International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Among friends: the role of academic-preparedness diversity in individual performance within a small-group STEM learning environment

Marina Micari, Zachary Van Winkle & Pilar Pazos

To cite this article: Marina Micari, Zachary Van Winkle & Pilar Pazos (2016) Among friends: the role of academic-preparedness diversity in individual performance within a small-group STEM learning environment, International Journal of Science Education, 38:12, 1904-1922, DOI: 10.1080/09500693.2016.1218091

To link to this article: http://dx.doi.org/10.1080/09500693.2016.1218091

Published online: 24 Aug 2016.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
VOL. 38, NO. 12, 1904–1922
http://dx.doi.org/10.1080/09500693.2016.1218091

Routledge
Taylor & Francis Group

Among friends: the role of academic-preparedness diversity in individual performance within a small-group STEM learning environment

Marina Micari[a], Zachary Van Winkle[b,c] and Pilar Pazos[d]

[a]Searle Center for Advancing Learning and Teaching, Northwestern University, Evanston, IL, USA; [b]Department of Microsociology, Humboldt-University Berlin, Berlin, Germany; [c]Demography and Inequality Research Group, WZB Berlin Social Science Center, Berlin, Germany; [d]Department of Engineering Management & Systems Engineering, Old Dominion University, Norfolk, VA, USA

ABSTRACT

In this study, we investigate the relationship between academic-preparedness diversity within small learning groups and individual academic performance in science, technology, engineering, and mathematics (STEM) university courses. We further examine whether academic-preparedness diversity impacts academically more- and less-prepared students differently. We use data from 5367 university students nested within 1141 science, engineering, and mathematics learning groups and use a regression analysis to estimate the effect of group diversity, measured in two ways, on course performance. Our results indicate that academic-preparedness diversity is generally associated with positive learning outcomes, that academically less-prepared students derive greater benefit, and that less-prepared students fare best when they are not alone in a group of highly prepared students. Implications for teaching and small-group facilitation are addressed.

ARTICLE HISTORY
Received 30 January 2016
Accepted 25 July 2016

KEYWORDS
Collaborative learning; small-group learning; higher education; STEM education; learning-group composition

Introduction

With a tremendous push over the past decade toward active learning in science, technology, engineering, and mathematics (STEM) fields, small-group learning has become a common teaching approach at colleges and universities across the U.S.A and abroad. Small-group learning is widely described as one of the best ways to achieve the sort of active learning that results in increased learning gains when compared to traditional lecture-based methods (e.g. Barkley, Cross, & Major, 2004; Bean, 2011; Michael & Modell, 2003). In particular, peer-led small-group learning has been used widely in STEM learning, and a growing number of colleges and universities have introduced peer-facilitated group learning components into their introductory science courses. In some cases, the small-group work happens in the course itself (e.g. Goertzen, Brewe, Kramer, Wells, & Jones, 2011; Otero, Pollock, & Finkelstein, 2010); in others, it happens outside of class (e.g. Born, Revelle, & Pinto, 2002; Drane, Smith, Light, Pinto,

CONTACT Marina Micari m-micari@northwestern.edu Searle Center for Advancing Learning and Teaching, Northwestern University, 627 Dartmouth Place, Evanston, IL 60208, USA

& Swarat, 2005; Lyle & Robinson, 2003; Swarat, Drane, Smith, Light, & Pinto, 2004; Tien, Roth, & Kampmeier, 2002). Typically in such models, small groups of undergraduates meet regularly, along with a more experienced peer leader, to solve problems collaboratively.

Numerous studies of peer-led group learning programs have shown enhanced academic performance and retention (Bonsangue & Drew, 1995; Drane et al., 2005; Drane, Micari, & Light, 2014; Freeman, 1995; Gosser et al., 2001; Lewis & Lewis, 2005; Pazos, Drane, Light, & Munkeby, 2007; Tien et al., 2002). Research has also shown that small-group learning in STEM contributes to increased retention, better relationships among peers, greater self-esteem, and improved attitudes toward the discipline (Drane et al., 2005, 2014; Springer, Stanne, & Donovan, 1999). The use of peer leaders is thought to further promote learning, as peers are not only close to student participants in terms of their own understanding, enabling them to explain concepts at the appropriate level (Vygotsky, 1978), but also serve as role models, providing a vicarious experience of success for participants (Bandura, 1997).

The practice of small-group learning is grounded in constructivist learning theories. The importance of the social environment in learning stems from the work of theorists who frame learning as a social act: interaction between and among individuals plays a critical role in cognitive development (Vygotsky, 1978). Peer learning is thought to work well because it forces students to engage in the cognitive conflict that arises from realizing that others' understandings differ from one's own. When this conflict occurs, learners seek to resolve it, and this process brings a deeper understanding (Piaget, 1977; also see Springer et al., 1999). Students engaging in discussion with peers are also likely to think through ideas carefully, view problems from different perspectives, and rework answers (Myers & Lamm, 1976). Further, students who learn together benefit not only from hearing others' explanations, but also by offering their own (Webb, Farivar, & Mastergeorge, 2002). Explaining material to a peer should be expected to encourage what is termed *cognitive restructuring*, or building meaningful mental relationships among ideas, which engenders future learning and long-term retention of knowledge (Ausubel, 1963).

However, the small-group model is not always effective in producing learning gains, because of variation in the composition of groups, differences in the context of implementation, and interpersonal dynamics within groups. Indeed, there has been – and continues to be – a certain amount of controversy over how and when collaborative learning is most effective (Holen, 2000; Micari & Pazos, 2014; Slavin, 1996; Van den Bossche, Gijselaers, Segers, & Kirschner, 2006). A number of factors have been shown in the research to have some impact on the effectiveness of small-group learning, including the tendency of students to engage in off-task conversation (Hastings & Schweiso, 1995); varied interaction levels (Cohen, 1994; Webb, 1991); differences in the type of task assigned (e.g. problems with well-structured vs. ill-structured solutions; application vs. learning new information) (Chizhik, 1999; Cohen, 1994; Kutnik, Blatchford, & Baines, 2002); and the mixture of preparedness levels in a group (Webb, 1989). It is this last factor that we address in this study.

Before describing our hypotheses and methods in detail, in the next section we provide an overview of the literature on small-group diversity as well as the relationship between learning-group composition and academic performance.

Theoretical considerations

Small-group diversity

In the literature on groups and teams, the term *group diversity* generally refers to the degree to which there are differences of various kinds among members of a group (Harrison & Klein, 2007). Scholars have invested great effort in investigating the impact of diversity on group outcomes, yet there is a general lack of agreement on the issue. There has been growing recognition that the paths linking team diversity to group functioning and group outcomes are complex and dynamic. Part of this complexity results from the many different types of diversity within groups, changes over time in a group, and interactions with other variables such as type of task, type of team, team context, team history, team hierarchy, and member goal orientation, among others (Harrison, Price, Gavin, & Florey, 2002; Pazos & Beruvides, 2011).

Given this complexity, identifying the types of diversity at play in a given situation is critical in investigating the effects of group diversity. There are a number of different kinds of group diversity identified in the literature, including demographic (variation in gender, ethnicity, etc.), psychological (variation in personality, values, etc.), and status diversity (variation in status, tenure, etc.). The type of diversity we are most concerned with in this study has been termed *informational diversity*, and refers to the variation in group members' education, knowledge, and skills (Jehn, Northcraft, & Neale, 1999). Informational diversity has been shown to increase difference of opinion and discussion on task-related issues (Jehn, Chadwick, & Thatcher, 1997), and this level of discussion might be expected to produce the cognitive conflict that leads to more substantive learning (Piaget, 1977; also see Springer et al., 1999).

Additionally, there are several theoretical perspectives that can be used to frame the ways in which groups operate. Some of these (e.g. similarity-attraction theory: see Klein, Lim, Saltz, & Mayer, 2004; Zatzick, Elvira, & Cohen, 2003) focus on what have been termed *surface-level characteristics* (e.g. race and gender), and studies in this vein have typically found that the more similar individuals in a group or dyad are to one another, the more readily they will communicate and engage in team tasks. (Note that the use of the word *surface* is not meant to suggest that such differences represent a superficial level of diversity.) By contrast, the approach we use in this study, the information-processing perspective, frames group diversity in what is termed *deep-level* variation, for example, variation in knowledge, skills, perspectives, and so on (Dahlin, Weingart, & Hinds, 2005; Kearney, Gebert, & Voelpel, 2009; see also Mannix & Neale, 2005). Research taking the information-processing perspective generally finds that diversity has a positive influence on group outcomes through increased access to a wide range of knowledge, skills, abilities, and cognitive perspectives (Jackson & Ruderman, 1995), and that group diversity can have positive consequences on cognitive, affective, and behavioral phenomena at the individual and group levels of analysis (Jackson, May, & Whitney, 1995). From an information-processing perspective, the positive effects of team diversity result from the enriched, elaborated exchange of information and ideas that emerge when individuals with differing knowledge bases and perspectives work together (van Knippenberg & Schippers, 2007). For example, diverse opinions have been shown to allow group members to gather information and combine ideas (Rohrbaugh, 1979), and diversity in skills and expertise has been found to positively affect goal selection, technical problem-solving activities, and execution of the group's task (Jackson & Ruderman, 1995).

Learning-group composition and academic performance

The issue of how to most effectively compose groups in academic settings has long been a topic of inquiry for teachers and researchers (Mello, 1993; Oakley, Felder, Brent, & Elhajj, 2004; Zhou & Pazos, 2014). Do we allow students to self-select into groups, or assign them to groups based on a given criterion? Is it better to group less-prepared students together and more-prepared students together, or to mix levels within a group? Which students – more or less prepared – will do better in academically homogenous vs. academically diverse groups? These questions lack clear answers in the literature.

Some studies have suggested that mixing preparedness levels has a negative effect on learning. (We note that different authors use different terms to describe a student's level of preparedness for a particular course or academic task, including not only *preparedness*, but also *prior knowledge*, *ability*, and so on. We use the term *preparedness* because it captures the student's likelihood for success without suggesting differences in inherent capabilities.) Indeed, in a meta-analysis, Lou, Abrami, and Spence (2010) found that effect sizes for learning gains were larger for groups composed of students with similar levels of preparedness than for groups composed of students with differing levels. This seems to be particularly true for students with low and medium levels of academic preparedness. For instance, Gijlersand de Jong (2005) found that in groups with students at extreme ends of the preparedness scale, less-prepared students retreated from the activity. Beane and Lemke (1971) likewise found that less-prepared students working in academically homogenous foursomes performed better on a transfer task than did less-prepared students working in academically diverse groups. Webb (1980) and Webb and Kenderski (1984) found a similar pattern, with students who had medium levels of preparation participating less actively and receiving less help than their peers in academically diverse groups. This may be a result of students 'in the middle' essentially being left out of the conversation that develops between students with high levels and those with low levels of preparation (Webb, 1995).

Other research, however, finds that less-prepared students learn more effectively in academically diverse groups. Hooper and Hannafin (1988) found that less-prepared students working in academically diverse math study groups performed better on a posttest than did less-prepared students working in homogenous groups. This effect may be particular to less-prepared students: Saleh, Lazonder, and De Jong (2005) found that less-prepared students fared better – in terms of both achievement and motivation – in academically diverse groups, whereas moderately prepared students showed the opposite pattern, performing better in more homogenous groups. In an experiment involving small-group learning in an electricity unit, Webb, Nemer, Chizhik, and Sugrue (1998) found that less-prepared students working in academically diverse groups performed better on a posttest than did those who worked with only other less-prepared students. However, in both studies this pattern was reversed for highly prepared students: they did better in academically homogenous compared to academically diverse groups.

When less-prepared students do perform better in academically diverse groups, their success may be due to the ways in which they interact in those groups. Hooper and

Hannafin (1991) found that less-prepared students interacted more frequently (nearly 30% more) in academically diverse than homogenous groups, while highly prepared students interacted at equal levels in both group conditions. Group discussion may also simply have a larger impact on less-prepared students: Webb et al. (1998) found that the quality of small-group discussion (as measured by quality of explanations and accuracy of answers) predicted posttest scores for less-prepared students, but in most cases *not* for highly prepared students. In fact, for less-prepared students, the quality of group discussion was a greater predictor of the posttest score than was the individual's own pretest score; the reverse was true for highly prepared students.

In short, there is a lack of clarity in the current literature – and very limited literature in the higher education context – as to whether increased diversity in academic preparedness in the group benefits student learning, as well as whether there are differential impacts for different types of students. That is, do academically less-prepared students fare better in homogenous rather than heterogeneous groups? What about highly academically prepared students? This study is an attempt to fill the research gap in this area by empirically estimating the effect of academic-preparedness diversity within a group on individual performance in the context of peer learning environments. From a theoretical perspective, this study can contribute to our understanding of how learning takes place within social contexts, in particular how STEM students' ability to learn may be affected by the presence of others who are similar or different from them in terms of academic level. In effect, this study is an opportunity to critique the assumption that small-group learning is by definition beneficial, by investigating one of the factors that may render it more or less beneficial.

Hypotheses

In line with the information-processing perspective in team research – which views deep-level, knowledge-based diversity as potentially enriching – and drawing from social constructivist views of education – which place cognitive conflict at the center of learning – we would expect academic-preparedness diversity to increase a group's access to a wide range of cognitive approaches to solving problems and thus complexity of discussion resulting from a range of levels of understanding. Therefore, we predict that increased academic-preparedness diversity within learning groups will be associated with increased individual academic performance (*Hypothesis 1*).

The question of whether students with different levels of preparation experience differing effects of within-group academic diversity is also central to this study. Because less-prepared students can be expected to engage to a greater degree in the cognitive restructuring that comes through rich group discussion of problems, we predict that academic-preparedness diversity will be more positively associated with individual academic performance for academically less-prepared students than for more-prepared students (*Hypothesis 2*).

Further, we are interested in whether a less-prepared student's status as the solo less-prepared student, versus being in the company of other less-prepared students, has an impact on that student's course performance. Because being 'solo' in this situation may lead to decreased willingness to engage fully in the group, we predict that students who are the sole less-prepared individuals in their groups will perform more poorly in the course than less-prepared students who are in the company of other less-prepared students within their groups (*Hypothesis 3*).

Data and methods

This study was conducted in the context of Northwestern University's Gateway Science Workshop program, one of the largest and most-researched peer-led team learning programs in the U.S.A (Drane et al., 2014; Light & Micari, 2013). In this program, students enrolled in a range of STEM courses meet weekly, outside of class, in groups of approximately 4–7 throughout the academic quarter with a peer leader – a fellow undergraduate who has already completed and excelled in the associated course. The peer leaders are trained through a two-quarter, credit-bearing course through the institution's school of education. This course focuses on group facilitation and small-group dynamics, as well as social-psychological issues such as student motivation, theories of intelligence, and stereotype threat. The leaders are volunteers, receiving course credit for training but no payment.

In the weekly sessions, participating students work collaboratively to solve advanced problems related to course material. Workshop participation is voluntary, and students receive a zero-credit notation on their transcripts for participating. Participating students are provided with an explanation of the goals of small-group learning at the outset of the term, and participate in discussion about the expectations of participants to contribute productively to a small-group learning experience. Groups are formed according to students' schedules; students sign up for the offered time slot that works best for them. It is possible that friends may sign up together, but on the whole the groups are made up of students who simply have the same weekly time period free. The goal of the program is to enhance students' understanding of crucial concepts, encourage critical thinking, and improve grades and student retention in the associated courses. Previous program evaluation (Drane et al., 2014) shows that students who participate in the program tend to earn higher grades in the associated courses, even after controlling for initial differences.

As part of a summative evaluation of the program, participant characteristics (including grade-point average (GPA) and demographic information) and course grade have been systematically collected. Student data such as course grade and demographic information were obtained through the institution's Registrar; data on group size have been recorded by the program managers. This study was reviewed and received approval from the sponsoring university's institutional review (human subjects) board.

The program data set used in this study contains information on 5367 program participants from the years 2001 to 2006, nested within 1141 workshops.

Variables

The dependent variable is academic performance, as measured by the student's numeric course grade at the end of the quarter being studied. We standardized the grade variable on the grade distribution of the particular course in question to test the robustness of our results. The standardization, as computed by equation in Figure 1, corrects two potentially

$$Std.\,Numeric\ Grade_i = \frac{(Numeric\ Grade_i - Mean\ Numeric\ Grade_j)}{Std.\ Deviation\ of\ Numeric\ Grade_j}$$

Figure 1. Formula for standard numeric grade.

problematic aspects of the unstandardized grade variable: the unstandardized dependent variable is highly negatively skewed, and multiple comparisons tests indicate significant differences among grade distributions by course. However, the regression results deviate only minutely between the model using a standardized and the one using an unstandardized grade. The types of assessments used to determine course grade in this study include problem-solving activities that are very similar to the problems assigned to the peer-led groups. The course instructor typically develops the course assessments as well as the problems solved by the groups.

The academic preparedness diversity variable is based on the group members' prior college achievement, as measured by GPA. While prior achievement is not necessarily equivalent to preparedness for a particular course, it is often a strong predictor of a student's potential academic success (e.g. Harackiewicz, Barron, Tauer, & Elliot, 2002), and thus a reasonable measure of the degree to which a student is prepared to do well in a given course.

There have been a wide variety of mathematical equations used to measure group diversity. Since there is no universal agreement as to which metric is more valid and reliable, this analysis uses two diversity measures to reduce the potential for measurement bias. We describe these two approaches below.

The first indicator of group diversity is the (unbiased) *standard deviation* of students' GPAs within a workshop group. Standard deviation represents the variation in GPA, a continuous variable used as an indicator of academic preparation. This is a measure of academic diversity conceptualized as *separation*, or a difference in position among team members along a continuum (Harrison & Klein, 2007). To account for a group-size bias in diversity measurements (Biemann & Kearney, 2010), we derived an unbiased standard deviation from the uncorrected sample variance using k values (Cureton, 1968; see Figure 2).

The second indicator, the (unbiased) *Blau index*, measures diversity in terms of variety of a categorical attribute, which in this case is a dichotomous variable classifying students either as 'prepared' (if their GPA is in the upper three quartiles) or 'less prepared' (if their GPA is in the lowest quartile). Blau's index conceptualizes diversity as variety, or differences in types or categories, and can theoretically vary from 0 to 1, reaching the maximum when the group consists of equal numbers of individuals within each category – in this case, when a group consists of equal numbers of academically prepared and less-prepared students. An unbiased index was calculated based on the approach recommended by Biemann and Kearney (2010) (see Figure 3).

Analysis

We estimated the effects of group academic diversity on academic achievement (as measured by course grade), controlling for the group and individual specific variables described above, using multivariate linear ordinary least squares regression analysis (hypothesis 1). The

$$SD_j = \sqrt{\frac{s^2{}_j N_j}{k}}$$

Figure 2. Formula for unbiased standard deviation.

$$Blau's\ Index_j = 1 - \sum_{i=1}^{R} \frac{N_i(N_i - 1)}{N_j(N_j - 1)}$$

Figure 3. Formula for Blau's index.

hierarchical nature of the data, that is, individuals being nested within workshop groups, violates the assumption of independent residuals, which can lead to imprecise regression estimates (Moulton, 1986). To remedy this, we clustered standard errors by groups, which allows for the relaxation of the independence assumption and produces more accurate regression estimates (Liang & Zeger, 1986; see Angrist & Pischke, 2009). An alternative solution often used is maximum likelihood hierarchical linear modeling, which provides accurate estimates through the estimation of group-specific error terms; however, in this case, the regression results from the ordinary least squares (OLS) and hierarchical models were equivalent.

An interaction term between group academic diversity and the student's prior academic achievement was used to test whether the impact of academic-achievement diversity within the group on individual grades differs across individual levels of academic achievement (hypothesis 2). *Individual prior academic achievement* was measured by the individual's GPA prior to joining the group.

In hypotheses 1 and 2, we expect that academic diversity will have a positive effect on students' academic achievement, and that this effect will decrease with higher prior student achievement. Therefore, the main effect of diversity is expected to be positive and the interaction term between diversity and prior achievement negative.

We include a number of additional control variables in the analyses that may confound the association between group academic diversity and individual academic performance. The omission of such control variables could bias the estimated effect of group diversity on individual outcomes. A number of studies have shown that, generally speaking, underrepresented minorities (Crisp, Nora,& Taggart, 2009; May & Chubin, 2003) and women (Miyake et al., 2010; Nosek, Smyth, Sriram, & Greenwald, 2009) tend to perform more poorly than majority students and men, respectively, in undergraduate STEM courses. Therefore, we include students' minority status (minority defined as African-American, Hispanic American, or Native American – groups underrepresented in STEM at the study institution) and gender (coded as male or female) in the analyses. We also include an interaction between minority status and gender in the analyses, as there is evidence that minority status female students perform worse than their majority status peers (Espinosa, 2011). Finally, while there is some disagreement about the effect of class size on student achievement, many studies have observed a negative relationship between class size and students' outcomes (Arias & Walker, 2004; Becker & Powers, 2001). Although we are analyzing small learning groups, and not classes, we control for group size as the number of students participating in the workshop group, derived from program data.

Results

The distributions of the individual-level variables are displayed in Table 1; distributions of the group-level variables are displayed in Table 2. As discussed above, students' prior

Table 1. Distribution of individual-level variables.

		Percentage
Gender	Male	41.87
	Female	58.13
Minority status	Majority	90.50
	Minority	9.50
Academic achievement	Prepared	75.44
	Less prepared	24.56
	Mean	Std. Deviation
Numeric course grade	3.098	0.781
Std. numeric grade	−0.009	1.024
Individual GPA	3.411	0.416
N individuals	5367	

academic achievement and individual academic performance are high overall, with an average 3.4 GPA and 3.1 (out of 4) numeric course grade. Average group academic-preparedness diversity is relatively low, as indicated by the mean and range values of Blau's index in Table 2. Using Blau's index, there is an average of just 1 'less-prepared' student per group. However, the actual number of such students ranges between 0 and 6 per group.

The results of the regression analyses are displayed in Table 3. Hypotheses 1 and 2 are tested here. Variables included here are students' prior academic achievement as measured by GPA, within-group academic diversity, and an interaction between these two variables. These are used to predict students' academic performance, as measured by numeric course grade. Academic diversity is measured as the unbiased standard deviation in column 1 and as Blau's index in column 2.

As can be seen in column 1 of Table 3, group academic diversity as measured by the within-group standard deviation in GPA is positively associated with individual academic performance. On average, an increase in group academic diversity by one standard deviation unit is significantly associated with 0.338 higher course grade points. Significant positive effects of within-group academic diversity on individual course grades are also found in column 2, using Blau's index as a diversity measure. The effect strength of group diversity on course grades is similar across the two models. These results support

Table 2. Distribution of group-level variables.

	Mean	Min–Max
Unbiased standard deviation	0.396 (0.198)	0–1.680
Unbiased Blau's index	0.350 (0.266)	0–1
Group size	5.825 (1.712)	2–9
N Workshops	1141	
		Percentage
Group diversity category	Homogenous	25.82
	1 less-prepared	34.12
	1 prepared	3.80
	Mixed group	36.26

Note: Standard deviation between groups in parentheses.

Table 3. OLS linear regression results of academic preparedness and group diversity on individual performance.

	Unbiased standard deviation	Unbiased Blau's index
GPA	1.230**	1.228**
	(44.17)	(37.04)
Group diversity	0.338**	0.201**
	(6.230)	(4.471)
GPA*Group diversity	−0.423**	−0.366**
	(−3.914)	(−2.857)
Ethnic identity	−0.367**	−0.369**
(Ref.: Majority)	(−5.671)	(−5.702)
Gender	−0.105**	−0.107**
(Ref.: Male)	(−6.238)	(−6.325)
Ethnic identity*Gender	0.182*	0.184*
	(2.477)	(2.500)
Group size	−0.004	−0.003
	(−0.696)	(−0.578)
Constant	3.172**	3.158**
	(221.7)	(200.6)
Observations	5367	5367
Adjusted R^2	0.424	0.419

Notes: $^{**}p < 0.01$, $^{*}p < 0.05$; Unstandardized coefficients displayed; robust *t*-statistics in parentheses; controls omitted from table; Variables centered.

our first hypothesis that increased academic-achievement diversity within the group is associated with increased individual academic performance.

The interaction effects between academic diversity and GPA are negative in both columns of Table 3. As seen in column 1, the effect of academic diversity on course grades (0.338) drops by 0.423 points to −0.085 for students with a one-point above-average GPA. This effect is displayed graphically in Figure 4, in which course grades are estimated for students with a high (4.0), medium (3.0), and low (2.0) GPA in workshops with minimum (standard deviation of 0) to maximum observed (standard deviation of 1.5) group academic diversity. While the predicted grades of students with high GPAs do not differ significantly between less- and more-academically diverse groups, the predicted grades of students with a medium GPA are 0.86 grade points higher in groups with maximum diversity compared to homogenous groups. The impact is even greater for students with low GPAs: estimated course grades are more than one grade point higher for students in groups with maximum observed diversity compared to homogenous groups. While there may be a ceiling effect at play here (i.e. students with higher GPAs tend to have high grades and so have little room to improve), these results provide some support for our second hypothesis, that academic achievement diversity will be more positively associated with individual academic performance of academically less-prepared students than that of academically more-prepared students.

As indicated by the explained variance (R^2) and the *t*-statistics of the academic diversity and interaction coefficients in columns 1and 2 of Table 3, the differences between the diversity measurements are minute when analyzing individual course grades. Both models are capable of explaining some 42% of course grade variation. In sum, the analysis results of academic diversity's association with individual course grade are highly consistent across the two diversity measurements.

In hypothesis 3, we expect that less-prepared students will fare better in groups in which they are with other less-prepared students; in other words, not among exclusively

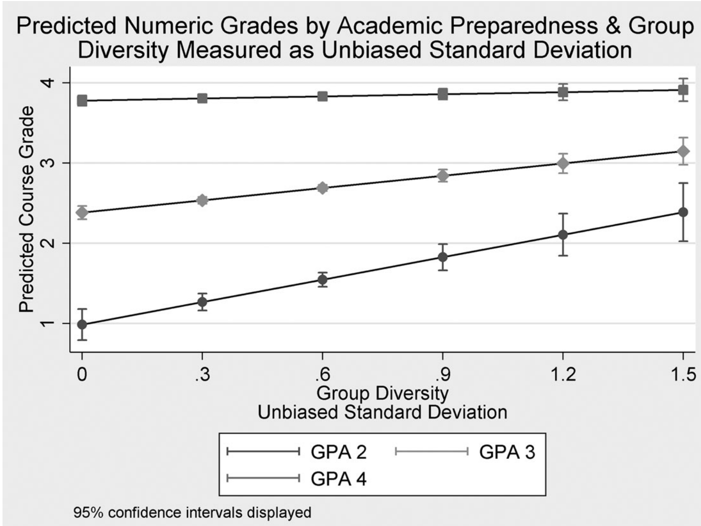

Figure 4. Predicted grades by preparedness and group diversity.

highly prepared students. For this analysis, we classified groups as being (a) homogenously prepared (all students in the top three GPA quartiles), (b) all prepared with the exception of one less-prepared (lowest quartile GPA) student, (c) all less-prepared students with the exception of one prepared student, or (d) mixed groups with more than two students of each category. We then estimated the effects of these group composition categories and interacted these effects with students' GPAs (see Table 4). As in the previous analysis, while all three types of diversity showed a positive effect as compared to homogenous groups, that effect becomes weaker for higher GPA students, with the strongest effect in mixed groups. For instance, the effect of group composition on course grades within a mixed group (0.123) drops by 0.176 points to −0.053 (main effect for mixed group minus interaction effect) for students with a one-point above-average GPA, whereas the mixed-group composition effect increases to 0.299 (main effect for mixed group plus interaction effect) for students with a one-point below-average GPA.

We used this regression model to predict grades for more- and less-prepared students within the four types of group composition. As can be seen in Figure 5, predicted course grades for less-prepared students alone in a group of prepared students are not statistically different from the predicted course grades of less-prepared students in homogenous (that is, consisting of exclusively less-prepared students) learning groups. On average, predicted course grades are statistically highest for less-prepared students in groups that include just one highly prepared student or in mixed groups – in other words, where they are not the solo less-prepared student.

Regarding our control variables, we find in accordance with the literature that minority-status students have lower course grades than majority-status students. Even after adjusting for differences in GPA, gender, and group size, African-American, Hispanic, and Native American students' course grades are 0.3 points lower than their White and Asian peers (see Tables 3 and 4). Also in line with existing literature, our adjusted data indicate that women have slightly lower grades than men, with statistical significance.

Table 4. OLS linear regression results of academic advantage and group diversity categories on individual performance.

GPA	1.302**
	(16.49)
Group diversity category (Ref.: Homogenous)	
1 Less-prepared student	0.0872**
	(2.831)
1 Prepared student	0.108+
	(1.755)
Mixed Group	0.123**
	(3.990)
GPA*Group diversity category	
GPA*1 Less-prepared student	−0.0629
	(−0.705)
GPA*1 Prepared student	−0.313*
	(−2.555)
GPA*Mixed group	−0.176*
	(−1.993)
Ethnic Identity	−0.369**
(Ref.: Majority)	(−5.740)
Gender	−0.108**
(Ref.: Male)	(−6.373)
Ethnic Identity*Gender	0.186*
	(2.528)
Group size	−0.0133*
	(−2.060)
Constant	3.092**
	(105.2)
Observations	5367
Adjusted R^2	0.419

Notes: ** $p < 0.01$, * $p < 0.05$, + $p < .1$ (note that sample size for groups with one prepared student was below 5% of the total); Unstandardized coefficients displayed; robust *t*-statistics in parentheses; controls omitted from table; Variables centered.

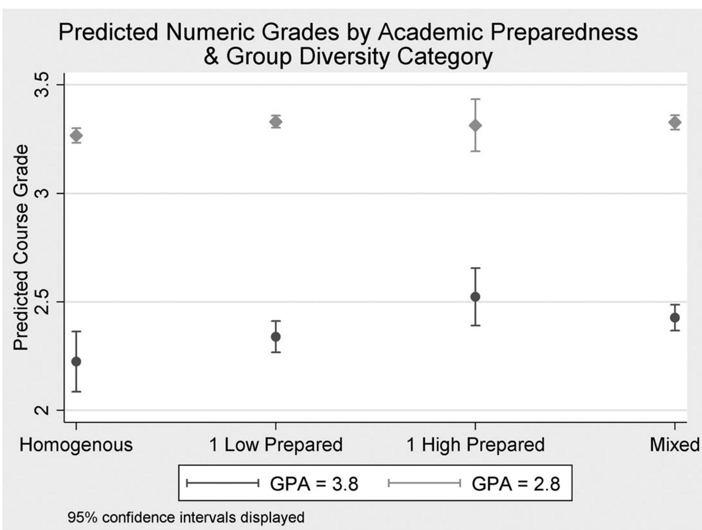

Figure 5. Predicted grades by preparedness and group diversity category.

However, there is no statistical difference between the adjusted course grades of male and female minority-status students. Finally, in our first two models (see Table 3), we find no significant association between group size and individual outcomes; however, we do find a significant negative relationship in our last model, consistent with the literature on class sizes (see Table 4). Thus, there is some evidence here that smaller group sizes, where presumably students are getting more individual attention, may be associated with better student outcomes.

Discussion

This research contributes to our understanding of the role of diversity in group learning by investigating the impact of group academic-preparedness diversity on individual academic achievement in the context of peer-led team learning groups.

The results of this study suggest benefits of increased within-group diversity on academic performance for all students (hypothesis 1), with greater benefit for students with lower levels of academic preparedness (hypothesis 2). We further found that less-prepared students fared better when they were in groups that include a mixture of other less-prepared students and prepared students (hypothesis 3), rather than being the lone less-prepared student in the group.

This study is in line with earlier studies by Hooper and Hannafin (1988, 1991) and Webb et al. (1998), finding that less-prepared students fared better in academically diverse study groups than in academically homogenous study groups. In other words, academically disadvantaged students seem to do poorly in groups where, like them, all other students lack background knowledge, *and* in groups where all other students are far ahead of them in background knowledge.

What might be the mechanism behind the greater benefit of academic diversity for less-prepared students? It may be that in academically diverse groups, the more advanced students are scaffolding (Wood, Bruner, & Ross, 1976) others' learning by operating within the less-advanced students' 'zone of proximal development,' defined by Vygotsky (1978) as the imaginary space between a student's real level of skill or development and her or his potential level; students can reach their potential levels with guidance from teachers or 'in collaboration with more capable peers' (p. 86). Students who are somewhat – but not drastically, as with a professor – more advanced in their understanding are uniquely positioned to provide the support that less-advanced members need, offering guidance in terms that are understandable but which still require learners to stretch beyond their current capabilities. Less-prepared students may also benefit more from hearing others' explanations than do more-prepared students, who may derive the most benefit from explaining to others (Saleh et al., 2005).

The finding that less-prepared students do better in groups that include at least one other less-prepared student suggests that there may be something about the company of similar others that boosts performance of less-prepared students. Micari and Drane (2011) found that feeling intimidated by others' (perceived) superior knowledge within a learning group had a negative effect on students' academic self-efficacy and academic performance. A lone disadvantaged student may retreat out of this kind of intimidation (Micari & Drane, 2011; Micari & Pazos, 2014) or be left behind in discussion, much like the students at the low end of the preparation scale in Gijlersand de Jong's (2005) study. Academic diversity may in fact have a negative impact when that diversity is perceived as threatening to the academic self-concept of lower achieving students (Brophy, 2005; Dijkstra, Kuyper, Van der Werf, Buunk, & Van der Zee, 2008; Mugny, Butera, & Falomir, 2001). As the current study did not measure perceived threat to self-concept, it is unclear whether there may be any difference between groups with a more egalitarian feel and groups with greater perceived divisions between 'advanced' and 'less-advanced' students.

It is important to note that all students may benefit from group diversity in ways not captured by course grade. For instance, interacting with others whose perspectives and experiences differ from one's own has been shown to have a positive effect on such factors as complexity of thinking, self-confidence, social agency, and more (Antonio et al., 2004; Laird, 2005).

Practical considerations

The results of this study can guide decision-making of teachers, faculty members, and others who determine learning-group composition. Our findings suggest not only that learning groups may work best when composed of students with different levels of preparation in the subject, but also that less-prepared students should *not* be isolated in groups of highly prepared students. However, if there is such an imbalance in groups, group facilitators – whether teachers or peers – can be used to help mitigate its potentially negative effects. For instance, facilitators can be trained to pay close attention to interaction patterns in groups, and to draw out students who appear to be retreating as well as keeping more dominant members at bay (see Jaques & Salmon, 2007; Light & Micari, 2013; Roth, Goldstein, & Mancus, 2001). Facilitators can also use techniques to lessen the anxiety that might be produced when less-prepared students are in the minority within a group, and by extension the impairment in cognitive processing ability that such anxiety can produce (Eysenck, Derakshan, Santos, & Calvo, 2007). Micari and Pazos (2014), for instance, reported success in using peer group facilitators to reduce students' concerns about being academically 'behind' others in the group, through an intervention focusing on malleability of intelligence (Dweck, 1999) and on normalizing struggle as a part of learning.

Limitations

Because the study context was limited to science and mathematics learning groups at the college level, the results should not be generalized to humanities or social sciences fields, or to education at the K-12 level. The study also looked only at groups working on solving traditional problems (that is, not ill-structured, real-world problems) within a classroom context, and findings should not be generalized to other kinds of group learning contexts. Course grades and GPA are the most commonly used proxy for academic achievement in quantitative studies, largely due to their demonstrated connections to academic persistence (York, Gibson, & Rankin, 2015). One possible limitation of these widely used metrics is their limited validity for measuring learning or growth in cognitive capabilities (Arum & Roksa, 2011). Further, because the institution in which the study was conducted is highly selective – in other words, only students with top secondary-school academic records are admitted to the institution – academic diversity within the learning groups is narrower in range than what might exist in a more moderately selective university, or in a high school setting. Put plainly, even individuals at the extreme low end of academic achievement in this study are highly academically talented students. It may be that these students had enough background knowledge and learning 'savvy' to be able to make productive use of the knowledge of their more-advanced peers. Students lacking in such background knowledge, however, might feel so far behind their more-advanced peers in a learning group, and be so unable to engage productively in discussion, that the experience would have a negative effect on their learning and performance. We also note that there may have been a ceiling effect for more highly prepared students, putting the findings for hypothesis 2 somewhat into question. Finally, while a strength of this study is the consistent findings regardless of the measure of diversity used, both of those measures are based on students' GPA. A study triangulating the academic-preparedness-diversity construct by using additional measures of academic performance would offer enhanced construct validity.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Angrist, J. D., & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Antonio, A., Chang, M., Hakuta, K., Kenny, D., Levin, S., & Milem, J. (2004). Effects of racial diversity on complex thinking in college students. *Psychological Science*, *15*(8), 507–510.

Arias, J. J., & Walker, D. (2004). Additional evidence on the relationship between class size and student performance. *The Journal of Economic Education*, *35*(4), 311–329.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Ausubel, D. (1963). *The psychology of meaningful verbal learning*. New York, NY: Grune & Stratton.

Bandura, A. (1997). *Self-efficacy, the exercise of control*. New York, NY: Freeman.

Barkley, E., Cross, P., & Major, C. (2004). *Collaborative learning techniques: A handbook for college faculty*. San Francisco, CA: Jossey-Bass.

Bean, J. (2011). *Engaging Ideas, the professor's guide to integrating writing, critical thinking, and active learning in the classroom*. San Francisco, CA: Jossey-Bass.

Beane, W., & Lemke, E. (1971). Group variables influencing the transfer of conceptual behavior. *Journal of Educational Psychology*, *62*(3), 215–218.

Becker, W. E., & Powers, J. R. (2001). Student performance, attrition, and class size given missing student data. *Economics of Education Review*, *20*(4), 377–388.

Biemann, T., & Kearney, E. (2010). Size does matter: How varying group sizes in a sample affect the most common measures of group diversity. *Organizational Research Methods*, *13*(3), 582–599.

Bonsangue, M., & Drew, D. (1995). Mathematics, opening the gates-increasing minority students' success in calculus. In J. Gainen & E. Willemsen (Eds.), *Fostering student success in quantitative gateway courses. New Directions for Teaching and Learning, 61* (pp. 23–33). San Francisco, CA: Jossey-Bass series.

Born, W., Revelle, W., & Pinto, L. (2002). Improving biology performance with workshop groups. *Journal of Science Education and Technology*, *11*(4), 347–365.

Brophy, J. (2005). Goal theorists should move on from performance goals. *Educational Psychologist*, *40*(3), 167–176.

Chizhik, A. W. (1999). Can students work together equitably? An analysis of task effects in collaborative group work. *Social Psychology of Education*, *3*, 63–79.

Cohen, E. (1994). Restructuring the classroom, conditions for productive small groups. *Review of Educational Research*, *64*(1), 1–35.

Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, *46*(4), 924–942.

Cureton, E. E. (1968). Unbiased estimation of the standard deviation. *The American Statistician*, *22* (1), 22.

Dahlin, K., Weingart, L., & Hinds, P. (2005). Team diversity and information use. *The Academy of Management Journal*, *48*(6), 1107–1123.

Dijkstra, P., Kuyper, H., Van der Werf, G., Buunk, A. P., & Van der Zee, Y. (2008). Social comparison in the classroom, a review. *Review of Educational Research*, *78*(4), 828–879.

Drane, D., Micari, M., & Light, G. (2014). Students as teachers, effectiveness of a peer-led STEM learning program over 10 years. *Educational Research and Evaluation*, *20*(3), 210–230.

Drane, D., Smith, H. D., Light, G., Pinto, L., & Swarat, S. (2005). The gateway science workshop program, enhancing student performance and retention in the sciences through peer facilitated discussion. *Journal of Science Education and Technology*, *4*(3), 227–352.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Espinosa, L. (2011). Pipelines and pathways: Women of color in undergraduate STEM majors and the college experiences that contribute to persistence. *Harvard Educational Review*, *81*(2), 209–241.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*, 336–353.

Freeman, M. (1995). *Mathexcel, a special opportunity in calculus*. Unpublished report, Department of Mathematics, University of Kentucky.

Gijlers, H., & de Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of Research in Science Teaching*, *42*(3), 264–282.

Goertzen, R., Brewe, E., Kramer, L., Wells, L., & Jones, D. (2011). Moving toward change: Institutionalizing reform through implementation of the learning assistant model and open source tutorials. *Physical Review Special Topics – Physics Education Research*, *7*(2), 020105.

Gosser, D. K., Cracolice, M. S., Kampmeier, J. A., Roth, V., Strozak, V. S., & Varma-Nelson, P. (2001). *Peer-led team learning, a guidebook*. Upper Saddle River, NJ: Prentice Hall.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, *94*(3), 562–575.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228.

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal*, *45*, 1079–1084.

Hastings, N., & Schweiso, J. (1995). Tables and tasks: The effects of seating arrangements on task engagement in primary classrooms. *Educational Research*, *37*(3), 279–291.

Holen, A. (2000). The PBL group, self-reflections and feedback for improved learning and growth. *Medical Teacher*, *22*, 485–488.

Hooper, S., & Hannafin, M. J. (1988). Cooperative CBI, the effects of academically diverse versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research*, *4*, 413–424.

Hooper, S., & Hannafin, M. J. (1991). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Educational Technology Research and Development*, *39*(3), 27–40.

Jackson, S. E., May, K. E., & Whitney, K. (1995). Under the dynamics of diversity in decision-making teams. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 204–261). San Francisco: Jossey-Bass.

Jackson, S. E., & Ruderman, M. N. (1995). *Diversity in work teams: Research paradigms for a changing workplace*. Washington, DC: American Psychological Association.

Jaques, D., & Salmon, G. (2007). *Learning in groups*. New York, NY: Routledge.

Jehn, K. A., Chadwick, C., & Thatcher, S. M. B. (1997). To agree or not to agree: The effects of value congruence, individual demographic dissimilarity, and conflict on workgroup outcomes. *International Journal of Conflict Management*, *8*(4), 287–305.

Jehn, K. A., Northcraft, G., & Neale, M. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative Science Quarterly*, *44*, 741–763.

Kearney, E., Gebert, D., & Voelpel, S. C. (2009). When and how diversity benefits teams: The importance of team members' need for cognition. *The Academy of Management Journal Archive*, *52*(3), 581–598.

Klein, K., Lim, B., Saltz, J., & Mayer, D. (2004). How do they get there? An examination of the antecedents of centrality in team networks. *Academy of Management Journal, 47*, 952–963.

van Knippenberg, D. L., & Schippers, M. (2007). Work group diversity. *Annual Review of Psychology*, *58*, 515–541.

Kutnik, P., Blatchford, P., & Baines, E. (2002). Pupil groupings in primary school classrooms: Sites for learning and social pedagogy? *British Educational Research Journal*, *28*(2), 187–206.

Laird, N. (2005). College students' experiences with diversity and their effects on academic self-confidence, social agency, and disposition toward critical thinking. *Research in Higher Education*, *46*, 365–387.

Lewis, S., & Lewis, J. (2005). Departing from lectures: An evaluation of a peer-led guided inquiry alternative. *Journal of Chemical Education*, *82*(1), 135.

Liang, K., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22.

Light, G., & Micari, M. (2013). *Making scientists*. Cambridge, MA: Harvard University Press.

Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *Journal of Educational Research*, *94*(2), 101–112.

Lyle, K. S., & Robinson, W. R. (2003). A statistical evaluation: Peer-led team learning in an organic chemistry course. *Journal of Chemical Education*, *80*(2), 132–134.

Mannix, E., & Neale, M. (2005). What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest*, *6*(2), 31–55.

May, G., & Chubin, D. (2003). A retrospective on undergraduate engineering success for underrepresented minority students, *Journal of Engineering Education*, *92*(1), 27–39.

Mello, J. A. (1993). Improving individual member accountability in small group settings. *Journal of Management Education*, *17*, 253–259.

Micari, M., & Drane, D. (2011). Intimidation in small learning groups: The roles of social-comparison concern, comfort, and individual characteristics in student academic outcomes. *Active Learning in Higher Education*, *12*, 175–187.

Micari, M., & Pazos, P. (2014). Worrying about what others think, a social comparison–concern intervention in small learning groups. *Active Learning in Higher Education*, *15*(3), 249–262.

Michael, J., & Modell, H. (2003). *Active learning in secondary and college science classrooms: A working model for helping the learner to learn*. Mahwah, NJ: Lawrence Erlbaum.

Miyake, A., Kost-Smith, L., Finkelstein, N., Pollock, S., Cohen, G., & Ito, T. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science*, *330*(6008), 1234–1237.

Moulton, B. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, *32*, 385–397.

Mugny, G., Butera, F., & Falomir, J. M. (2001). Social influence and threat in social comparison between self and source's competence, relational factors affecting the transmission of knowledge.

In F. Butera & G. Mugny (Eds.), *Social influence in social reality* (pp. 225–248). Gottingen: Hogrefe and Huber.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*, 602–627.

Nosek, B., Smyth, F., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597.

Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning*, *2*(1), 9–34.

Otero, V., Pollock, S., & Finkelstein, N. (2010). A physics department's role in preparing physics teachers: The Colorado learning assistant model. *American Journal of Physics*, *78*(11), 1218–1224.

Pazos, P., & Beruvides, M. (2011). Incorporating training and feedback into the study of patterns in group decision making: The impact of communication medium. *Team Performance Management*, *17*(1/2), 83–101.

Pazos, P., Drane, D., Light, G., & Munkeby, A. (2007). *A peer-led team learning program for freshmen engineering students, Impact on retention*. Proceedings of American Society for Engineering Education Conference, Honolulu, HI, USA.

Piaget, J. (1977). *The development of thought, equilibration of cognitive structures*. New York, NY: Viking Press.

Rohrbaugh, J. (1979). Improving the quality of group judgment: Social judgment analysis and the Delphi technique. *Organizational Behavior and Human Performance*, *24*, 73–92.

Roth, V., Goldstein, E., & Mancus, G. (2001). *Peer-led team learning: A handbook for team leaders*. Upper Saddle River, NJ: Prentice Hall.

Saleh, M., Lazonder, A. W., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science*, *33*(2), 105–119.

Slavin, R. (1996). Research on cooperative learning, what we know and what we need to know. *Contemporary Educational Psychology*, *21*, 43–69.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). *Effects of small-group learning on undergraduates* in science, mathematics, engineering, and technology, a meta-analysis. *Review of Educational Research*, *69*(1), 21–51.

Swarat, S., Drane, D., Smith, H. D., Light, G., & Pinto, L. (2004). Opening the gateway, increasing student retention in introductory science courses. *Journal of College Science Teaching*, *34*(1), 18–23.

Tien, L., Roth, V., & Kampmeier, J. (2002). Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching*, *39*(7), 606–632.

Van den Bossche, P., Gijselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments team learning beliefs and behaviors. *Small Group Research*, *37*(5), 490–521.

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

Webb, N. M. (1980). An analysis of group interaction and mathematical errors in heterogeneous ability groups. *British Journal of Educational Psychology*, *50*, 266–276.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, *13*, 21–39.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*, 366–389.

Webb, N. M. (1995). Group collaboration in assessment, multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis*, *17*(2), 239–61.

Webb, N. M., Farivar, S. H., & Mastergeorge, A. M. (2002). Productive helping in cooperative groups. *Theory Into Practice*, *41*, 13–20.

Webb, N. M., & Kenderski, C. M. (1984). Student interaction and learning in small group and whole class settings. In P. L. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of* *instruction, group organization and group processes* (pp. 153–170). New York, NY: Academic Press.

Webb, N. M., Nemer, K. M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment, group composition and performance. *American Educational Research Journal*, *35*, 607–651.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology*, *17*(2), 89–100.

York, T. T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research & Evaluation*, *20*(5), 1–20.

Zatzick, C. D., Elvira, M. M., & Cohen, L. E. (2003). When is more better? The effects of racial composition on voluntary turnover. *Organization Science*, *14*, 483–496.

Zhou, Z., & Pazos, P. (2014). *Managing engineering capstone design teams: A review of critical issues and success factors*. Proceedings of the Industrial and Systems Engineering Research Conference, Montreal, PQ, Canada.