Journal of Advanced Academics
jaa
Volume 21 ✤ Number 3 ✤ Spring 2010 ✤ pp. 500–529

What Determines High- and Low-Performing Groups?

The Superstar Effect

Priya K. Nihalani
University of Texas at Austin

Hope E. Wilson
Stephen F. Austin State University

Gregory Thomas
University of Texas

Daniel H. Robinson
University of Texas at Austin

In Vygotsky's (1978) seminal essay on education he remarked that, "learning awakens a variety of internal developmental processes that are able to operate only when the [student] is interacting with people in his environment and with his peers" (p. 90). Then, in the 1960s, several decades after Vygotsky's death, the emerging cognitive revolution coupled with breakthrough research on how people learn prompted interest in the form of meaning-making that only occurs within a social culture (Greenwood, 1999). Today, educators are urged to promote these sorts of social encounters within their classrooms to foster learning amongst students. Further, education researchers have mostly adopted the notion that learning does not occur in a vacuum and encouraged the influence of social constructivist principles when designing

summary

The notion that greater learning outcomes will be achieved if the cognitive work is distributed amongst a group of individuals working together versus working alone has received mixed support when explored empirically (e.g., Daiute & Dalton 1993; Johnson & Johnson, 1991). This study examined the relationship between small-group collaborative learning structures and the potential predictors of groups' overall academic performance. We sought to identify specific factors that distinguished high-performing groups from low-performing groups in the classroom. Class attendance and individual-level academic performance were positively related to group-level academic performance. Further, it was predicted that groups consisting of an exceptionally high-performing member, or superstar, would achieve greater group-level academic performance than groups consisting of members who performed similarly. However, the greater the distance between the highest-performing member's score and the average of the other group members' scores on individual-level tasks, the lower the score on group-level tasks. This difference between the highest scoring group member and the rest of the members is referred to as the Superstar Difference Score. Qualitative and quantitative analyses indicated that the Superstar Difference Score is a reliable, negative predictor of group-level academic performance. Practical implications for classroom instructors and future directions for education research resultant from this study's superstar effect are discussed.

Nihalani, P. K., Wilson, H. E., Thomas, G., & Robinson, D. H. (2010). What determines high- and low-performing groups? The superstar effect. *Journal of Advanced Academics, 21*, 500–529.

classroom environments (Bransford, Brown, & Cocking, 2000). Social constructivism describes the student as an active "maker of meanings" who participates in a small culture to co-construct knowledge with other students (Voss, Wiley, & Carretero, 1995).

As inquiry into how people learn moved toward considering individual differences in cognitive demands and determining how learning environments could address these differences, social constructivism gained prominence as a learning theory. Researchers hypothesized that by distributing the cognitive work amongst a group of individuals working together, the groups could attain more success than individuals working alone (Bruner, 1990). Further, observations of students working together have found that peer-to-peer interactions may be even more facilitative for active meaning-making than teacher-student interactions, given the shared perspectives and life experiences (Daiute & Dalton, 1993; Vygotsky, 1978; Yilmaz, 2008).

The awareness of the social dimension of learning that has evolved over the last half-century has resultantly changed the nature of effective pedagogy. Where the emphasis was once on direct instruction, today's teachers utilize a multitude of varying instructional techniques, and those that are influenced by social constructivism require equal or greater involvement by students in the learning environment (Jonassen, 1991). Examples of such instructional techniques include collaborative or cooperative learning, reciprocal learning, distributed cognition, and cognitive apprenticeship (Brown, Collins, & Duguid, 1989; Hutchins, 1995; Resnick, Levine, & Teasley, 1991; Slavin, 1995). However, prior research on these strategies have reported discrepancies in work products produced by a group, or pair, of students working together and those produced by the individual members, as well as work products produced by groups over a semester (Johnson & Johnson, 1991). The present study sought to contribute to this line of research by examining the relationship between collaborative learning and potential predictors of groups' overall academic performance.

Collaborative Learning

As a well-established instructional strategy, *collaborative learning* refers to a small group of students who cognitively and cooperatively engage in a common task to achieve a shared goal (Brandon & Hollingshead, 1999; Johnson & Johnson, 1989; Slavin, 1995). Today, collaborative learning procedures have widespread use at every tier of the education process from preschool to graduate school, and across various subject domains (Johnson & Johnson, 2000). The general effectiveness of this strategy has been supported theoretically, validated empirically, and operationalized into practical procedures for educators (Onwuegbuzie, Collins, & Elbedour, 2003). For example, when collaborative learning has been combined with short periods of formal lecture at the postsecondary level, students have demonstrated increased involvement in class discussions, heightened motivation, and generally more positive attitudes toward learning (Munoz & Huser, 2008; Slavin, Hurley, Chamberlain, Reynolds, & Miller, 2003). Also at the university level, collaborative strategies have been utilized to improve both creative thinking and communication skills in addition to mental organization of novel information (Mason, 2006; Rhys & Fetherston, 2008; Zuheer, 2008). In undergraduate courses with larger numbers of students enrolled, such benefits cannot be achieved when students work individually (Slavin & Karweit, 1981). Johnson, Johnson, and Smith's (1991) meta-analysis of 164 studies compared collaborative learning to individual learning strategies. They found evidence of greater achievement levels, more instances of advanced reasoning, and increased transfer in the collaborative learning conditions. In addition to academic effectiveness, collaborative learning has been shown to increase females' self-efficacy in competitive environments (Rodger, Murray, & Cummings, 2007). Similar results have been demonstrated in educational settings outside the United States (e.g., Gillies & Boyle, 2006; Zakaria & Iksan, 2007) and through the research in special education (e.g., Wolford, Heward, & Alber, 2001).

Collaborative learning tasks are meant to foster shared thinking and co-construction of meaning amongst students (Murphy

& Alexander, 2005; Ormrod, 2008). Collaborative efforts to learn, to understand, and to solve a range of problems are central for constructing knowledge structures that can be efficiently applied to novel tasks. To maximize learning outcomes and create an environment that is conducive to positive group interaction, three conditions must be met. First, group members must feel positively interdependent and, at the same time, individually accountable for their own academic goals. Second, individual members should demonstrate support of fellow members' efforts toward task completion (i.e., by offering corrective feedback). Finally, members need to be reflective of their group's smaller achievements as they continue to work toward a greater common goal (Johnson & Johnson, 2002). Instructors can support such an environment by allocating enough class time for face-to-face interaction and encouraging practices such as guided peer questioning (Hagman & Hayes, 1986; Ormrod, 2008).

Issues in Collaborative Learning

Despite the mass of literature praising collaboration amongst peers, sometimes the use of collaborative learning structures, rather than delivering impressive results, is ineffective (Slavin et al., 2003). Simply placing individuals in groups and assigning tasks does not necessarily lead to the positive results discussed previously (Johnson & Johnson, 1989). Several studies have reported discrepancies between performance scores reported at the group level and those reported at the individual member level (e.g., Hatano & Inagaki, 1991; Johnson & Johnson, 1991; Salomon & Perkins, 1998; Salomon, Perkins, & Globerson, 1992; Webb, 1992). One possible explanation of this phenomenon may be that, within a given group, a highly variant individual member's abilities cause variable, or difficult to predict, group-level performance. A second possibility may be that a single high-ability group member dominates group-level tasks such that group-level performance is more reflective of this particular member's ability opposed to a composite of all members' ability. The latter explanation has been termed the *Matthew effect* (Merton, 1973) and is discussed next.

The Matthew effect. Originating within the field of sociology, the Matthew effect describes a situation wherein an individual possesses an excessive amount of resources and then leverages those resources to obtain even more (Merton, 1973). In other words, the Matthew effect describes an unequal distribution of resources among individuals in a community. Within education, the term *Matthew effect* was adopted by Canadian psychologist Keith Stanovich to describe trends he observed between new readers as they developed skills associated with reading: Early success in acquiring reading skills appeared to correlate with future successes in reading, while students who failed to acquire those same skills before the third or fourth grade exhibited similar difficulties in other areas (Cunningham & Stanovich, 1997; Stanovich, 1986). Longitudinal data revealed a fan-shaped spread in achievement; when educational outcomes were plotted against time, students' rates of achievement were relative and proportional to their initial skill acquisition (Adams, 1990; Cunningham & Stanovich, 1997).

Within a collaborative learning environment, members perceived as more capable are often deferred to and this deference may support those members in obtaining greater comprehension of novel information. In other words, if the discourse within collaborative groups supports meaning-making, and thus comprehension, and, if the members who are deferred to speak more often, then the more capable members may achieve greater levels of comprehension (Fogarty & Pete, 2007; Jonassen & Kwon, 2001). This process would become cyclical resulting in a variation of the Matthew effect.

To avoid what some refer to as "the rich getting richer," instructors can aim to maximize inter-group homogeneity by assigning groups based on members' cultural background, gender, prior knowledge, and academic major (Onwuegbuzie et al., 2003). Even distribution of these variables across groups may be especially important in learning settings where competition between groups (i.e., exam bonus points) can reveal an unfair distribution of these variables (Michaelsen, Knight, & Fink, 2002). Nevertheless, prior research has also demonstrated that heterogeneous groups have greater difficulty establishing cohesion than homogeneous groups; however, they tend to catch up to homogeneous groups within a semester (Watson, Kumar, & Michaelsen, 1993).

Team-based learning. A form of collaborative learning that has received praise is team-based learning (TBL). Within this specific strategy, students experience an increased sense of responsibility, engagement in the learning process, and develop interpersonal skills (Michaelsen, Watson, Schwartzkopf, & Black, 1992). TBL's social constructivist roots are evidenced by class time use; time is shifted away from direct lecture and instead allocated toward meaning-making activities that are completed in groups. This method also includes assessment opportunities for both individual and group-level performance.

TBL consists of three interrelated phases. In Phase 1, students study independently outside of class to master content objectives identified by the instructor. In Phase 2, students independently complete a multiple-choice exam to assure their readiness to apply their Phase 1 knowledge. Groups consisting of 6 to 7 members then retake the exam together and submit their consensus answers for immediate feedback scoring. Two performance scores are calculated; the first is for students' independent exam responses and the second is for groups' exam responses. Phase 3 lasts several class periods during which groups complete in-class assignments that require application of knowledge obtained during Phases 1 and 2. During preestablished times, all groups within a class will share their answers, which stimulates an energetic discussion between groups as they defend and explain their answers. The instructor plays a supportive role in helping facilitate discussion and consolidate content for comprehension.

The present study utilized a variation of TBL, referred to as team-based testing (TBT). TBT includes Phases 1 and 2 from the TBL process. The Phase 3 activities that take several classes to complete are replaced with activities that are completed during one class (Michaelsen et al., 1992). During the semester that this study took place, the common goal for groups was extra credit points to increase exam scores. The activities that teams worked on outside of the exams (Phase 2) ensured that groups received enough face-to-face interaction and promoted an encouraging learning environment through the group discourse that took place.

Our research team has been using TBT during the semester leading up to the present study in undergraduate educational psychology and introductory statistics courses. Anecdotally, during these previous experiences using TBT there appeared to be a trend where several groups in each class were unable to receive extra credit and, as the semester progressed, seemed to essentially give up trying. As observers, it seemed these groups were composed of members that were discordant from the very beginning. In contrast, the more successful groups were composed of members that interacted more harmoniously with each other. These observations represent the motivation for further exploring the distinguishing characteristics of high- and low-performing collaborative groups in the present study.

Present Study

Prior research on classroom small-group structures has described emergent patterns that clearly separate high- and low-performing groups. Johnson and Johnson (1996) found that success in collaborative groups result from members embodying collectivist values (i.e., interdependence, joint effort). As a class progresses through the TBL process, group-level performances are reflective of more than simply the sum of individual members' performances as they work together to become a team (Watson et al., 1993). For this reason, the instructor of the classes used for this study adapted the term *team* instead of *group*.

The purpose of this study was to identify the distinguishing characteristics among members of high- and low-performing teams that may predict group-level performance to inform instructors on their use of small-group learning structures. Individual- and group-level performance data were collected, discussions amongst members were recorded and transcribed, and groups were observed in an attempt to ascertain whether certain individual characteristics inhibit or promote successful group, or team, performance.

Based on prior collaborative learning research, we predicted that higher performing teams would have a greater degree of homogeneity than the poor performing teams. However, the instructor formed teams to maximize inter-group homogeneity so the only characteristic that could emerge as the semester progressed was members' individual performance record. Thus, the teams comprised of members that exhibited similar individual performance scores would predictably obtain higher team performance scores than teams consisting of one member that performed significantly higher than the other members. We dubbed this phenomenon "the superstar effect."

Hypothesis 1. We expected teams comprised of similarly performing members to achieve higher team-level performance scores than teams that experience a superstar effect where one member obtains significantly higher individual-level scores than the other team members.

Secondly, we hypothesized that strong patterns in team discourse would reflect high-performing teams' strong collective mentality. Low-performing teams, on the other hand, were hypothesized to have more widely varying scores within the team, evidence that they were simply a collection of individuals without the collective mentality of high-performing teams.

Hypothesis 2. We hypothesized that discourse patterns of high-performing teams would be reflective of a stronger collective mentality than low-performing teams.

Method

Participants

The study took place within introductory educational psychology and statistics courses at a large Southwestern university. The instructor assigned a total of 101 students to teams during the first week of class. The assignment mechanism was mostly random. However, the instructor made an effort to balance the teams by gender and cultural background. This university and, by extension, these two courses, have an ethnically diverse population that includes many international and nontraditional students. Both courses were also composed of students from widely varying disciplines such as nursing, social work, education, communication, liberal arts, and business. The primary difference between the two courses was the calculation component in the statistics course. Teams consisted of 5 to 7 members. The collaborative learning literature recommends groups consist of 3 or 4. However, this decision was based on previous attendance records: Some students were absent during days allocated for team activities. Students remained in the same teams throughout the semester. Performance data collected reflected 9 teams of 7, 8 teams of 6, and 1 team of 5 students. The basic demographics upon which teams were created were fairly comparable between the two classes chosen for this study, thus allowing for the combining of the two samples for purposes of analysis. The teams selected for the qualitative analyses portion of the present study were those in which all members consented to being observed and recorded.

Procedure

We collected data on class attendance, individual and team performance data on quizzes and tests, and recordings of discussions during team activities over the course of the semester. Neither the instructor nor teaching assistant was privy to this data until final semester grades were submitted.

TBT method. The structure of both classes was essentially identical. Before the lecture on a new unit began in the class, students individually took a 15-point multiple-choice quiz on the assigned reading. After students completed the quiz individually, they turned in their individual answer sheets but retained the quiz sheet upon which they had also recorded their answer choices. Having turned in their quizzes, the students then met in their teams to discuss their answers. The team task was to collaboratively complete the same quiz using Immediate Feedback Assessment Technique (IF-AT) answer sheets that allowed immediate feedback question by question. The IF-AT assessment forms provide affirmative and/ or corrective immediate feedback and are available for commercial purchase (see http://www.epsteineducation.com).

For each question, when a team had made a choice among alternatives, it would scratch the chosen answer on the IF-AT form and, if correct, a star was revealed. If a team chose an incorrect answer, an empty spot was revealed and they had the opportunity to choose again whereby two "tries" were given for each question. If the correct answer was chosen on the first attempt, one point was awarded for that specific question. If the correct answer was not chosen on the first attempt but rather the second, half a point was awarded for that question. If the correct answer was still not chosen on the second attempt, no points were awarded. Points were then tallied for the team's quiz.

Point calculation. For a complete breakdown of grade points, see Table 1. For quizzes, the team that received the highest score in the class (or teams if there was a tie) received 2 bonus points that were added to each team member's individually obtained scores. Members of the team that earned the second highest score in the class received 1 additional bonus point. Across the semester there were 6 quizzes, each worth 15 points. If students received a perfect score on their quiz and were members of the highest performing team for each of the 6 quizzes, the quiz points toward the final grade would be 102 points (90 plus 12 bonus points).

For tests, members of the team (or teams) that achieved the highest score in the class were each awarded 9 bonus points on their individually obtained test scores, 6 points were earned by the second highest team, and 3 bonus points were earned by the third highest team. Tests were administered three times over the semester and were worth 60 points each. A student who received a perfect test score and was a member of the highest scoring team on each of the three administrations would receive a score of 207 points (180 plus 27 bonus points) toward the final course grade.

Table 1

Grade Point Breakdown

	Points worth	Occurrences in semester	Total	Possible bonus points	Occurrences in semester	Total	Grand total
Tests	60	3	180	9	3	27	207
Quizzes	15	6	90	2	6	12	102
Practice exercises	N/A	N/A	N/A	4	6	24	24

***Note*. Grading scale: 90% and higher = A; 80-89% = B; 70-79% = C; 60-69% = D; less than 60 % = F**

Students absent during classes when a quiz or test was given were given the opportunity to take a make-up. However, if their team had received bonus points, the absentee was not awarded those bonus points because he or she had not contributed to the team's discussion and achievement.

In addition to quizzes and tests, games that consisted of practice exercises were also administered, or played, for additional extra credit points. These games required team participation only and represented the team-level activities of Phase 3 in the TBL strategy. Team members on the highest scoring team received 4 points, those on the second highest scoring team received 3 points, and so on. Games were played 6 times over the semester. Thus, students could receive up to 24 bonus points. Students who were absent during these days did not receive bonus points. To some extent then, attendance on game days were a measure of students' motivation to improve their grade and contribute to the achievements of their team.

Variables. The *team score* is the average score of the team-level quizzes and tests completed across the semester; games were excluded from calculation of this variable. *Individual score* is the average score of the individual-level quizzes and tests completed across the semester. *Superstar score* is the highest individual score for each team and thus represents each team's highest-scoring member across the semester. *Superstar difference* is the distance between the superstar score and the average of the remaining team member's individual scores. *Bonus score* is the sum of the team's bonus points given for top-performing groups on tests and quizzes. *Game score* is the sum of the teams' scores on the game/review days. See Tables 2 and 3 for individual- and team-level descriptive statistics.

Analyses

Regression

Two different quantitative analyses were carried out to explore the potential superstar effect. Multiple regression was first con-

Table 2

Descriptive Statistics for Student-Level Variables

Variable	*N*	Percent	Mean	*SD*
Male	23	22.3		
Female	74	71.8		
Below 18	1	1.0		
18-21	75	72.8		
22-25	21	20.4		
26-29	1	1.0		
30 and over	3	2.9		
White	52	50.5		
Black	9	8.7		
Asian or Pacific Islander	17	16.5		
Hispanic	18	17.5		
Multi-racial	5	4.9		
Individual Quiz Score	101		11.10	1.76
Individual Test Score	101		45.92	6.56
Individual Score	101		22.71	2.91

Table 3

Descriptive Statistics for Team-Level Variables

Variable	*N*	Mean	*SD*
Team Score	18	4.88	2.72
Team Bonus	18	20.83	9.67
Team Games	18	15.75	6.32
Super Star Difference	18	3.26	1.93
Super Star Score	18	25.37	2.71

ducted to identify team-level performance predictors and a multilevel model was then used to account for team-level effects on individual scores.

To explore the superstar effect on team-level achievement, we conducted a multiple regression to predict average team achievement. The following variables were included as predictors in the

Table 4

Bivariate Correlations of Team-Level Variables

	Team Score	Team Bonus	Team Games	Super Star Score	Super Star Difference
Team Score	1.00				
Team Bonus	.821**	1.00			
Team Games	.426**	.646**	1.00		
Super Star Score	.288**	.168	-.088	1.00	
Super Star Difference	-.093	-.040	-.211*	.814**	1.00

Note. *$p \le .05$. **$p \le .001$.

model: individual scores, superstar scores, superstar difference, game scores, and bonus points. There were significant bivariate correlations between team scores and individual scores, superstar scores, bonus scores, and game scores (see Table 4). Because the superstar difference was of theoretical significance to this research, it was also included in the multiple regression model.

Multilevel Model

Although the regression model predicted team-level performance, the multilevel model used individual, or student, scores as the dependent variable to predict average team score and analyze the potential superstar effect on student-level achievement. This is a two-level model, with level 1 being student and level 2 being team. An additional variable, *superstar status*, was created to indicate each team's superstar member. Superstar status was included as a level-1 variable and coded dichotomously with 1 indicating the student was the superstar, or highest scoring member within their respective team, and 0 indicating that the student was not the superstar within his or her respective team. Therefore the level-1 model is,

$$Y_{ij} = \beta_0 + \beta_1(SuperstarStatus) + r_{ij}. \quad (1)$$

The level-2 model includes: the team score, which is the average of each of the scores on the quizzes as they were taken by the entire team; the superstar score, which is the average of the highest scoring student's individual scores on the quizzes; and the superstar difference score, which is the difference between the average of the superstar's scores on individual quizzes and the average of the rest of the team's scores on individual quizzes. The outcome variable is the average of the individual quizzes. Therefore, the team score is the average of a different set of scores from the individual scores used to calculate the superstar score and the superstar difference score. The correlation between the superstar score and the superstar difference score was .812.

Team score, superstar difference, and superstar score were included as level-2 variables in the level-2 model for the intercept (see Equation 2). Team score, superstar difference, and superstar score were also included as continuous variables in the level-2 model for the superstar status slope (see Equation 3).

$$\beta_0 = \gamma_{00} + \gamma_{01}(TeamScore) + \gamma_{02}(SuperstarDifference) + \gamma_{03}(SuperstarScore) + u_{0j} \quad (2)$$

$$\beta_1 = \gamma_{10} + \gamma_{11}(TeamScore) + \gamma_{12}(SuperstarDifference) + \gamma_{13}(SuperstarScore) + u_{1j} \quad (3)$$

The superstar status slope was included to measure the differing effect of these variables for superstars as compared to their team members.

Qualitative Analysis

The transcribed recordings of team discussions were used to investigate the relationship between team's discourse patterns and academic performance. The qualitative analysis was a recursive process between the team's discourse, members' individual-level response choices, and team-level response choices on quizzes and tests.

Table 5

Summary of the Regression Analysis for Variables Predicting Team Score

Variable	*B*	*SE (B)*	β
Constant	22.673**	.541	
Individual Quiz Scores	.008	.024	.017
Individual Test Scores	-.009	.007	-.072
Super Star Difference	-.287**	.037	-.661
Super Star Score	.220**	.029	.708
Bonus Score	.070**	.005	.792
Games Score	-.021*	.008	-.156

Note. * $p \leq .05$. ** $p \leq .001$.

Results

Regression Results

The results of the regression analysis showed that individual scores were not significant predictors of team scores, when considering the other variables in the model (see Table 5). However, the superstar scores and superstar differences were significant predictors of team scores. Superstars, or highest-performing team members, had a positive effect on the team score. Conversely, the superstar difference, which illustrated the distance between the superstar score and the average of the remaining team members' individual scores, negatively affected the team score, when considering the other variables in the model. Finally, the bonus and game scores also had a positive relationship with the team score, after accounting for the other variables in the model. This five predictor model accounted for 81.9% of the variance in team achievement, $F(6, 93) = 70.34, p < .001$.

Multilevel Results

During preliminary analyses, an intraclass correlation coefficient of 0.22 indicated that 22% of the variance in individual scores was between teams and that multilevel modeling would be

Table 6

Fixed Effects for Multilevel Model

Fixed Effect	Coefficient (*SE*)	*p*
Model for team average individual scores (β_0)		
Intercept (γ_{00})	22.16 (.21)	<.001
Team Score (γ_{01})	-.001 (.32)	.997
Super Star Difference (γ_{02})	-1.00 (.23)	.001
Super Star Score (γ_{03})	1.00 (.17)	<.001
Model for Super Star Slope (β_1)		
Intercept (γ_{10})	3.33 (.49)	<.001
Team Score (γ_{11})	.10 (.76)	.898
Super Star Difference (γ_{12})	1.01 (.55)	.088
Super Star Score (γ_{13})	-.004 (.41)	.993

Table 7

Random Effects for Multilevel Model

Random Effects	Variance	*df*	Chi-square
Variance in team average individual scores (τ_{00})	<.001	14	.009 ($p > .5$)
Variance within teams (σ^2)	3.630		

appropriate to explain the variance in scores at both levels. The model was estimated using Restricted Maximum Likelihood, and each of the level-2 variables were grand-mean centered indicating that "0" is the average score across all teams. Thus, the intercept represents the predicted score for a student on an "average" team.

The fixed and random effects for this model are reported in Tables 6 and 7. In examining the fixed effects, team score was not a significant predictor of individual scores after accounting for the other variables in the model. However, superstar score and superstar difference were both significant predictors of individual score. A negative relationship between superstar difference and individual score ($\gamma_{02} = -1.00$) indicated that for every 1 point difference between the superstar score and the team score, the predicted individual score fell by 1 point, after accounting for the other variables in the model. Predictably, there was a positive relationship between superstar score and individual score (γ_{03} = 1.00), indicating that for every 1 point increase in the highest individual score on a team, individual scores increased by 1 point. Interestingly, none of these variables were significant predictors of the slope in the model, indicating that there were no additional effects of these variables on the superstars, or highest scoring team members. Thus, while the model for the slope could be eliminated from analyses, we chose to include this in the analysis because the result may be due to a lack of power due to the small sample size (i.e., there were only 18 groups, and thus, only 18 highest-scoring team members). The variance between teams on individual scores was not statistically significant after accounting for the variables in the model (τ_{00} = <.001) and the variance within teams (σ^2) was 3.63. Thus, the model adequately explains the variance between teams.

Qualitative Results

There were indeed emergent discourse patterns between the high- and low-performing teams, and these patterns served as support of our hypotheses. Within the low-performing teams, the superstar, or highest-performing team member, dominated the discourse that took place. Both excerpts below are from two different low-performing team discussions from the educational psychology class. The first excerpt, Team A, took place during the middle of the semester and the second, Team B, took place closer toward the end of the semester. Superstars' dialogue are italicized in an effort to visually illustrate their level of discourse over their other members in their team.

Excerpt 1. Team A (low-performing team)

4: *Yeah, I thought it was E.*

5: 'Cause yeah I was looking over it before I turned it in and I changed it from B to E, like right before I handed it in, so . . .

4: *'Cause E could totally be . . . yeah, goal setting.*

5: Um . . .

6: I said E. But I mean, it's just like
3: I said B.
4: [reads aloud] *"distinguishing between important and unimportant information"*
2: I said E, because to me it seemed like you weren't focusing on main ideas . . .
4: I'll go with E, I'm going with E.
5: Maybe they'll remember that more than something else . . . Let's just—
1: I say E. But maybe B.
4: *We'll go with E.* [Scratching] *And it's wrong.*

Excerpt 2. Team B (low-performing team)

5: We're on 14, right? C, definitely.
3: You checked already [speaking to Student 5]? It's not C?
1: No! We got it wrong, what else could it be, A?
2: *But it doesn't have anything to do with relationships. I said D because . . .*
5: (interrupts) Well I put C, that's what I thought.
2: *Isn't that impossible to have both? Because if they have a high motive to avoid failure, he's never taken a chemistry class before, so he has no idea . . .*
6: (interrupts) He obviously has a low motive for failure.
2: *Well, if he is going to take it on, he wants to be really successful and at the same time, he's got to really try to not fail at it, cause then that would be, you know. But he doesn't know if he's not good at it or not, so . . .*
4: But I thought it was like a low motive to avoid failure just because he knew he was taking on something that was hard.
2: *That's a clue right there.*
4: When you like try to avoid failure, you do something that's easy . . .
2: (interrupts) *Hold on, look, or you do something . . . wait . . . what's the one where it's like two different opposite ends of the spectrum where you either do something that's impossible so when you fail you can attribute it to the fact that it was*

impossible, so it wasn't really you that failed, if it's possible. Hold on let me think.
1: Yeah.
2: *Between ability and effort, right? So how much effort you put forth and like the abilities you have, you really well, it's opposite so you either have high ability and low efforts, or what is it? I can see that page in my notes in my head. Ugh hold on.*
3: Let's go on.
4: You want to skip 14? I am all confused.
5: Then let's just put B.
2: *No, let's go with D.*
5: I'm trying B. Yeah, it's B, I knew it was B.
3: So it's D or B?
5: It was B.
2: *Right, so let's just go to 15.*

Notice that in both excerpts, the two superstars were incorrect in their answer choice. However, their performance scores indicate that across the semester, these students obtained the highest quiz and test scores. These excerpts may actually be representative of the superstars' making sense of course content through discourse dominance, providing evidence for the Matthew effect.

The following excerpts are from high-performing teams C and D within the educational psychology class during the middle and toward the end of the semester respectively. Our hypotheses stated that, for high-performing teams, a pattern would emerge that revealed the joint effort of team members who achieve success through collaboration. Discourse of these two teams supported our hypothesis. Superstars' dialogues are again italicized in an effort to visually illustrate their level of discourse for comparison with other members in their team.

Excerpt 3. Team C (high-performing team)
5: I was torn between D and E. I just didn't remember reading anything about information being presented in a simpler manner with Piaget.
4: I was between C or D, so.

5: I was torn between E and D.
4: I was between A and D.
6: *I was between C and D . . . and A.*
2: I was between A and D, too, but I put D.
4: What do you want to do?
6: *I don't back mine 100%, so . . . we can go with A.*
1: A is fine . . .
5: Yeah if y'all want to go with A. Since I'm not convinced . . .
6: *We can come back.*

2: Okay, I still don't know . . . See, I am still learning toward A. Is anyone else?
5: Yeah. I'm going towards A.
6: *Yeah, I'd say A.*
4: [laughs] Okay. Are y'all sure?
2: Yeah.
1: Now I'm going to feel bad if it…*scratching*
4: Yeah, it's A.
[various Yay! exclamations]
1: Fantastic.

Excerpt 4. Team D (high-performing team)

5: Okay, so, 24. I got A.
4: I put D.
1: I think all of these confused me; I got C.
3: It is B.
1: Why is it B? What was the question? Well, luck is uncontrollable, right? John's is uncontrollable. Controllable is what you have control over.
2: *So you have no control over the situation? So, between Ken and Frank? So let's see what the difference is between them.*
1: That's what I was thinking too.
3: Ken's IS controllable because he's saying: If I would've cracked a book, I probably would've done better.
6: Right. So that knocks off three answers.
5: That doesn't mean it's uncontrollable, though.
2: *I couldn't figure that out, maybe he's saying that was just easy?*

6: I mean that's what I took from it, I don't know.
1: Yeah, I see.
2: *Let's see what it's not: A? Why wouldn't it be A?*
6: It's not A, C, or E, it can't be Ken.
5: Yeah, I don't know why I put A, you're right.
3: It could be external because of the task difficulty, not because of his ability; Ability would be internal. External will be task difficulty.
5: That's true. He said anyone could have aced this test. But also, like in the book, like I just read it before . . .
3: No. I mean, I would argue that . . .
5: (interrupts) No, no, I definitely agree with you, I'm just saying that's such a hard question because in the book it clearly states it is external locus, due to luck.
2: *Yeah.*
1: So do we want to guess Frank for 24?
(all consent)
5: OK. Yay! We rock!

Through these excerpts it becomes evident that the low-performing teams' superstars took charge but their team failed to benefit from this. The high-performing teams, on the other hand, displayed a democratic discussion through which they were able to come to a consensus on a correct answer choice.

Discussion

Results from both the quantitative and qualitative analyses provide support that the superstar score and superstar difference are reliable predictors group-level performance. As educators, we would hope that high scoring group members would positively affect the overall group achievement. However, after controlling for the superstar's score, the superstar difference was negatively related to group-level achievement. Thus, after controlling for the superstar's score, the farther the highest scoring individual member is from the rest of the group, the lower the team scores on group exams. In predicting team-level performance, this effect is actually equal to the positive effect the team gains from having a higher superstar score. In other words, having a superstar on your team is only beneficial if the rest of the team also scores relatively high. If they do not, then the superstar does not help the team as much. This has implications for pedagogic endeavors and utilization of small-group structures in the classroom.

The results of this study indicate the practice of heterogeneously grouping students may be detrimental to overall team achievement if there is a great discrepancy between a high-performing individual and the rest of the team members. Thus, not all teams seem to benefit from having a top student, or superstar, in their group. The greater the superstar status, as compared with the rest of the group, the lower is the achievement of the overall team. This implies that more homogeneously matched ability groups may better allow all students to achieve at higher levels and contribute more equally to the team. Homogeneous groups may allow for greater social constructivism and learning to occur, thus maximizing the benefits of cooperative work.

In contrast to popular conceptions of collaborative learning, low-performing groups included an individual whose performance was considerably higher than other members' performance (a larger superstar difference value). This point is most evident in the discourse patterns that emerged amongst teams. Teams that had an individual with a high superstar status failed to display balanced team discussions whereby members contributed equally. Teams whose superstar score was closer to their team average score appeared to display more democratic discussions.

These results are not surprising, and in fact may appear to be common sense from an educator's perspective. High-performing students tend to perform consistently across environments. In addition to looking for reliable predictors of high-performing learning teams, it was expected that this implementation of TBL would support the widely held views of Johnson and Johnson (1996), Michaelsen et al. (2002), and Watson et al. (1993) that cooperative learning nurtures the development of joint effort, mutual responsibility, and democracy. The evidence here seems to suggest that lower performing teams are less cooperative than higher performing teams, which would be expected under the above assumptions. One might argue that the differences observed support the contention that the diversity of the team slows the cohesiveness development process.

Advantages and Limitations of Present Study

As Johnson and Johnson (1989) have asserted, cooperative efforts might be expected to be more productive than competitive or individualistic efforts only under certain conditions. The conditions that seem to promote successful cooperation are clearly perceived positive interdependence, considerable face-to-face interaction, and frequent use of relevant interpersonal skills. This implementation of TBL provided students with clearly perceived positive interdependence as they clearly understood that bonus points were available to high-performing teams.

The presence of considerable face-to-face interaction and frequent use of relevant interpersonal skills is somewhat more questionable as the terms "considerable" and "frequent" are ambiguous. There is a limited amount of time available in a semester-long undergraduate course. The total amount of time spent on team activities in these classes was approximately 10 hours. Interpersonal and team skills were practiced to a greater or lesser extent by all team members. The task structure, wherein the team must come to consensual answers on multiple-choice test items, limits the interpersonal and small-team skills that are necessary; however, the limited team interaction time could be seen by some as an argument for the need for specific interpersonal and small-team skills training.

One downside of the superstar phenomenon is the rich-get-richer or Matthew effect (Merton, 1968) as documented by Onwuegbuzie et al. (2003). Similar to the present study, teams containing high-performing individuals outperformed those teams with lower performing individuals. From one perspective, the superstars can be said to be depriving their lower performing counterparts of the opportunity to learn by taking over control of the team, thereby reaping the rewards of better understanding the content at the expense of the other team members.

On the other hand, Merton (1968) noted an upside of the Matthew effect. In the world of science, the well-known researcher often popularizes ideas that, if put forward by a lesser known scientist, might go unnoticed. Similarly, coauthoring a paper with a well-known scientist gets the other more attention, recognition, and the opportunity to work with a senior scientist in the field. Along this reasoning, the formation of groups with heterogeneous levels of ability may still be a beneficial practice (although not in the present study). The lower performing individuals may get bonus points awarded their team, simply due to the superstar's performance, that they would not otherwise receive. Additionally, lower performing team members are in a position to observe some part of the superstar's cognitive process by their close proximity, even if the superstar is not explicitly "teaching" them.

Recommendations for Future Research

Will knowing that team performance is contingent on the performance of specific individuals and their commitment to the process assist instructors in assigning students to teams so that competition for bonus points is relatively even among all teams, resulting in maximum learning? Perhaps a future intervention study could examine whether adding the results of a pretest of course material and the answers to a couple of questions to accurately probe commitment to the criteria already used for team assignment will allow for more even competition among teams. Future research should also examine whether assigning students to teams in such a way that minimizes the superstar effect would facilitate learning.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Brandon, D. P., & Hollingshead, A. B. (1999). Collaborative learning and computer-supported groups. *Communication Education, 48,* 109–126.

Bransford, J. D., Brown, A. L., & Cocking R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: Nation Academy Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning, *Educational Researcher, 18,* 32–42.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934–945.

Daiute, C., & Dalton, B. (1993). Collaboration between children learning to write: Can novices be masters? *Cognitive and Instruction, 10,* 281–333.

Fogarty, R., & Pete, B. M. (2007). *How to differentiate learning: Curriculum, instruction, and assessment*. Washington, DC: Corwin Press.

Gillies, R., & Boyle, M. (2006). Ten elementary teachers' discourse and reported pedagogical practices during cooperative learning. *The Elementary School Journal, 106,* 429–451.

Greenwood, J. (1999). Understanding the "cognitive revolution" in psychology. *Journal of the History of the Behavioral Sciences, 35,* 1–22.

Hagman, J., & Hayes, J. (1986). *Cooperative learning: Effects of task, reward, and group size on individual achievement.* Alexandria, VA: Army Research Institute for the Behavioral and Social Sciences.

Hatano, G., & Inagaki, K. (1991). Sharing cognition through collective comprehension activity. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 331–348). Washington, DC: American Psychological Association.

Hutchins, E. (1995). How a cockpit remembers its' speeds. *Cognitive Science, 19,* 265–288.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book.

Johnson, D. W., & Johnson, R. T. (1991). Social interdependence theory and university instruction: Theory into practice. *Swiss Journal of Psychology, 61,* 119–129.

Johnson, D. W., & Johnson, R. T. (1996). Cooperative learning and traditional American values: An appreciation. *NASSP Bulletin, 80*(579), 63–65.

Johnson, D. W., & Johnson, R. T. (2000). *Joining together: Group theory and group skills* (7th ed.). Boston, MA: Allyn & Bacon.

Johnson, D. W., & Johnson, R. T. (2002). Social interdependence theory and university instruction: Theory into practice. *Swiss Journal of Psychology, 61,* 119–129.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1991). *Active learning: Cooperation in the college classroom.* Edina, MN: Interaction Book.

Jonassen, D. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology, Research, & Development, 39*(3), 5–14.

Jonassen, D., & Kwon, H. (2001). Communication patterns in computer-mediated vs. face-to-face group problem solving. *Educational Technology, Research, and Development, 49*(10), 35–52.

Mason, L. H. (2006). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology, 96,* 283–296.

Merton, R. K. (1968). The Matthew effect in science. *Science, New Series, 159,* 56–63.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations.* Chicago, IL: University of Chicago Press.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (2002). *Team-based learning: A transformative use of small groups*. Westport, CT: Praeger.

Michaelsen, L. K., Watson, W. E., Schwartzkopf, A., & Black, R. H. (1992). Group decision-making: How you frame the question determines what you find. *Journal of Applied Psychology, 77,* 106–108.

Munoz, C., & Huser, A. (2008). Experiential and cooperative learning: Using a situation analysis project in principles of marketing. *Journal of Education for Business, 83,* 214–220.

Murphy, P. K., & Alexander, P. A. (2005). *Understanding how students learn: A guide for instructional leaders.* Thousand Oaks, CA: Corwin Press.

Onwuegbuzie, A., Collins, K., & Elbedour, S. (2003). Aptitude-by-treatment interactions and Matthew effects in graduate-level cooperative-learning groups. *Journal of Educational Research, 96,* 217–230.

Ormrod, J. E. (2008). *Human learning* (5th ed.). Alexandria, VA: Prentice Hall.

Resnick, L., Levine, S., & Teasley, L. (1991). *Perspectives of socially shared cognition*. Washington, DC: American Psychological Association.

Rhys, K., & Fetherston, B. (2008). Productive contradictions: Dissonance, resistance, and change in an experiment with cooperative learning. *Journal of Peace Education, 5,* 97–111.

Rodger, S., Murray, H. G., & Cummings, A. L. (2007). Gender differences in cooperative learning with university students. *Alberta Journal of Educational Research, 53,* 157–173.

Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. *Review of Research in Education, 23,* 1–24.

Salomon, G., Perkins, D. N., & Globerson, T. (1992). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher, 20*(3), 2–9.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Boston, MA: Allyn & Bacon.

Slavin, R. E., Hurley, E. A., Chamberlain, A., Reynolds, W. M., & Miller, G. E. (2003). Cooperative learning and achievement: Theory and research. In W. M. Reynolds, G. J. Miller, & I. B. Weiner (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 177–198). New York, NY: John Wiley & Sons.

Slavin, R. E., & Karweit, N. (1981). Cognitive and effective outcomes of an intensive student team learning experience. *Journal of Experimental Education, 50,* 29–35.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–407.

Voss, J. F., Wiley, J., & Carretero, M. (1995). Acquiring intellectual skills. *Annual Review of Psychology, 46,* 155–181.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watson, W. E., Kumar, K., & Michaelsen, L. K. (1993). Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups. *Academy of Management Journal, 36,* 590–602.

Webb, G. (1992). On pretexts for higher education development activities. *Higher Education, 24,* 351–361.

Wolford, P. L., Heward, W. L., & Alber, S. R. (2001). Teaching middle school students with learning disabilities to recruit peer assistance during cooperative learning group activities. *Learning Disabilities Research & Practice, 16,* 161–173.

Yilmaz, K. (2008). A vision of history teaching and learning: Thoughts on history education in secondary schools. *The High School Journal, 92,* 37–46.

Zakaria, E., & Iksan, Z. (2007). Promoting cooperative learning in science and mathematics education: A Malaysian perspective. *Eurasia Journal of Mathematics, Science, and Technology, 3,* 35–39.

Zuheer, K. M. M. (2008). *The effect of using a program based on cooperative learning strategy on developing some oral communication skills of students, at English department, faculty of education, Sana's University* (ERIC Document Reproductive Services No. 502845). Retrieved from ERIC database.

Priya K. Nihalani is currently in her fourth year as a doctoral student in educational psychology. Her research interests and expertise lie in innovative instructional design and assessment through educational technology. *Contact information:* The University of Texas at Austin, Department of Educational Psychology, 1 University Station D5800, Austin, TX 78712-0383; p_k_nihalani@yahoo.com

Karen Rambo is completing a Ph.D. in measurement, evaluation, and assessment in educational psychology at the University of Connecticut. Before beginning her work at UConn in 2007, she taught mathematics for 10 years and spent 8 of those years working with gifted students in Texas. Her research interests include large scale assessments, growth mixture models, multilevel modeling, and gifted and mathematics education. *Contact information:* University of Connecticut, 249 Glenbrook Road, Unit 2064, Storrs, CT 06269; karen.rambo@uconn.edu

Sally M. Reis is a professor and the past department head of the Educational Psychology Department in the Neag School of Education, where she also serves as a principal investigator for the National Research Center on the Gifted and Talented. She has authored or coauthored more than 200 articles, books, book chapters, monographs, and technical reports. She recently was honored with the highest award in her field, Distinguished Scholar of the National Association for Gifted Children; she was also named a Board of Trustees Distinguished Professor at the University of Connecticut. *Contact information:* National Research Center on Gifted and Talented, University of Connecticut, 2131 Hillside Road Unit 3007, Storrs, CT 06269; sally.reis@uconn.edu

Daniel H. Robinson has been a faculty member in educational psychology at the University of Texas since 1999. He is editor of *Educational Psychology Review* and serves on the board of directors of the Meadows Center for Preventing Educational Risk. *Contact information:* The University of Texas at Austin, Department of Educational Psychology, 1 University Station D5800, Austin, TX 78712-0383; dan.robinson@mail.utexas.edu

Emily J. Shaw is an associate research scientist in the Higher Education Research and Testing group at The College Board. Her research primarily focuses on test validity, college preparation and success, and college admission practices. *Contact information:* The College Board, 45 Columbus Avenue, New York, NY 10023; eshaw@collegeboard.org

Erin Sullivan is a doctoral student in educational psychology with dual concentrations in gifted education and school psychology at the University of Connecticut. *Contact information:* University of Connecticut, 249 Glenbrook Road, Unit 2064, Storrs, CT 06269-2064

Gregory Thomas received his Ph.D. in educational psychology from the University of Texas in 2007. His research interests include team-based learning and group dynamics in education. He has taught college courses in Taiwan and Saudi Arabia. *Contact information:* University of Texas, 1 University Station D5800, Department of Educational Psychology, Austin, TX 78712-0383; drgreg.sa@gmail.com

Hope E. Wilson is an assistant professor at Stephen F. Austin State University. She earned her Ph.D. in educational psychology-gifted education from the University of Connecticut. She is coauthor of the recent book, *Letting Go of Perfect: Overcoming Perfectionism in Children*. *Contact information:* Stephen F. Austin State University, Department of Elementary Education, SFA Box 13017, Nacogdoches, TX 75962; wilsonhe2@sfasu.edu

Frank C. Worrell is a professor of education at the University of California, Berkeley. He is the Program Director of the UC Berkeley School Psychology program and the faculty director of the Academic Talent Development Program. *Contact information:* Cognition and Development, 4511 Tolman Hall, Berkeley, CA 94720-1670; frankc@berkeley.edu