December, 1999

Making Use of Difference: Diversity, Debate, and Decision Comprehensiveness in Top Management Teams

Tony L. Simons, Cornell University School of Hotel Administration
Lisa Hope Pelled, University of Southern California
Ken A. Smith, Syracuse University

Available at: https://works.bepress.com/tony_simons/13/
This study examined how top management team diversity variables and debate interacted to influence two measures of company financial performance. Further, it assessed the degree to which decision comprehensiveness mediated those interaction effects. Multi-informant data from the top management teams of 57 manufacturing companies revealed that more job-related types of diversity interacted with debate to influence financial performance, but a less job-related type (age diversity) did not. Decision comprehensiveness partially mediated those effects.

Recent reviews of the extensive group diversity literature have noted extremely mixed empirical results and, accordingly, have described diversity as a two-edged sword (Milliken & Martins, 1996) or as a mixed blessing (Williams & O'Reilly, 1998). Although many of the studies discussed in these reviews investigated lower-level work groups, equivocal findings are characteristic of top management team (TMT) diversity studies as well. For example, Roure and Keeley (1990) found that top managers' functional background diversity predicted firm financial performance, but West and Schwenk (1996) found no relationship between TMT diversity and performance. Smith, Smith, Olian, Sims, O'Bannon, and Scully (1994) found that TMT educational diversity was positively associated with company financial performance. However, in the same study, experience diversity was negatively associated with performance. In the present study, we endeavored to clarify such equivocal findings. We developed a theory about when TMT diversity will have primarily positive effects on firm performance and when it will have primarily negative effects. We then tested that theory in a sample of 57 top management teams.
Previous attempts to explain group diversity effects have generally suggested that process variables such as communication mediate relationships between diversity and performance outcomes (e.g., Ancona & Caldwell, 1992; Felled, Eisenhardt, & Xin, 1998; Smith et al., 1994). Although these theories helped move diversity research in an important direction, additional steps are needed, as empirical support for mediated explanations has been relatively modest (Lawrence, 1997). In the current study, we moved beyond prior work by exploring the additional possibility that team process moderates the relationship between TMT diversity and firm performance. We propose that debate among team members moderates the effect of TMT diversity on firm performance, with interactions between debate and more job-related forms of diversity having stronger effects than interactions between debate and less job-related forms of diversity. We then propose that decision comprehensiveness will mediate those interactive effects on performance.

CONCEPTUAL BACKGROUND AND HYPOTHESES

Key Concepts

*Top management team diversity.* A core component of our theory is TMT diversity, defined here as the extent to which a top management team is heterogeneous with respect to members' demographics and cognitions. Recently, scholars have begun to identify important distinctions among kinds of diversity. Accordingly, our theory differentiates diversity types by the degree to which they are directly related to the task at hand. *Job relatedness* is the extent to which a type of diversity captures distinct experiences, skills, or perspectives relevant to cognitive tasks at work (Felled, 1996). Job relatedness is an important property because it determines whether a particular type of diversity constitutes an increase in a group's total pool of task-related skills, information, and perspectives. The magnitude of this pool, in turn, represents a potential for more comprehensive or creative decision making (Milliken & Martins, 1996).

In the present study, we examined four kinds of demographic diversity that tend to exist in top management teams: diversity in functional background, education level, tenure, and age. Of these demographic variables, functional, education level, and tenure diversity are generally classified as more job-related because they largely capture experiences, information, and perspectives relevant to cognitive tasks (Pelled, 1996; Williams & O'Reilly, 1998). Although age diversity may capture experiences and perspectives relevant to the workplace, those
experiences constitute a small portion of the total set of experiences and perspectives age diversity captures (Zenger & Lawrence, 1989). Thus, age diversity is less job-related than the other three demographic variables.

Although this study's primary focus was demographic diversity, it also included a type of nondemographic diversity: perceived environmental uncertainty (PEU) diversity. PEU diversity is the degree to which TMT members differ in their perceptions about the level of uncertainty in their company's external business environment. Such perceptions constitute a critical input for strategic decisions (Calori, Johnson, & Sarnin, 1994). Further, they relate more directly than demographic measures to team members' cognitions about their decision context. Hence, PEU diversity is perhaps the most job-related variable in this study. Since our theory addresses the degree to which diversity variables are job-related, it was important to ensure that the diversity variables in this study represented a range of job relatedness. Including PEU diversity would help determine whether job-relatedness played the role our theory suggested.

**Debate.** Several researchers have suggested that team process should be considered a moderator of the link between TMT diversity and performance (e.g., Ancona & Caldwell, 1989; Hambrick, 1994). Their argument is that diversity must be effectively expressed and integrated into decisions if a group is to derive benefit from it. If diversity is not integrated in this manner, then the coordination costs or other social costs that accompany diversity will drive performance down. Thus, as mentioned above, we examined the role of team process—specifically, debate—as a moderator in this study. **Debate** is defined here as an open discussion of task-related differences and the advocacy, by different top management group members, of differing approaches to the strategic decision-making task (Schweiger, Sandberg, & Rechner, 1989). Debate is similar to the construct **controversy**, which Tjosvold (1985) suggested as a form of interaction in which group members have opposing views, preferences, or approaches when solving a problem or reaching a decision. In the context of top management groups, the affective tone of debate may be negative, but it is not necessarily so. This distinction separates debate as defined here from many group process measures of conflict, which—as Sessa (1996) observed—have often relied on negative affective tone as part of the definition of conflict. Although debate has some overlap with **task conflict** as defined by Jehn (1995), the two concepts are not identical. Questions used to measure debate focus on specific behaviors rather than on a general perception of difference or conflict.

**Decision comprehensiveness.** The process variable proposed as a mediator in this study is **decision comprehensiveness**, which Fredrickson defined as "the extent to which
organizations attempt to be exhaustive or inclusive in making and integrating strategic decisions" (1984: 445). Both academic and popular writing has suggested that decision comprehensiveness is a highly influential process in top management teams (e.g., Miller, Burke, & Glick, 1998; Taylor, 1987). In accord with this notion, we expected comprehensiveness to play a critical role in the relationship between diversity-by-debate interactions and performance.

Although debate and decision comprehensiveness may be related, they are distinct constructs. Debate is a process in which team members challenge and oppose one another on task issues. Decision comprehensiveness, in contrast, is a process in which team members look at an issue with a wide lens, considering multiple approaches, multiple courses of action, and multiple decision criteria. Debate may sometimes encourage members to weigh alternatives and take a broader, more open-minded approach to decision making (to be more comprehensive), but it is also possible for members to have debate without decision comprehensiveness. Members may propose different approaches but never compare those approaches and weigh them against each other. Members may also disagree with each other's opinions without offering substantive reasons for that disagreement—that is, without seriously discussing pros and cons. By the same token, it is possible to have decision comprehensiveness without having debate. For example, members may brainstorm about possible alternatives and systematically list pros and cons without ever disputing one another's ideas.

**Performance.** Because top management decisions can have a critical impact on firm success, team performance has typically been conceived of as firm performance in studies of top management teams (e.g., Eisenhardt & Schoonhoven, 1990; Smith *et al.*, 1994). Following this tradition, we conceptualized TMT performance as increased firm profitability and sales.

**Hypotheses**

A central argument of this article is that teams may need debate to make constructive use of diversity. Our first hypothesis advances the notion that debate will thus moderate the impact of team diversity on both team process (decision comprehensiveness) and outcome (performance) variables.

According to this argument, diversity represents a potential for more thoughtful decision making. Quality of decision making is difficult to measure in real-world groups but may be estimated by decision comprehensiveness. Debate may increase the tendency for diversity to enhance decision comprehensiveness. Without debate, a team's diversity may remain an
untapped resource, existing but never used. If team members debate each other, however, they are likely to draw on their diversity—that is, on their divergent knowledge sets—to bolster their arguments. As they are confronted by new information from people with different backgrounds, members are forced to rethink their points of view and consider factors they had not previously considered. In this manner, the decision-making process becomes more comprehensive.

Similar reasoning suggests that debate will moderate the impact of diversity on company performance. In the absence of debate, a TMT may not be able to draw on the diverse experiences of its members to make decisions that optimize performance. For example, a team member from the marketing area may have seen previous cases suggesting a particular course of action is flawed but, without conflictual discussion, other members may never hear about those cases—and thus, they may make costly mistakes. The need for teams to use a debate process to bring out the positive potential of diversity is especially critical given that diversity can lead to difficulties of communication and coordination (Williams & O'Reilly, 1998).

Debate can also moderate the performance impact of diversity through the experience of voice (group members' sense that they have actively participated in a decision). Without debate, members of a heterogeneous team may infer that their expertise has not been considered. As a result, they may feel bypassed and frustrated. Hence, even if the team makes a good choice, these members may not fully accept and implement that choice (Werther, 1988). Such poor implementation can impair company performance (Beer, Eisenstat, & Spector, 1990; Bryant, 1997). Thus, debate may be necessary to unleash the performance benefits of TMT diversity.

Hypothesis 1. Debate will moderate the impact of TMT diversity, increasing the tendency for diversity to enhance decision comprehensiveness and performance.

It is possible that the performance impact described above emerges primarily through the effect of debate and diversity on decision comprehensiveness. Debate and diversity together may influence the comprehensiveness of a decision-making process, which in turn affects company performance. Most empirical research has indicated that, particularly in turbulent industries, the comprehensiveness of strategic decision processes is positively related to firm performance (e.g., Boyd, 1991; Miller & Cardinal, 1994). As Lant, Milliken, and Batra (1992) suggested, a decision process involving considerable investigatory activity helps prevent "groupthink" (Janis, 1982). Along the same lines, Eisenhardt (1989: 558) noted that considering multiple alternatives "reduces the escalation of commitment to any one option
"We therefore propose that decision comprehensiveness acts as a vehicle through which debate-by-diversity interactions enhance performance.

Hypothesis 2. Decision comprehensiveness will mediate the interactive effects of diversity and debate on company performance.

One must remember, however, that diversity is not a unitary construct. Different kinds of diversity may interact with debate to influence decision comprehensiveness and firm performance in different ways. Some types of diversity—those that are high on the job-relatedness dimension—by definition capture heterogeneous information and perspectives that are highly relevant to organizational performance. As debate occurs between TMT members who are diverse in job-related ways, the discussion is likely to include a wide range of task-relevant perspectives and information; hence, subsequent process and performance gains are apt to be particularly high. If, however, debate is driven by less job-related forms of diversity, then the expressed differences between TMT members may ultimately have little to do with accomplishing the task at hand. Consequently, the debate is more likely to incorporate irrelevant opinions or thoughts. Team process and performance gains may then be only modest or negligible. Thus, measures of TMT diversity that are more job-related should, upon interacting with TMT debate, have the greatest impact on decision comprehensiveness and company financial performance.

Hypothesis 3. Interactions between more job-related forms of diversity and debate will have stronger positive associations with decision comprehensiveness and company performance than will interactions between less job-related forms of diversity and debate.

METHODS

Sample and Data Collection

Our analysis was conducted with multi-informant data from 57 electronic components manufacturing firms. Data were collected in 1992 as part of a broader research study on the structure and functioning of TMTs. We solicited participation in the overall study by mailing questionnaire packets to the CEOs of 925 electronic components manufacturing companies selected from industry association directories.

The packets included three copies of a questionnaire for TMT members and a company financial information survey. The items used for the present study constituted only a portion of
the whole questionnaire, which was eight pages long. Because the overall questionnaire placed heavy demands on top managers' time, we elected to follow a sampling plan that balanced minimizing these demands with meeting the data requirements of the study. We believed that in so doing, we increased the likelihood of participation. Thus, we asked CEOs to complete and return one survey and to ask two fellow TMT members, defined as individuals who actively participated in strategic decision making, to do the same. This sampling approach is similar to that of Tsui, Egan, and O'Reilly (1992), who chose to reduce the burden on the organizations in their study by sampling to obtain a minimum of 20 employees from each work unit, the average size of which was 834. This approach is also similar to that of Smith and colleagues (1994), who relied on CEOs to identify members of TMTs in an effort to capture Cyert and March's (1963) concept of the dominant coalition. Including the CEO—arguably the most influential member of a TMT—and two other managers whom the CEO identified as playing important roles in strategic decision making helped ensure that our sampling plan captured the most relevant data effectively. Eligibility for participation in the study was limited to companies that had not experienced turnover of the CEO or of two or more other senior executives within the previous year.

Measures of financial performance were collected in a separate survey instrument that requested self-reported, objective (numeric) financial data. Each CEO was instructed that the financial survey should be completed by him- or herself or by the company's chief financial officer (CFO).

Responses were received from 83 companies, and complete data were received from 57 of these. This response rate of approximately 6 percent is, although somewhat low by general standards, similar to that in other TMT research that has used multiple questionnaires rather than archival data only (e.g., West & Schwenk, 1996). A response rate of less than 10 percent can result in sample selection bias (Pedhazur & Schmelkin, 1991). However, in studies that assess complex relationships among measured variables, such as the current one, sample selection bias is unlikely to pose a threat (Simons, 1995). For selection bias to affect the validity of a moderated linear analysis, a sample idiosyncrasy would have to alter a moderated relationship to form a three-way interaction. Such an effect is improbable.

We only analyzed teams from which we received two or more surveys and financial information. Although we asked for three executive responses from each company, a comparison between the number of responses received and the reported total top team size showed that our sample represented almost 50 percent of all top management team members.
for the participating companies. The average top management team size for our sample companies was 6.06, and the median size was 5. We received an average of 2.84 responses (range = 2-5, s.d. = .68) from the companies included in the sample. Although each TMT was originally sent only three surveys, five companies requested or photocopied additional surveys.

Measures

*TMT diversity.* The top management team questionnaire was the source of data used to compute the diversity measures in this study. The questionnaire asked each respondent to estimate the ages of the oldest and youngest members on his/her team. The difference between these two numbers was the age range on the team, and the mean of the two numbers estimated the mean age on the team. Dividing the first figure (age range) by the second figure (estimated age mean) approximates a coefficient of variation. We averaged team members’ coefficient of variation estimates to obtain a measure of team age diversity. To assess the appropriateness of aggregating respondents’ estimates in this manner, we computed an eta-square statistic, which indicates whether respondents from a single group are more likely to have similar responses than respondents from different groups. Eta-square was .41, which exceeds Georgopoulos’s (1986: 40) criterion for aggregation of .20.

The questionnaire asked respondents to report how long they had been working for their companies. *Company tenure diversity* was then computed as the coefficient of variation of company tenure among TMT members.

Respondents were asked to indicate which of the following categories reflected their functional specialty: finance, human resources, general management, marketing, operations, research and development, or strategic planning. We then computed functional background diversity using Blau’s index \(1 - \sum i^2\), where \(i\) is the proportion of respondents in the \(i\)th category.

The questionnaire asked respondents to report the highest educational degree they had attained. To be consistent with the measure used by Smith and colleagues (1994), we transformed highest degree into years of formal education and then computed the group’s coefficient of variation to estimate education-level diversity.

To calculate perceived environmental uncertainty (PEU) diversity, we crafted a version of Duncan’s (1972) scale. Survey respondents were asked to indicate whether or not each of
ten environmental factors (product technology, manufacturing technology/equipment costs, client needs, new materials/product part costs, labor supply/costs, cost of capital, number of competitors, behavior of competitors, government regulation, and access to new markets) had undergone a major (greater than 10 percent), nonpredictable and noncyclic change during the 1991 calendar year. For each team, we summed the number of environmental factors that were marked by at least one member as "changing" and by at least one member as "not changing." For example, if (1) one TMT member indicated that product technology had changed but client needs had not, (2) at least one other member indicated product technology had not changed but client needs had, and (3) there was no divergence on any of the other eight items, then the sum would equal 2. We then statistically controlled this sum for the number of respondents on a team (as more respondents meant more opportunity for discrepancy), and we saved the residuals as the PEU diversity variable.

**TMT debate.** To form a team-level index of debate, we first computed each individual's perceptions of debate using a Likert-type scale and then averaged those perceptions. Each team member was asked to consider one recent, important strategic decision that his or her team had made and to rate the level of debate that occurred using four items. These items are the top four shown in Table 1. We did not specify a particular decision so as to let general patterns of team interaction emerge. In essence, we sampled from the pool of recent team decisions. Following the suggestions and findings of previous researchers (Fredrickson & Iaquinto, 1989; Tushman & Romanelli, 1985), in our measurement of debate we assumed that work groups' approaches to decision making tend to be relatively consistent over time and that recent TMT discussion patterns therefore reflect historical behavior patterns.

Cronbach's alpha for this four-item scale was .75. After adding responses to these items to obtain individual-level debate perception scores, we averaged individual-level scores within each team to obtain team-level scores. The eta-square for debate was .46, which exceeds Georgopoulos's criterion.

**Decision comprehensiveness.** As with TMT debate, we aggregated individual perceptions of the comprehensiveness of the described decisions. Each team member was asked to describe the specified decision process (the one for which he or she had described the level of debate) in terms of comprehensiveness, using three scale items (the lower three items shown in Table 1). This scale, which was based on Miller's (1990) comprehensiveness measure, had a Cronbach's alpha of .78. The eta-square for decision comprehensiveness was .39, which indicates that aggregation was statistically permissible.
A principal components analysis with "varimax" rotation suggested that debate and
decision comprehensiveness were distinct constructs. Results of that analysis are shown in
Table 1. Each of the two eigen-values was above one, and the two factors accounted for 64.1
percent of the variance in responses.

**Organizational performance.** Following previous TMT research (cf. Murray, 1989;
Smith et al., 1994), we measured organizational performance using profitability and sales
growth data reported in the financial information survey. We computed change scores, as
change in profitability or sales is likely to be more closely linked to recent behavior than overall
profitability or sales, which are more heavily influenced by distant historical factors.

*Change in profitability* was measured as the change in gross profit expressed as a
percentage of sales between 1990 and 1991. We first asked respondents to indicate 1991 gross
profit, before executive draw, as a percentage of sales. Next, we asked whether gross profit in
1991 was higher or lower than 1990 gross profit, and by how many percentage points. Then, to
obtain our measure, we standardized this change-in-profitability figure by dominant product
line. Specifically, we divided our sample according to dominant product line (passive
components, semiconductors, or printed circuit boards), and we computed the mean and
standard deviation of change in profitability across all teams having the same product line. We
then used these figures to standardize the change-in-profitability score of each team. For
example, we standardized change in profitability for a semiconductor team by using the mean
and standard deviation of change in profitability for all semiconductor teams.

*Change in sales* was the change in gross sales from 1990 to 1991 expressed as a
percentage of 1990 sales. Respondents were asked to report sales growth using the following
scale: 1, down more than 25 percent; 2, down 15-25 percent; 3, down 5-15 percent; 4, down
1-5 percent; 5, no change; 6, up 1-5 percent; 7, up 5-15 percent; 8, up 15-25 percent; 9, up
over 25 percent. As with change in profitability, we standardized change in sales by dominant
product line.

To assess the validity of these self-reported performance measures, we used Disclosure
Global Access to find financial information on the 13 publicly traded, stand-alone companies in
the sample. Working from these data, we computed change in profitability and change in sales
in the same manner asked of the respondents. Our calculated change in profitability correlated
with the self-reported data at .70 (p < .01, two-tailed test), and our calculated sales growth
correlated with the self-reported figure at .61 (p < .05, two-tailed test). These correlations
indicated that our self-reported financial performance measures approximated the financial data available from third parties.

**Control variables.** We controlled for team size, company age, company size, and team tenure because these variables have frequently been identified as factors that can influence strategic decision processes and outcomes (e.g., Hambrick & D'Aveni, 1992; Miller et al., 1998).

*TMT size* was the total number of managers on a company's top management team reported on the CEO's questionnaire. *Company size* was the number of full-time employees or equivalents reported on the financial information survey by the CEO or CFO. *Company age* was the years since founding reported on the financial information survey. *TMT tenure* was the average number of years TMT members had belonged to their teams, as reported on their own questionnaires.

**Analysis**

To test Hypothesis 1, which predicts that debate will moderate the effects of TMT diversity on performance and decision comprehensiveness, we regressed performance and decision comprehensiveness on the control variables, main effect variables, and interaction terms in sequential steps.

To test Hypothesis 2, our prediction about the mediating effect of decision comprehensiveness, we followed a procedure Baron and Kenny (1986: 1177) outlined, (1) regressing the mediator on the independent variables, (2) regressing the dependent variable on the independent variables, and (3) regressing the dependent variable on both the mediator and independent variables. The following conditions indicate mediation: the independent variables affect the mediator in the first equation; the independent variables affect the dependent variable in the second equation; the mediator affects the dependent variable in the third equation; and the effects of the independent variables on the dependent variable are weaker in the third equation than in the second equation (Baron & Kenny, 1986).

To test Hypothesis 3, our prediction about the relative magnitudes of interactive effects, we examined the equations used to test Hypothesis 1. We compared the more job-related diversity-by-debate interactive effects with the less job-related diversity-by-debate interactive effect (age diversity by debate). "Pairwise" comparisons of interactions were made with Kmenta's (1980) one-tailed *t*-test for comparing coefficients within a single regression equation.
RESULTS

Table 2 shows the means, standard deviations, and correlations among study variables. The largest correlation among predictor variables was .64, and the median and mean correlations were .03 and .05, respectively. These magnitudes suggest that multi-collinearity was not a serious problem in this study (Kennedy, 1980; Tsui, Ashford, St. Clair, & Xin, 1995).

Table 3 presents the results of the multiple regression analyses used to test the hypotheses. A significant measure of support was found for Hypothesis 1, which predicts that interactions between diversity and debate will have positive associations with decision comprehensiveness and with company performance. As models 1 and 2 revealed, 9 out of 15 interaction terms achieved marginal or full significance.

The results of separate, controlled hierarchical regressions for each of the individual debate-by-diversity interactions were largely consistent with the results of the combined models. The most noteworthy exception was that the interaction between debate and functional background diversity was very strongly associated with changes in profitability in a separate regression equation ($R^2 = .18$, $p < .001$) but was reduced to marginal significance in a pooled regression with other debate-by-diversity interactions. This pattern suggests a degree of substitutability between functional background diversity and one or more of the other diversity types.

Hypothesis 2, which predicts that decision comprehensiveness will act as a mediator, received partial support. As shown in model 2 of Table 3, the regression of decision comprehensiveness on predictors revealed that the interaction of debate and education diversity was positively associated with decision comprehensiveness ($\beta = .20$, $p < .10$), as was the debate-by-tenure-diversity interaction ($\beta = .36$, $p < .01$). As shown in model 3, decision comprehensiveness predicted change in profitability ($\beta = .37$, $p < .01$) and change in sales ($\beta = .42$, $p < .01$). A comparison of the interactions in model 1 with the interactions in model 3 shows support for a partial mediating effect. Adding decision comprehensiveness as a predictor in the performance regressions diminished the effects of debate crossed with education diversity and debate with tenure diversity on change in profitability and change in sales. This result suggests that decision comprehensiveness partially mediates the effects on firm performance of the interactions of debate and education diversity and of debate and tenure diversity.
Hypothesis 3 predicts that interactions between more job-related forms of diversity and debate will have stronger positive effects on decision comprehensiveness and firm performance than will interactions between less job-related forms of diversity and debate. This hypothesis received considerable support when firm performance was the dependent variable. A visual inspection of model 1 (Table 3) revealed that each of the more job-related diversity variables interacted with debate to have significant and positive effects on at least one performance measure. The interaction between PEU diversity and debate had significant associations with change in profitability ($\beta = .27, p < .05$) and change in sales ($\beta = .38, p < .01$). Similarly, the interaction between education-level diversity and debate had significant associations with change in profitability ($\beta = .34, p < .05$) and change in sales ($\beta = .29, p < .05$), and the interaction between functional background diversity and debate had marginally significant associations with change in profitability ($\beta = .22, p < .10$) and change in sales ($\beta = .20, p < .10$). The interaction between tenure diversity and debate had a marginally significant association with change in sales ($\beta = .21, p < .10$). In contrast, the less job-related type of diversity, age diversity, did not have a significant, positive interaction with debate affecting either performance outcome.

A visual inspection of model 2 (Table 3) revealed that several of the more job-related diversity types interacted with debate to have significant, positive effects on decision comprehensiveness. The education-level-diversity-by-debate interaction had a marginally significant effect ($\beta = .20, p < .10$), and tenure diversity by debate had a significant effect ($\beta = .36, p < .01$). The less job-related age diversity, however, did not have a significant interaction with debate.

Kmenta's (1980) t-test tentatively confirmed that many of the coefficients of the interactions between debate and more job-related forms of diversity were significantly more positive than the coefficient for the debate-by-age-diversity interaction. When change in profitability was the dependent variable, the difference was significant for debate crossed with education-level diversity ($t = 2.51, p < .01$), and it was marginally significant for debate with PEU diversity ($t = 1.50, p < .10$) and debate with functional background diversity ($t = 1.56, p < .10$); however, the t-test was not significant for debate with tenure diversity ($t = 1.15, \text{n.s.}$). For change in sales as the dependent variable, a similar pattern of contrasts between the more job-related diversity types and age diversity emerged: the contrast was significant for the debate-by-education-level-diversity interaction ($t = 2.09, p < .05$) and for debate by tenure diversity ($t = 1.74, p < .05$), and it was marginally significant for debate by PEU diversity ($t =$
1.64, \( p < .10 \)) and for debate by functional background diversity (\( t = 1.66, \ p < .10 \)). Thus, in the case of company performance, the statistical comparison of coefficients was consistent overall with the visual inspection results and lent further support to Hypothesis 3. In the case of decision comprehensiveness, however, the statistical comparison did not corroborate the visual findings: the effects of the interactions of the more job-related diversity variables and debate were not significantly different from the effects of the interaction of age diversity and debate.

**DISCUSSION**

We set out to address the inconsistency of findings in previous TMT diversity studies by proposing that debate moderates diversity's effects on performance and that decision comprehensiveness mediates these debate-by-diversity effects. Our data revealed that debate increased the tendency for diversity to enhance TMT performance. Further, debate-by-diversity interactive effects were strongest for more job-related forms of diversity. Decision comprehensiveness partially mediated these interactive effects.

The observed positive interactive effects of debate and diversity on performance are a particularly critical finding. Although recent studies have focused on process variables as mediators of diversity effects, this finding suggests that the interactions of such process variables and diversity are also significant and that process may thus be usefully viewed as a moderator of diversity effects. The conclusion to be drawn from these results is that for diversity to benefit a company's bottom line, there must be a process by which the positive aspects of diversity are brought to bear.

Interestingly, the interactive effects of debate and diversity on performance were stronger for several more job-related types of diversity than for one less job-related type (age diversity). PEU diversity, education-level diversity, and company tenure diversity lacked main effects on performance, yet all of these variables had significant moderated effects (as debate-by-diversity interactions) on performance. Functional background diversity also had a significant moderated effect, but age diversity did not. Apparently, debate is more likely to be fruitful when it draws on different experiences and perspectives that are relevant to a task rather than on less relevant viewpoint differences. This finding provides some preliminary support for the notion that more job-related forms of diversity have greater potential impact on organizational performance, as Pelled (1996), Milliken and Martins (1996), and Williams and O'Reilly (1998) have suggested.
Although the primary contribution of our study is its theory and evidence that team process may act as a moderator of TMT diversity effects, this phenomenon does not preclude the possibility that team process acts as a mediator, as emphasized in prior research. Indeed, our findings indicated that decision comprehensiveness partially mediated the effects on performance of the interactions of debate and education diversity and debate and tenure diversity. It is surprising that this mediating effect was not present in the case of debate and PEU diversity or of debate and functional background diversity. One possibility is that other, unmeasured, mediators, such as decision acceptance-the degree to which TMT members "buy into" a team's decision-may account for these interactive effects on performance.

Taken together, the results of this study point to the importance of considering both the potential moderating and the potential mediating roles of team process in the effects of diversity on performance. In particular, the moderating role of debate and the mediating role of decision comprehensiveness can help account for the performance consequences of diversity. Further, our findings demonstrate the utility of treating diversity as a multifaceted construct whose different facets interact with team process to shape performance in different ways.

This study has implications for practitioners. First, it suggests the wisdom of fostering diversity in top management teams, for such diversity is apt to help TMT members’ debates become more constructive. Second, it reinforces the notion that team members may do well to strive for comprehensive decision making, which can enhance firm performance. Finally, it indicates that group diversity must be matched by appropriate processes for the group to benefit.

Like any investigation, ours has several limitations. First, the sample was in a high-technology manufacturing industry, so its results might not generalize to low-technology or service industry settings. Second, the response rate among the contacted companies was low. Although it is unlikely that sample selection bias drove these results, they should be replicated for more confident generalization. Third, our sample included, on the average, less than 50 percent of the TMT members from each company. Although our sampling procedure was designed to capture the dominant coalition in each team, it may nevertheless be a source of some inaccuracy in our measures. Finally, the study’s cross-sectional design means that we cannot be absolutely certain that the causality of the relationships among variables is in the direction that our theory suggests. Addressing these limitations is a worthwhile endeavor for future research.
Other opportunities for future research exist as well. For example, additional process variables can be considered as potential moderators and mediators of TMT diversity effects. It would also be interesting to assess the degree to which debate moderates the effects of diversity on other outcome variables, such as turnover, which has frequently been a dependent variable in TMT studies (e.g., Jackson, Brett, Sessa, Cooper, Julin, & Peyronnin, 1991; Pfeffer & O'Reilly, 1987). Another important step is to see how properties of diversity other than job relatedness shape diversity's interactive effects. As organizational scholars continue to look beyond archival data in their studies of top management teams, they are likely to find a rich set of dynamics open to exploration.
<table>
<thead>
<tr>
<th>Item</th>
<th>Debate</th>
<th>Decision Comprehensiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>In discussions of this issue, executives stated clear disagreement with each other.</td>
<td>.71</td>
<td>-.05</td>
</tr>
<tr>
<td>Different executives proposed different approaches to the issue.</td>
<td>.01</td>
<td>.15</td>
</tr>
<tr>
<td>Executives openly challenged each other’s opinions.</td>
<td>.28</td>
<td>.19</td>
</tr>
<tr>
<td>Discussions of the issue became heated.</td>
<td>.76</td>
<td>-.06</td>
</tr>
<tr>
<td>To what extent did the group weigh multiple approaches against each other?</td>
<td>.03</td>
<td>.80</td>
</tr>
<tr>
<td>To what extent did the group examine the pros and cons of several possible courses of action?</td>
<td>.03</td>
<td>.62</td>
</tr>
<tr>
<td>To what extent did the group use multiple criteria for eliminating possible courses of action?</td>
<td>.09</td>
<td>.86</td>
</tr>
<tr>
<td><strong>Eigenvalue</strong></td>
<td>2.57</td>
<td>1.91</td>
</tr>
<tr>
<td><strong>Percentage of variance explained</strong></td>
<td>36.8</td>
<td>27.3</td>
</tr>
<tr>
<td><strong>Cumulative percentage of variance explained</strong></td>
<td>36.8</td>
<td>64.1</td>
</tr>
</tbody>
</table>

*Values shown bold are those defining a factor.*
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>s.d.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. TMT size</td>
<td>6.06</td>
<td>2.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Company age</td>
<td>4.52</td>
<td>1.46</td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Company size</td>
<td>22.47</td>
<td>15.70</td>
<td>.64*</td>
<td>.37**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. TMT tenure</td>
<td>4.34</td>
<td>0.69</td>
<td>-.20</td>
<td>-.26*</td>
<td>-.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. TMT debate</td>
<td>11.59</td>
<td>1.97</td>
<td>.23</td>
<td>-.08</td>
<td>.24</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. PEU diversity&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.09</td>
<td>2.10</td>
<td>-.05</td>
<td>.11</td>
<td>-.10</td>
<td>.22</td>
<td>-.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Education-level diversity</td>
<td>2.32</td>
<td>0.78</td>
<td>-.24</td>
<td>-.35**</td>
<td>-.31*</td>
<td>.00</td>
<td>-.11</td>
<td>-.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Functional background diversity</td>
<td>0.09</td>
<td>0.07</td>
<td>.02</td>
<td>-.01</td>
<td>-.20</td>
<td>.03</td>
<td>-.15</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Company tenure diversity</td>
<td>0.59</td>
<td>0.36</td>
<td>.14</td>
<td>.12</td>
<td>-.02</td>
<td>-.39**</td>
<td>-.22</td>
<td>.16</td>
<td>-.11</td>
<td>.29*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Age diversity</td>
<td>0.40</td>
<td>0.16</td>
<td>.30**</td>
<td>.06</td>
<td>-.02</td>
<td>-.22</td>
<td>-.17</td>
<td>-.11</td>
<td>.16</td>
<td>.17</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Change in profitability</td>
<td>-.01</td>
<td>1.01</td>
<td>.08</td>
<td>-.21</td>
<td>.03</td>
<td>.04</td>
<td>.07</td>
<td>-.09</td>
<td>-.16</td>
<td>-.22</td>
<td>-.18</td>
<td>-.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Change in sales&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.01</td>
<td>1.02</td>
<td>.05</td>
<td>-.15</td>
<td>.10</td>
<td>.00</td>
<td>.05</td>
<td>.12</td>
<td>.01</td>
<td>-.24</td>
<td>-.12</td>
<td>-.06</td>
<td>.58**</td>
<td></td>
</tr>
<tr>
<td>13. Decision comprehensiveness</td>
<td>9.34</td>
<td>1.79</td>
<td>.03</td>
<td>.03</td>
<td>.17</td>
<td>.06</td>
<td>.47**</td>
<td>-.16</td>
<td>-.20</td>
<td>.07</td>
<td>-.13</td>
<td>-.08</td>
<td>.22</td>
<td>.31*</td>
</tr>
</tbody>
</table>

<sup>a</sup> N = 57.

<sup>b</sup> Standardized and residual variables do not have zero mean values because of "listwise" case deletion.

<sup>c</sup> Value is standardized.

* p < .05

** p < .01
### TABLE 3
Results of Hierarchical Regression Analysis for Firm Performance and TMT Decision Comprehensiveness*

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Change in Profitability</td>
<td>Change in Sales</td>
<td>Decision Comprehensiveness</td>
</tr>
<tr>
<td>Team size</td>
<td>.35†</td>
<td>-.06</td>
<td>-.21</td>
</tr>
<tr>
<td>Company age</td>
<td>.00</td>
<td>.07</td>
<td>.04</td>
</tr>
<tr>
<td>Company size</td>
<td>-.22</td>
<td>.15</td>
<td>.12</td>
</tr>
<tr>
<td>Team tenure</td>
<td>-.04</td>
<td>-.06</td>
<td>-.08</td>
</tr>
<tr>
<td><strong>ΔR²</strong></td>
<td>.10</td>
<td>.06</td>
<td>.06</td>
</tr>
<tr>
<td>F</td>
<td>1.43</td>
<td>.74</td>
<td>.61</td>
</tr>
</tbody>
</table>

**Step 2: Main effects**

- **Debate**
  - .25
  - .20
  - .47**
  - .07
  - .00

- **PEU diversity**
  - -.07
  - .14
  - -.15
  - -.01
  - .21

- **Education-level diversity**
  - .23
  - .08
  - -.30*
  - .34*
  - .20

- **Functional background diversity**
  - -.23*
  - -.34*
  - .06
  - -.20
  - -.31*

- **Company tenure diversity**
  - -.10
  - -.07
  - .92
  - .00
  - .06

- **Age diversity**
  - -.29*
  - -.04
  - .11
  - -.33*
  - .08

- **ΔR²**
  - .10
  - .07
  - .03
  - .10
  - .07

- **Partial F**
  - .04
  - .09
  - 2.47*
  - .94
  - .59

- **β**
  - .20
  - .12
  - .29
  - .20
  - .12

- **F**
  - 1.13
  - .03
  - 1.86*
  - 1.13
  - .03

**Step 3: Interactions**

- **Debate × PEU diversity**
  - .27*
  - .38**
  - .91
  - .27*
  - .38**

- **Debate × education-level diversity**
  - .34*
  - .29*
  - .20*
  - .23*
  - .21*

- **Debate × functional background diversity**
  - .22*
  - .20*
  - .30*
  - .18*
  - .18*

- **Debate × tenure diversity**
  - -.02
  - .21*
  - .36**
  - -.15
  - .06

- **Debate × age diversity**
  - -.20
  - -.24
  - -.94
  - -.10
  - -.22

- **ΔR²**
  - .27
  - .25
  - .13
  - .27
  - .25

- **Partial F**
  - 4.11**
  - 3.29*
  - 1.84
  - 4.11**
  - 3.29*

- **β**
  - .47
  - .37
  - .42
  - .47
  - .37

- **F**
  - 2.38*
  - 1.62
  - 1.97*
  - 2.38*
  - 1.62

**Step 4: Mediator**

- **Decision comprehensiveness**
  - .37**
  - .42**

- **ΔR²**
  - .08
  - .10

- **Partial F**
  - 7.13*
  - 7.08**

- **β**
  - .55
  - .47

- **F**
  - 3.01**
  - 2.25*

---

*All standardized regression coefficients (βs) are from the final step in the hierarchical regression. N = 57. One-tailed tests were used for interactive and mediator effects, which are directionally predicted in the hypotheses. Model 1 includes two regressions, one for each financial performance measure. Model 2 includes one regression for decision comprehensiveness. Model 3 includes two regressions, one for each performance measure. Hypotheses 1 and 3 are tested with models 1 and 2. Hypothesis 2 is tested with models 1, 2, and 3.

* PEU is perceived environmental uncertainty.

† p < .10
* p < .05
** p < .01
REFERENCES


