The influence of grades is estimated using the highest individual grade point average (GPA) of students on the team and the average GPA of all team members. For the team RAT scores, the maximum individual RAT score was found to be a significant explanatory variable as well.

The TBL literature recommends teams between 5 and 7 in size, suggesting that there are not enough resources for smaller teams to address complex application exercises while it is harder for all students to contribute substantially in larger teams. Given that all but two teams are within the range of TBL recommended team sizes, the team size variable is not expected to be statistically significant, yet finding this empirically is also valuable in that it would support the TBL literature. Class size is also not expected to be a significant factor influencing either peer evaluations or team success, as it is interactions within the teams, rather than between teams, that would be expected to matter most. However, since many potential adopters of TBL question their ability to use it in large or even medium size classes, this variable is included in the analysis.

### **Team Cohesiveness**

Team cohesiveness is also modeled as a function of surface-level and deep-level variables, as well as team size and class size:

$$(2) \quad \text{Cohesiveness} = f(\text{class level diversity, gender, geographic diversity, team success, team size, class size})$$

In general, the relationship between diversity and team cohesiveness, or lack thereof as reflected in greater variance in peer evaluations, is ambiguous, with theory suggesting that diversity brings more experiences and different perspectives to the table that can contribute to success, but also creates more differences and possible lack of harmony. Gender and class level diversity, both as percentages and as indices, are used to measure these aspects of diversity. The out-of-state variable was also included. While combining out-of-state and in-state students increases diversity, it is also possible that in-state students from different parts of the state may provide sufficient diversity.

In addition to surface-level diversity, variation in academic ability and/or effort as reflected in GPA might be expected to influence peer evaluations. A high GPA student would be expected to be a stronger than average contributor to a team, perhaps garnering significantly higher peer evaluations as a result. On the other hand, some high GPA individuals are initially less enthusiastic about TBL (Espey 2010), thus might be perceived as aloof or as less than ideal teammates by their peers. Low GPA students are more inclined to miss class or not be prepared when they do attend. Unfortunately, given the lack of grade information for first semester freshman and transfer students, a quarter of the students overall and more than fifty percent of many teams, variance in team GPA was not deemed to be a reliable measure. However, team success, whether stimulated by one strong student on the team or by several capable students, may make all the students on the team more enthusiastic about working hard for continued success, increasing team cohesiveness and reducing variation in peer evaluations. Thus the relationship between academic success and team cohesiveness was estimated using the highest GPA on the team, lowest GPA on the team, and the overall team score.

## **Individual Success**

Individual success in the class is measured by students' performance on the comprehensive final exam, expressed generally as:

$$(3) \quad \text{Final exam score} = f(\text{individual characteristics, team characteristics, environmental characteristics})$$

Individual characteristics include gender, class level, whether or not the student is a first semester freshman, whether or not the student is a transfer student, whether or not the student is from out of state, cumulative grade point average at the start of the course, and end-of-semester peer evaluations as a measure of the student's effort and engagement in the class. Team characteristics include class level diversity measured both as percentage of the team at each level and using the diversity index, gender measured both as the percent of the team that is female and using the diversity index, team size, average GPA of team members, overall team performance, and team cohesiveness. Class size is used as the control for the environment.

## **EMPIRICAL RESULTS**

### **Team Success**

Models were estimated for team RAT scores, in-class team application activities, and the overall team score. Results are shown in Table 3. In the model for team RAT and activity scores, percentages are used to measure class and gender first, then diversity indices are used. The overall team score is a combination of the RAT and activity scores, more heavily weighted

toward the activity scores (70 to 75 percent) and is shown for comparison, using the variables from the best fit for the model of activities.

The only variable estimated to be statistically significant at the five percent level or better in explaining RAT scores was the highest individual RAT score on the team, not surprisingly contributing to a higher team RAT score. The RATs are intended to test at the basic conceptual level and do not involve application or calculations, as they are given before any lecture or discussion of the material takes place in the class. Interestingly, the most academically successful students were not necessarily the highest scoring on the RATs and did not have a statistically significant impact on team RAT performance, but did have significant impact on the team activity scores. Class level diversity was positive and significant at the ten percent level, while gender diversity was negative and significant at the ten percent level. Team cohesiveness does not significantly impact team RAT scores, but as students work in the same teams throughout the semester, they likely learn quickly who on the team does well on the RATs and typically go with the majority vote when deciding on team answers for the RATs.

Neither the gender nor the class level diversity indices were statistically significant in explaining team activity scores, but there was a positive and significant correlation between the percent of the team being female and team success. This result corresponds with some of the sociological literature that finds that females invite more discussion and openness on teams, factors that are important for many of the application exercises that involve choosing “the best” among a set of options and those involving justification of answers when more than one choice might be correct.

Neither team size nor class size were significant. The lack of significance of team size suggests that teams of five to seven do not differ significantly in terms of the resource

availability, but the limited range of team sizes in this study precludes making conclusions about teams larger or smaller than this range. Similarly, class sizes ranged from 19 to 60, so the results are not necessarily generalizable outside of this range. Finally, team cohesiveness was not found to be statistically significant in any of the models of team performance.

### **Team Cohesiveness**

The results of three models estimated to explain the variance in peer evaluation scores are shown in Table 4. The percent female, gender diversity, and class level diversity are not statistically significantly different from zero. A larger percentage of freshman on a team correlates with statistically significantly higher measure of team cohesiveness, perhaps reflective of a greater desire to get along with others in a new situation, not just in the classroom but also in their first year at college.

Neither the maximum GPA nor the minimum GPA on a team was statistically significant in the model of team cohesiveness. In addition to the impact of individuals' academic ability on team cohesiveness, it is also of interest to know if greater team success contributes to greater team cohesiveness. The results of this analysis suggest overall team performance, however, does not have a significant impact on team cohesiveness. Sometimes teams get along well even when they do not perform particularly well on team activities, while other teams do well in spite of significant variance in contributions. Altogether, the results of this analysis suggest an explanation of team cohesiveness in terms of quantitative measures available to an instructor is elusive. Nonetheless, when it occurs in the TBL classroom, lack of cooperative effort in team activities is typically obvious and can often be addressed relatively easily, in or out of the classroom as appropriate. For example, instructors can encourage quiet students to speak up and

more talkative students to pause to allow others to provide input, ensure seating arrangements are conducive to every student participating, and otherwise facilitate communication.

### **Individual Success**

Comprehensive final exam scores are used to measure student success in the class. These scores are modeled as a function first of individual characteristics, then individual and team characteristics, and finally individual and team characteristics as well as team performance and team cohesiveness.

The estimated influence of individual characteristics is consistent across all models, as shown in Table 5. There is not a statistically significant difference in performance between males and females, or among freshman, sophomores, and juniors. Seniors averaged about 2.5 percentage points higher than other students. Out-of-state students scored on average about two percentage points higher than in-state students.

The coefficient on GPA implies about a 3.3 percentage point increase in final exam score per average grade level, with B students (3.0 GPA), for example, scoring 3.3 points higher out of 100 than C students (2.0 GPA) on the final exam and A students (4.0 GPA) scoring 6.6 points higher, on average, than C students, all else constant. Finally, the coefficient on students' average peer evaluation score indicates almost a three percentage point higher final exam score per unit change in the average peer evaluation. Given the structure of the peer evaluations, the average across a team is 10, so a student receiving an average of 9 on the peer evaluation would be expected to score about three percent lower, and a student earning an average of 11 on the peer evaluation score about three percent points higher than the average teammate.



The positive, significant coefficient on “first semester freshman” indicates that the average first semester freshman performs on the final exam at a level equivalent to a student with about a 2.60 GPA. In contrast, transfer students perform on average at a level equivalent to a student with a 1.65 GPA.

Among the team characteristics, only gender diversity and team cohesiveness were statistically significant at the five percent level or better. Gender diversity was positive and significant, suggesting students on teams with an equal number of men and women score about 3.5 percent higher on the final exam than students on all-male or all-female teams. However, the grade improvement for a student on a team of maximum gender diversity compared to the average gender diversity in this study is less than one percentage point.

Somewhat surprisingly, team performance does not have a statistically significant impact on individual performance in the course, whether estimated as the only team characteristic, with team cohesiveness, or with other team characteristics as in Model 5. In contrast, team cohesiveness is consistently positive and significant at the one percent level. The estimated value on this coefficient implies that students on a team in the average of lowest 25 percent for cohesiveness perform about two percentage points lower than students in the middle in terms of cohesiveness. Students on a team at the average of the top 25 percent for cohesiveness score about one half a percent better on the final exam.

In order to better understand the impact of teams and team composition on different types of students, Model 5 was estimated separately for males and females, for each class level, and for four groupings by GPA: below 2.5, 2.5 to 2.99, 3.0 to 3.49, and 3.5 to 4.0. These results are shown in Tables 6-8. GPA and peer evaluations are positive and significant predictors of individual success for nearly all of these subgroups. GPA is not a statistically significant

predictor of the performance of freshman, while peer evaluations are not a significant predictor of performance for seniors. Neither of these are particularly surprising; freshmen have limited college experience with first semester grades likely not as representative of their capabilities as GPA for other students, while seniors taking an introductory level course likely care less about team interactions than they do simply getting the class out of the way, and often miss classes for job interviews, particularly in the spring semester. Further, with only 78 seniors in the data set, few coefficients are statistically significant in the model for seniors.

First semester females are estimated to score higher than first semester males, but male transfer students are estimated to perform better than female transfer students, although there were only eleven female transfers and twenty-eight males. There were not enough transfer students to provide reliable estimates for the finer subdivisions of the data. Team gender diversity, measured using percent female or the index, does not have a significant impact on females' exam scores but a significant, positive impact on males' exam scores.

Team cohesiveness was positive and significant for both males and females but only for freshmen when the data was divided by class level. When divided by GPA, team cohesiveness was most positive and significant for the lowest GPA group (below 2.5), not statistically significant for the middle GPA groups, and positive and significant at the 10 percent level for students in the 3.5 to 4.0 GPA range. While overall team grade appeared to be a significant predictor of individual performance for the highest GPA group, there is likely an endogeneity issue in estimating the model for this particular group in that the vast majority of the students in this category have the highest GPA on their team and significantly influence the team performance. However, removing this variable from this model does not significantly change

any of the other coefficient estimates except team cohesiveness, which becomes larger and more statistically significant.

Within the chosen GPA ranges, GPA is only significant for the 3.0 to 3.49 group. On the other hand, the coefficient on peer evaluations is positive and significant for all four groups, with the largest estimated value for the highest GPA group. All students are predicted to perform better on the final exam if they are rated as better teammates by their peers, regardless of average academic performance, with those below a 3.0 GPA estimated to score about 2 percentage points better on the final exam for each one point higher average peer evaluation, and those with a GPA 3.5 or above estimated to score 3.3 percentage points higher for each one unit higher peer evaluation. In the context of team-based learning, students typically get rated lower for missing class often, not being prepared when attending, and/or not being actively involved with team activities and decision-making. Students are typically rated higher when they “know their stuff”, are engaged with the team, and/or help facilitate team activities. Regardless of academic standing, students who work better with others in the team-based learning environment actually appear to learn more in comparison with similarly situated students who don’t engage as actively or work as well with others.

## **CONCLUSIONS**

Awareness of characteristics of team characteristics that contribute to team and individual success in the classroom can help faculty members more carefully design teams to enhance outcomes and learning. For example, more effective teams will have members with complementary skills, but effectiveness also appears to be enhanced by team cohesiveness, thus

faculty should be attuned to divisions or conflicts on teams and attempt to alleviate such problems to the extent possible.

Explaining to students the value beyond the classroom of learning to work in teams helps students see the bigger picture. Businesses use teams for many projects; the literature addressing the team composition in organizations often addresses who *not* to put on a team, yet in the classroom, everyone must participate if an instructor chooses to use teams or groups. Knowing what team components matter and what don't, in terms of team and individual performances, can help instructors focus energy and attention on the appropriate variables in designing teams and working to maximize outcomes. If students learn to work better in teams, in addition to performing better in the classroom, they may improve their employability and work productivity as well.

## REFERENCES

- Bell, S.T. (2007). "Deep-Level Composition Variables as Predictors of Team Performance: A Meta-Analysis", *Journal of Applied Psychology*, 92(3): 596-615.
- Bunderson, J.S. and K.M. Sutcliffe (2002). "Comparing Alternative Conceptualizations of Functional Diversity on Management Teams: Process and Performance Effects," *Academy of Management Journal*, 45: 875-893.
- Campion, M.A., G.J. Medsker, and A.C. Higgs (1993). "Relations Between Work Groups Characteristics and Effectiveness: Implications for Designing Effective Work Groups," *Personal Psychology*, 46: 823-850.
- Cossé, T.J., D.N. Ashworth, and T.M. Weisenberger (1999). "The Effects of Team Size in a Marketing Simulation," *Journal of Marketing Theory and Practice*, 7(Summer): 98-106.
- Cummings, T.G. (1981). "Designing Effective Work Groups," *In Handbook of Organizational Design, Vol. 2*, edited by P.C. Nystrom and W.H. Starbuck. New York: Oxford University Press.
- Dawe, H.C. (1934). "The Influence of the Size of Kindergarten Group on Performance," *Child Development* 5: 295-303.
- Deeter-Schmeltz, D.R., K.N. Kennedy, and R.P. Ramsey (2002). "Enriching Our Understanding of Student Team Effectiveness", *Journal of Marketing Education*, 24(2): 114-124.
- Denton, H.G. (1994). "Simulating Design in the World of Industry and Commerce: Observations from a Series of Case Studies in the United Kingdom," *Journal of Technology Education*, 6(1): 1045-64.
- Devine, D.J. and J.L. Phillips (2001). "Do Smarter Teams Do Better: A Meta-analysis of Cognitive Ability and Team Performance," *Small Group Research*, 30: 678-711.
- Dommeyer, C.J. (1986). "A Comparison of the Individual Proposal and the Team Project in the Marketing Research Course," *Journal of Marketing Education*, 8(Spring): 30-38.
- Espey, M. (2010), "Valuing Teams: What Influences Student Attitudes?" *NACTA Journal*, 54(1):31-40.
- Hackman, J. R. (1987). "The Design of Work Teams," *In Handbook of Organizational Behavior*, edited by J.W. Lorsch. Englewood Cliffs, NJ: Prentice Hall.
- McGrath, J.E. (1964). *Social Psychology: A Brief Introduction*, New York: Holt.

Meyer, J. (1994). "Teaching teams through teams in communication courses: Letting structuration happen." Paper presented at the annual meeting of the Speech Communication Association, New Orleans.

Michaelsen, L.K., A.B. Knight, and L.D. Fink (2002). Team-Based Learning: A Transformative Use of Small Groups in College Teaching, Praeger Publishers, Westport, CT.

Michaelsen, L.K. and M. Sweet (2008). "Team-based learning: Small group learning's next big step," *New Directions in Teaching and Learning*, No. 116.

Miller, N.E. (1951). *The effect of group size on decision-making discussions*, Ph.D. dissertation, University of Michigan, Ann Arbor.

National Association of Colleges and Employers, <http://www.nacweb.org/>

Nichols (2002). "The Effects of Cooperative Learning on Student Achievement and Motivation in a High School Geometry Class," *Contemporary Educational Psychology*, 21(4): 467-476

Pelled, L.H., K.M. Eisenhart, and K.R. Xin (1999). "Exploring the black box: an analysis of work group diversity, conflict, and performance," *Administrative Science Quarterly*, 44(1): 1-28.

Salas, E., T.L. Dickinson, S.A. Converse, and S.I. Tannenbaum (1992). "Toward an understanding of team performance and training." In *Teams: Their Training and Performance*, edited by R.W. Swezey and E. Salas, 3-30. Norwood, NJ: Ablex.

Stewart, G.L. (2006). "A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance." *Journal of Management* 32: 29-55.

Sweet, M. and L.K. Michaelsen (2011). *Team-Based Learning in the Social Sciences and Humanities: Group Work That Works to Generate Critical Thinking and Engagement*, Stylus Publishing, LLC.

Warren E. Watson, W.E., K. Kumar and L.K. Michaelsen (1993). "Cultural Diversity's Impact on Interaction Process and Performance: Comparing Homogeneous and Diverse Task Groups." *The Academy of Management Journal*, 36(3): 590-602 .

Webber, S. S., and L.M. Donahue (2001). "Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis." *Journal of Management*, 27: 141-162.

Yeatts, D.E. and Hyten, C. (1998). *High-Performing Self-Managed Work Teams: A Comparison of Theory and Practice*. Thousand Oaks, CA: Sage.

Table 1: Summary Statistics for Teams (N=106)

	Mean	Standard deviation	Minimum	Maximum
Surface Level Variables:				
Female (%)	0.390	0.194	0	1.00
Gender diversity	0.40	0.12	0	0.5
Class Level (%):				
Freshmen	0.295	0.210	0	0.83
Sophomore	0.374	0.165	0	0.80
Junior	0.217	0.176	0	0.80
Senior	0.116	0.123	0	0.50
Class diversity	0.60	0.10	0.28	0.73
Out-of-state (%)	0.274	0.187	0	0.83
Deep Level Variables:				
Max GPA	3.64	0.27	2.70	4.00
Max iRAT	0.868	0.56	0.74	1.00
Average GPA	2.90	0.20	2.35	3.22
Lowest GPA	2.06	0.51	0.33	3.00
Team cohesiveness	-2.87	3.13	-20.66	-0.02
Environmental Variables:				
Team size	6.15	0.74	4.0	7.0
Class size	43.1	10.29	19.0	60.0

Table 2: Summary Statistics for Individuals (N=641)

	Mean	Standard deviation	Minimum	Maximum
Final Exam Score	0.80	0.106	0	1.00
First semester freshman	0.189	0.392	0	1
Transfer student	0.063	0.242	0	1
GPA	2.90	0.66	0.33	4.00
Peer Evaluation	9.92	1.26	2.3	13.8



Table 3: Team Performance (N=106)					
	RATs	RATs	Activities	Activities	Overall
% freshmen	-0.004 (0.011)		-0.056** (0.027)		-0.039* (0.020)
% juniors	-0.002 (0.012)		-0.044 (0.029)		-0.030 (0.022)
% seniors	0.015 (0.018)		0.010 (0.041)		0.016 (0.032)
% female	0.002 (0.009)		0.052** (0.021)		0.031* (0.016)
Class level diversity		0.031* (0.017)		0.061 (0.043)	
Gender diversity		-0.023* (0.013)		-0.000 (0.034)	
% out-of-state	-0.010 (0.009)	-0.008 (0.009)	-0.040* (0.022)	-0.030 (0.022)	-0.029* (0.017)
Max GPA	-0.002 (0.006)	-0.001 (0.006)	0.041*** (0.015)	0.035** (0.015)	0.032*** (0.012)
Max iRAT	0.323*** (0.031)	0.333*** (0.030)			
Team cohesiveness	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)
Environmental					
Team size	0.001 (0.003)	0.001 (0.002)	0.001 (0.006)	-0.000 (0.006)	0.002 (0.005)
Class size	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Constant	0.680*** (0.036)	0.663*** (0.033)	0.769*** (0.069)	0.748*** (0.069)	0.807*** (0.053)
Adj. R-squared	0.5310	0.5622	0.1269	0.0334	0.1150

Standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significant at the 1%, 5%, and 10% levels respectively.

Table 4: Team Cohesiveness (N=106)			
	Model 1	Model 2	Model 3
% freshman	5.12** (2.02)		5.56*** (2.05)
% juniors	3.79* (2.18)		3.93* (2.17)
% seniors	5.54* (3.16)		5.21 (3.16)
% females	-0.83 (1.60)		-1.11 (1.61)
Class level diversity		4.56 (3.12)	
Gender diversity		-1.77 (2.57)	
% out-of-state	-1.30 (1.70)	-0.84 (1.68)	-0.93 (1.72)
Overall team grade			12.12 (9.89)
Environmental			
Team size	-0.36 (0.47)	-0.31 (0.47)	-0.37 (0.47)
Class size	0.016 (0.037)	-0.005 (0.036)	0.021 (0.037)
Constant	-3.64 (2.87)	-6.47 (9.22)	-14.88 (9.60)
Adj. R-squared	0.0067	-0.0202	0.0118

Standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significant at the 1%, 5%, and 10% levels respectively.

Table 5: Individual Performance on Final Exam (N=635)					
	Model 1	Model 2	Model 3	Model 4	Model 5
<b>Individual:</b>					
Female	-0.006 (0.008)	-0.012 (0.009)	-0.009 (0.008)	-0.005 (0.008)	-0.008 (0.008)
Freshman	-0.005 (0.010)	-0.003 (0.011)	-0.003 (0.010)	-0.008 (0.010)	-0.007 (0.010)
Junior	-0.003 (0.010)	-0.004 (0.011)	-0.003 (0.010)	-0.006 (0.010)	-0.006 (0.010)
Senior	0.026** (0.013)	0.025* (0.014)	0.027** (0.013)	0.025* (0.013)	0.026** (0.013)
First semester freshman	0.083*** (0.020)	0.079*** (0.020)	0.079*** (0.020)	0.088*** (0.019)	0.084*** (0.020)
Transfer	0.056** (0.023)	0.050** (0.024)	0.050** (0.024)	0.059** (0.023)	0.053** (0.024)
Out-of-state	0.015* (0.008)	0.021** (0.009)	0.021** (0.009)	0.015* (0.008)	0.021** (0.009)
GPA	0.033*** (0.006)	0.032*** (0.006)	0.031*** (0.006)	0.034*** (0.006)	0.033*** (0.006)
Peer Eval	0.029*** (0.003)	0.029*** (0.003)	0.029*** (0.003)	0.027*** (0.003)	0.028*** (0.003)
<b>Team</b>					
% freshman		0.006 (0.027)			
% junior		0.008 (0.028)			
% senior		0.017 (0.040)			
% female		0.031 (0.022)			
Class level diversity			0.010 (0.040)		-0.003 (0.041)
Gender diversity			0.070** (0.032)		0.074** (0.031)
% out-of-state		-0.038* (0.022)	-0.036* (0.022)		-0.038* (0.022)
Average GPA		0.020 (0.020)	0.019 (0.020)		0.020 (0.020)
Overall team grade				-0.092 (0.117)	-0.115 (0.119)
Team cohesiveness				0.004*** (0.001)	0.004*** (0.001)
<b>Environmental</b>					
Team size		-0.003 (0.006)	-0.003 (0.006)		-0.002 (0.006)
Class size	-0.001*	-0.001	-0.001	-0.001*	-0.001

	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.448***	0.396***	0.386***	0.554***	0.512***
	(0.036)	(0.073)	(0.072)	(0.113)	(0.130)
Adj. R-squared	0.2227	0.2224	0.2282	0.2319	0.2384

Standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significant at the 1%, 5%, and 10% levels respectively.

Table 6: Individual Performance by Gender		
	Females (n=246)	Males (n=388)
<b>Individual:</b>		
Freshman	0.016 (0.015)	-0.020 (0.014)
Junior	-0.006 (0.015)	-0.008 (0.014)
Senior	0.035* (0.019)	0.020 (0.017)
First semester freshman	0.105*** (0.030)	0.074*** (0.027)
Transfer	0.059 (0.039)	0.067** (0.031)
Out-of-state	0.024** (0.012)	0.020 (0.013)
GPA	0.036*** (0.009)	0.035*** (0.008)
Peer Eval	0.036*** (0.006)	0.025*** (0.004)
<b>Team</b>		
% out-of-state	-0.050 (0.031)	-0.037 (0.031)
Average GPA	0.009 (0.027)	0.026 (0.029)
Team size	0.005 (0.008)	-0.007 (0.008)
Class level diversity	0.046 (0.061)	-0.021 (0.056)
Gender diversity	0.042 (0.047)	0.087** (0.042)
Overall team grade	-0.045 (0.175)	-0.147 (0.159)
Team cohesiveness	0.005*** (0.002)	0.004** (0.002)
<b>Environmental</b>		
Class size	-0.000 (0.001)	-0.001 (0.001)
Constant	0.311 (0.196)	0.593*** (0.173)
Adj. R-squared	0.2765	0.2147

Standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significant at the 1%, 5%, and 10% levels respectively.

Table 7: Individual Performance by Class Level				
	Freshman (n=213)	Sophomore (n=212)	Junior (n=131)	Senior (n=78)
<b>Individual:</b>				
Female	0.015 (0.015)	-0.016 (0.014)	-0.016 (0.015)	-0.006 (0.022)
First semester freshman	0.008 (0.029)			
Out-of-state	0.022 (0.018)	0.025 (0.016)	0.011 (0.018)	0.040* (0.24)
GPA	0.004 (0.009)	0.025*** (0.006)	0.042*** (0.008)	0.034** (0.014)
Peer Eval	0.046*** (0.007)	0.026*** (0.005)	0.021*** (0.006)	0.012 (0.008)
<b>Team</b>				
% out-of-state	-0.079 (0.047)	-0.022 (0.037)	-0.017 (0.041)	-0.059 (0.052)
Average GPA	0.027 (0.035)	0.027 (0.034)	0.087** (0.039)	0.035 (0.062)
Team size	-0.002 (0.012)	-0.003 (0.010)	0.007 (0.011)	-0.004 (0.017)
Class level diversity	-0.037 (0.080)	0.004 (0.072)	0.039 (0.074)	-0.006 (0.125)
Gender diversity	0.050 (0.061)	0.052 (0.053)	0.116* (0.061)	0.074 (0.087)
Overall team grade	-0.268 (0.226)	-0.172 (0.209)	-0.127 (0.236)	0.213 (0.329)
Team cohesiveness	0.005** (0.002)	0.002 (0.002)	0.001 (0.003)	0.006 (0.004)
<b>Environmental</b>				
Class size	0.000 (0.001)	-0.001 (0.001)	-0.002*** (0.001)	-0.001 (0.001)
Constant	0.528** (0.242)	0.601** (0.236)	0.330 (0.247)	0.383 (0.404)
Adj. R-squared	0.2327	0.2198	0.3123	0.1542

Standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significant at the 1%, 5%, and 10% levels respectively.

Table 8: Individual Performance by Grade Point Average				
	Below 2.5 (n=115)	2.5 to 2.99 (n=119)	3.0 to 3.49 (n=156)	3.5 to 4.0 (n=90)
<b>Individual:</b>				
Female	-0.015 (0.029)	-0.010 (0.020)	-0.016 (0.012)	-0.010 (0.016)
Freshman	-0.021 (0.030)	-0.004 (0.024)	0.014 (0.018)	-0.038* (0.021)
Junior	-0.005 (0.031)	0.003 (0.021)	0.003 (0.014)	0.003 (0.020)
Senior	-0.011 (0.043)	0.038 (0.31)	0.030* (0.016)	0.013 (0.022)
Out-of-state	0.002 (0.030)	-0.001 (0.021)	0.027* (0.014)	-0.006 (0.017)
GPA	-0.012 (0.031)	-0.075 (0.064)	0.103** (0.041)	-0.003 (0.044)
Peer Eval	0.018* (0.011)	0.018*** (0.006)	0.014** (0.005)	0.033*** (0.009)
<b>Team</b>				
% out-of-state	-0.002 (0.067)	-0.009 (0.052)	-0.021 (0.033)	0.055 (0.043)
Average GPA	0.105 (0.073)	0.004 (0.045)	0.016 (0.032)	-0.054 (0.044)
Team size	0.008 (0.017)	0.018 (0.014)	0.005 (0.009)	-0.027* (0.014)
Class level diversity	-0.035 (0.123)	-0.024 (0.090)	0.036 (0.064)	0.020 (0.079)
Gender diversity	0.171 (0.116)	0.142** (0.069)	0.058 (0.051)	0.059 (0.062)
Overall team grade	-0.583 (0.376)	-0.186 (0.260)	-0.072 (0.173)	0.792*** (0.267)
Team cohesiveness	0.009** (0.004)	0.003 (0.004)	0.002 (0.002)	0.004* (0.002)
<b>Environmental</b>				
Class size	-0.001 (0.001)	-0.002 (0.001)	-0.002** (0.001)	0.001 (0.001)
Constant	0.828** (0.379)	0.890*** (0.336)	0.366 (0.223)	0.076 (0.297)
Adj. R-squared	0.0839	0.0601	0.1098	0.2888

Standard errors are in parentheses. \*\*\*, \*\*, and \* indicate significant at the 1%, 5%, and 10% levels respectively.