Exploring the Black Box: An Analysis of Work Group Diversity, Conflict, and Performance

Lisa Hope Pelled
University of Southern California
Kathleen M. Eisenhardt
Stanford University
Katherine R. Xin
University of Southern California

If we cannot now end our differences, at least we can help make the world safe for diversity.
—John F. Kennedy

In this paper we present an integrative model of the relationships among diversity, conflict, and performance, and we test that model with a sample of 45 teams. Findings show that diversity shapes conflict and that conflict, in turn, shapes performance, but these linkages have subtleties. Functional background diversity drives task conflict, but multiple types of diversity drive emotional conflict. Race and tenure diversity are positively associated with emotional conflict, while age diversity is negatively associated with such conflict. Task routineness and group longevity moderate these relationships. Results further show that task conflict has more favorable effects on cognitive task performance than does emotional conflict. Overall, these patterns suggest a complex link between work group diversity and work group functioning.

In the past decade, demographic diversity has become one of the foremost topics of interest to managers and management scholars. The term demographic diversity refers to the degree to which a unit (e.g., a work group or organization) is heterogeneous with respect to demographic attributes. Attributes classified as demographic generally include "immutable characteristics such as age, gender, and ethnicity; attributes that describe individuals' relationships with organizations, such as organizational tenure or functional area; and attributes that identify individuals' positions within society, such as marital status" (Lawrence, 1997: 11). The heightened concern with demographic diversity (hereafter referred to simply as diversity) stems not only from the growing presence of women and minorities in the work force (Buhler, 1997) but also from modern organizational strategies that require more interaction among employees of different functional backgrounds (e.g., Dean and Snell, 1991). One of the most significant bodies of research to arise from this trend is a stream of field studies linking group composition to cognitive task performance—i.e., performance on tasks that involve generating plans or creative ideas, solving problems, or making decisions. The impact of diversity on cognitive task performance has been examined in studies of top management teams (e.g., Bantel and Jackson, 1989; Murray, 1989; Eisenhardt and Schoonhoven, 1990) and lower-level work groups (e.g., Kent and McGrath, 1969; Murnighan and Conlon, 1991).

Despite this spotlight on diversity in work groups, there is more to be done. Investigations of diversity and work group performance have largely been what Lawrence (1997) referred to as "black box" studies, which do not measure intervening process variables. Further, the effects on performance are still unclear. Some studies (e.g., Bantel and Jackson, 1989) have linked diversity to favorable performance on cognitive tasks, and some (e.g., Murnighan and Conlon, 1991) have linked it to unfavorable performance on such tasks. Others (e.g., Watson, Kumar, and Michaelson, 1993) have shown that group diversity both enhances and diminishes cognitive task performance. To

0001-8392/99/4401-0001/$1.00.

We thank Deborah Ancona, Blake Ashforth, Domenic Colabella, Christopher Earley, Connie Gersick, Jean Kahwajy, Gerald Ledford, Nandini Rajagopalan, Robert Sutton, and Allen Weiss for their helpful input. We are especially grateful for the valuable feedback provided by Daniel Brass, Linda Johanson, and three anonymous reviewers. This research was funded by the Stanford Integrated Manufacturing Association.

capture fully the complex relationship between work group diversity and performance, we need more sophisticated theories and empirical work incorporating intervening variables and multiple types of diversity. The objective of the current investigation is to begin to meet these needs, offering an intervening process theory—one that attempts to untangle the complicated set of relationships among five types of diversity and performance—and providing a test of that theory.

Two prior studies that have empirically assessed whether process variables intervene between group diversity and performance are particularly important to our efforts. Ancona and Caldwell (1992) examined the intervening role of internal task process (i.e., the setting of goals and priorities) and external communication. Later, Smith et al.'s (1994) top management team study looked at three potential intervening variables (social integration, informality of communication, and communication frequency). The authors of both studies discovered that the process variables they measured did not fully explain the observed effects of diversity on performance, and both then suggested that the mediating effect of conflict should be assessed in future research. Hence, in the model we propose and test here, conflict plays an intervening role.

The model proposes that work group diversity indirectly affects cognitive task performance through two kinds of conflict: intragroup task conflict and intragroup emotional conflict. Task conflict is a condition in which group members disagree about task issues, including goals, key decision areas, procedures, and the appropriate choice for action, and emotional conflict is a condition in which group members have interpersonal clashes characterized by anger, frustration, and other negative feelings (Jehn, 1994; Eisenhardt, Kahwajy, and Bourgeois, 1997a).[1] We suggest that job-related types of diversity largely drive task conflict. In contrast, emotional conflict is shaped by a complex web of diversity types that increase emotional conflict based on stereotyping and decrease emotional conflict based on social comparison. Task routineness and group longevity moderate these diversity-conflict relationships. Each type of diversity indirectly affects performance via its relationship with conflict: task conflict tends to enhance performance, while emotional conflict tends to diminish performance. Thus, we offer a model that postulates that the black box between diversity and performance contains a more elaborate set of relationships than previously thought.

THEORETICAL BACKGROUND AND HYPOTHESES

Link between Diversity and Task Conflict

When the members of a work group have different demographic backgrounds, they may have dissimilar belief structures (Wiersema and Bantel, 1992), i.e., priorities, assumptions about future events, and understandings of alternatives (Hambrick and Mason, 1984: 195), based on previous training and experiences. As Eisenhardt, Kahwajy, and Bourgeois (1997b: 48) recently noted, executives "who have grown up in sales and marketing typically see opportunities and issues

1

These two types of conflict have been given a variety of labels, such as substantive and affective conflict (Guetzkow and Gyr, 1954; Pelled, 1996), cognitive and affective conflict (Amason, 1996), substantive conflict and interpersonal conflict (Eisenhardt, Kahwajy, and Bourgeois, 1997a), and task and emotional conflict (Jehn, 1994). Although they have used different labels, these studies have offered similar definitions for the two dimensions, essentially describing the same constructs.

from vantage points that differ from those who have primarily engineering experience." Such distinct perspectives may stem, in part, from resource allocation and reward disparities (Donnellon, 1993), which encourage contrasting views of what is important. Due to their respective belief structures, group members with different demographic backgrounds may have divergent preferences and interpretations of tasks (Dearborn and Simon, 1958; Walsh, 1988; Waller, Huber, and Glick, 1995). These divergences are likely to manifest themselves as intragroup task conflict. As diversity within a work group increases, such task conflict is likely to increase. Increased diversity generally means there is a greater probability that individual exchanges will be with dissimilar others. Members are more likely to hear views that diverge from their own, so intragroup task conflict may become more pronounced.

While any type of diversity may trigger task conflict, some are more likely to do so than others, based on the relevance of their corresponding belief structures. People hold multiple belief structures about a variety of information domains (Walsh, 1988); those belief structures most relevant to the information processing task at hand tend to influence interpretation of that task (Wickens, 1989; Waller, Huber, and Glick, 1995). Thus, demographic attributes corresponding to highly relevant belief structures should be especially influential in the perception of work group tasks. The job-relatedness of a demographic attribute is the degree to which that attribute captures experiences and skills germane to cognitive tasks at work (Pelled, 1996; Pelled, Cummings, and Kizilos, 1999). If work group members differ with respect to a demographic attribute that is low in job-relatedness, then their divergent experiences and knowledge may not pertain to the work they do, and opposing task perceptions may not emerge in the group. If work group members differ with respect to a highly job-related demographic attribute, however, then their divergent experiences and knowledge are apt to be pertinent to the task, and incongruent task perceptions are likely to emerge. Diversity with respect to highly job-related attributes is therefore apt to have a stronger relationship with task conflict than is diversity with respect to less job-related attributes.

Functional background and tenure are highly job-related attributes. Both are defined by one's workplace experiences, specifically, whether one is exposed to a particular functional area and how much time one has worked for a company. Also, cognitive tasks in organizations typically demand the experience and knowledge obtained through exposure to functional areas and organizational tenure. Ancona and Caldwell (1992) noted that for tasks such as those of product development teams, functional background and company tenure are likely to be particularly important because they determine one's technical skills, information, expertise, and one's perspective on an organization's history. Others (Sessa and Jackson, 1995; Milliken and Martins, 1996; Pelled, 1996) have similarly argued that functional background and tenure are especially pertinent to work group tasks.

Age, gender, and race, in contrast, are low in job-relatedness. In a recent editorial, one scholar even went as far as to argue, "there is no such thing as a woman's approach to mathematics or an African American approach to physics" (Heriot, 1996: M5). A less extreme and perhaps more realistic assertion is that of Zenger and Lawrence (1989: 357): "Although age similarity may produce similarity in general attitudes about work . . . , such attitudinal similarity is unlikely to have much direct bearing on conversations about technical work." The same logic applies to race and gender. For example, the attribute *race* tends to capture a broad collection of experiences, such as traditions followed, treatment received from teachers, and clubs joined. Work experiences may only be a fraction of the total set of experiences it captures (Pelled, Ledford, and Mohrman, 1998). Sessa and Jackson (1995: 137) have observed that race, gender, and age "form the context of more general social relationships" and, compared with tenure and department membership, are less directly associated with team objectives. Functional background and tenure, then, are apt to have a stronger impact on perceptions of work group tasks than are race, gender, and age:

Hypothesis 1 (H1): Functional background and tenure diversity will have stronger positive associations with intragroup task conflict than will diversity in age, gender, and race.

Link between Diversity and Emotional Conflict

While task conflict is largely shaped by the job-relatedness of diversity, emotional conflict is shaped by a more complex set of forces. One key factor is categorization, the subconscious tendency of individuals to sort each other into social categories, often on the basis of demographic attributes (Tajfel et al., 1971; Tajfel, 1972, 1982). Because there is an abundance of information about people and things in our environment, categorization is a useful way to simplify and "make our perceived world more predictable and controllable" (Zimbardo and Leippe, 1991: 236). Once categorization takes place, people strive for self-esteem by developing positive opinions of their own category and negative opinions of other categories (Turner, 1975; Tajfel, 1978). They perceive members of their own social category as superior and engage in stereotyping, distancing, and disparaging members of other categories (Tajfel, 1982). Members of other social categories, in turn, resent such stereotyping and disparaging treatment, and hostile interactions erupt in the group (Reardon, 1995). These hostile interactions constitute emotional conflict, clashes characterized by anger, resentment, and other negative feelings. As diversity within a work group increases, individuals generally will have more exchanges with those in different social categories. People in different social categories will be directly confronted with each other's negative stereotypes and self-serving biases, and emotional conflict may become more pronounced.

Although any kind of diversity may provoke categorization—and thus emotional conflict—in this manner, some kinds have a greater tendency to do so than others, depending on

the permeability of their defining demographic attributes. The permeability of an attribute is the degree to which that attribute can be altered, moving a person from one social category to another social category. Diversity based on relatively impermeable attributes is particularly likely to yield intercategory clashes. When attributes are not easily penetrated, it is difficult for employees to "stand in the shoes" of those in another social category. Consequently, employees feel especially polarized—and are therefore especially likely to stereotype—members of another category (Nelson, 1989; Kramer, 1991).

Race, gender, age, and tenure are not easily permeated. A person cannot change his or her race or gender and must wait a period of time for noteworthy increases in age and tenure. Also, a person can never regress to a younger age or a lesser amount of tenure in the same company. Functional background, in contrast, is more permeable. Employees often can transfer from one functional area to another if they simply want exposure to different areas or if the conditions in another area are better, and many companies "rotate employees in and out of both technical and business-oriented positions to help them round out their skills" (Ryan, 1991: 76). Hence, employees may find it easier to identify with those of a different functional background than with those of a different race, gender, age, or tenure. Thus, categorization theory leads to the following prediction:

Hypothesis 2a (H2a): Diversity in race, gender, age, and tenure will have stronger positive associations with intragroup emotional conflict than will diversity in functional background.

The previous hypothesis, derived from categorization theory, suggests that diversity tends to increase emotional conflict, but it is also possible that diversity diminishes emotional conflict. Festinger's (1954) social comparison theory suggests this competing prediction.[2] According to Festinger, humans have an innate tendency to evaluate themselves and their qualities—e.g., their opinions, abilities, and progress. Festinger contended that when objective information is unavailable, people prefer to compare themselves with similar others. This preference exists because comparison with a similar other is more meaningful and informative than comparison with someone who is very different. Festinger further asserted that, in Western culture, people feel a pressure to improve their abilities or other qualities continually, and, as a result, they strive to be slightly better than the targets of their comparisons. This process, according to Festinger, leads to competition among similar others.

In his original formulation of social comparison theory, Festinger used the term similarity to refer to similarity on the ability or other quality being evaluated. Over the years, however, researchers have found that people also compare themselves with those who are similar on other dimensions, such as attractiveness or gender (e.g., Miller, 1982; Major and Forcey, 1985). Based on social comparison theory, then, we can expect demographic similarity to be associated with rivalry or professional competition in work groups. Feelings of jealousy and hostility may accompany competitive, rivalrous interactions (Tjosvold, 1991). Hence, demographic

2
Personal communication from Blake Ashforth, 1997.

similarity (group homogeneity) may ultimately be linked to intragroup emotional conflict. As a work group becomes more homogeneous (i.e., as diversity decreases), individuals generally have more exchanges with similar others. Employees may be confronted with more social comparisons that precipitate rivalrous clashes, and emotional conflict in the group may become more pronounced.

While any kind of similarity may trigger social comparison—and thus emotional conflict—in the work context, some kinds of similarity are stronger triggers than others, depending on the career-relatedness of their defining demographic attributes. The career-relatedness of an attribute is the degree to which that attribute is considered in formal and informal assessments of career progress. In a work context, people are especially attuned to career achievements, rather than outside accomplishments. Hence, when employees engage in social comparison at work, they primarily look at those demographic attributes tied to career progress evaluations. Similarity with respect to highly career-related attributes is therefore particularly likely to yield the jealous rivalry that characterizes emotional conflict.

Age, tenure, and functional background are highly career-related. There are powerful age norms encouraging employees to expect that career progress comes over time, with age (Lawrence, 1988). Hence, workers may view people who are similar in age—people at the same stage in life—as yardsticks with which to measure their own career progress, and they may be concerned about falling behind (e.g., not being as successful or powerful as) those persons. For example, a group member who is 25 years old is apt to be more concerned about a 27-year-old "shining" in the group than about a 50-year-old shining. By the same token, there are implicit career timetables and expectations for know-how associated with tenure, so employees are inclined to look to others of the same tenure to see who has achieved greater recognition, acquired more expertise, or made more career progress in other ways. Additionally, since formal evaluations of individual employees typically compare employees in the same functional area or department (Kirkpatrick, 1986), group members may be especially inclined to focus on persons from the same functional area when making social comparisons.

Race and gender, in contrast, are less highlighted in informal and formal assessments of career progress. For example, a female employee will not necessarily expect the same degree of influence and recognition as another female employee simply because both are women; each may bring a different degree of experience to a group. Moreover, comparisons based on gender and race similarity are, in general, not an explicit part of formal evaluations. Thus, consideration of social comparison processes leads to the following expectation:

Hypothesis 2b (H2b): Diversity in age, tenure, and functional background will have stronger negative associations with intragroup emotional conflict than will diversity in race and gender.

Link between Task and Emotional Conflict

Researchers have previously suggested that task and emotional conflict may influence each other. As Ross (1989:

140) observed, it is possible for task-related disagreements "to generate emotionally harsh language, which can be taken personally. We then have both task and psychological conflicts occurring at the same time." Group members may feel strongly that their views on a particular issue are correct, and they may show impatience or intolerance when others express different views. Moreover, members whose ideas are disputed may feel that others in the group do not respect their judgment. Tjosvold (1991) observed that group members sometimes assume their competence is being challenged when their ideas are criticized. Conversely, emotional conflict may sometimes lead to task conflict. Individuals who feel frustrated or angry with other members of their group may have a propensity to dispute the ideas of those other members, for angry people are generally less compliant and agreeable than those who are cheerful (Milberg and Clark, 1988). Also, there may be a negative halo effect, such that when one feels irritated by or hostile toward another person, one is more inclined to find fault with that person's ideas. In their multiple case study of top management teams, Eisenhardt and Bourgeois (1988) concluded that executives who were engaged in political infighting tended to have distorted perceptions of each other's ideas. Even though such conflict is affective in origin, it may evolve into substantive debate:

Hypothesis 3 (H3): There will be a positive association between task conflict and emotional conflict in work groups.

Moderators of Diversity-Conflict Linkages

Jackson (1992: 155) observed that an important but as-yet-unanswered question is, "Does the nature of the task moderate the impact of group composition?" In addressing this query, a key task feature that warrants attention is task routineness, the extent to which a task has low information processing requirements, set procedures, and stability (Van de Ven, Delbecq, and Koenig, 1976; Gladstein, 1984; Jehn, 1995). This feature is especially relevant because it determines the richness of information required for a group's task, that is, whether a group needs to draw on different knowledge bases.

There are two possible effects that task routineness can have on the positive association between diversity and task conflict. One possibility is that task routineness diminishes the association between diversity and task conflict. When tasks are well-defined and straightforward, group members have little need to exchange opinions or challenge each other. Hence, in groups with routine tasks, even if members have diverse backgrounds, there is only minimal room for task conflict based on those backgrounds. In groups with nonroutine tasks, there is more room for task conflict, so group members with diverse backgrounds are more likely to exchange opposing opinions and preferences derived from their backgrounds. The tendency for diversity to trigger task conflict may therefore be heightened by task nonroutineness, or diminished by task routineness:

Hypothesis 4a (H4a): Task routineness will reduce the positive associations between diversity variables and task conflict in work groups.

An alternative possibility is that task routineness will have the opposite effect. Early theories of optimal arousal (Hebb, 1955; Fiske and Maddi, 1961; Berlyne, 1967) suggested that people have a preferred level of arousal, a preference for stimulation that is neither too low nor too high. Drawing on these theories, Zuckerman (1979) postulated that people engage in behaviors that decrease stimulus input when their optimal level of arousal is exceeded and increase stimulus input when they are underaroused. Empirical evidence, including results of sensory deprivation research (e.g., Vernon and McGill, 1960) and Zuckerman's own studies (1979, 1984), is consistent with this notion. An implication of Zuckerman's theory, referred to as sensation-seeking theory, is that people who are understimulated will seek experiences and interactions that offer them greater arousal. Since group members performing routine tasks may experience suboptimal levels of stimulation, they may seek opportunities to debate about their tasks to make their work more exciting. More specifically, they may elicit opposing task perspectives from people with different backgrounds, hoping to engage in cognitively stimulating discourse. In contrast, group members performing nonroutine tasks may be sufficiently aroused by the group's task and may be less motivated to draw out additional task conflict for the sake of excitement:

Hypothesis 4b (H4b): Task routineness will enhance the positive associations between diversity variables and task conflict in work groups.

Sensation-seeking theory is not likely to apply in the case of emotional conflict. When tasks are routine, group members are unlikely to seek emotional conflicts with people of different backgrounds, for researchers have found that people typically do not seek unpleasurable arousal when understimulated (Zuckerman, 1979; Gallagher, Diener, and Larsen, 1989). Still, task routineness may influence the relationship between diversity and emotional conflict. People performing complex tasks may be more anxious and, consequently, rely more heavily on cognitive mechanisms for simplifying information processing (Staw, Sandelands, and Dutton, 1981). In contrast, when tasks are routine, people have less need for such cognitive mechanisms (e.g., categorization). Thus, the tendency for diversity to trigger categorization and, ultimately, emotional conflict will be weaker when tasks are routine.

Research on displaced aggression also suggests that the association between diversity and emotional conflict will be weaker when tasks are routine. Studies have shown that frustrating work conditions lead to more interpersonal aggression among employees (Storms and Spector, 1987; Chen and Spector, 1992). Because routine tasks tend to be less frustrating than complex tasks, members of groups with routine tasks may have less frustration to vent and, consequently, less inclination to blame or "pick on" people of different backgrounds, compared with members of groups with complex tasks:

Hypothesis 5 (H5): Task routineness will reduce the positive associations between diversity variables and emotional conflict in work groups.

In addition to task routineness, a second moderator likely to operate on diversity-conflict relationships is group longevity, the length of time group members have spent working together (Katz, 1982). After a period of time, group members may become familiar with the different perspectives in a diverse group. If a group member has an idea, he or she may be able to anticipate other members' criticisms and, consequently, either frame the idea to make it more acceptable or avoid expressing the idea altogether. Alternatively, through informational social influence, group members may begin to share each other's perspectives, arriving at a common understanding of the group's tasks (Katz, 1982). In this manner, group longevity may diminish any tendency for diversity to trigger task conflict.

Group longevity may also weaken any positive associations between work group diversity and emotional conflict, for social categories based on demographic attributes may eventually become blurred. The boundaries of the "in" category and the "out" category may change, so that individuals who were once considered outsiders become insiders (Kramer, 1991). Over time, as members grow accustomed to being in the same work group, the perceived "in" category is apt to become the whole work group, while other work groups are perceived as "out" categories. Since people have a limited focus of attention (Kahneman, 1973; Fiske and Taylor, 1991), their enhanced focus on the group as a category is likely to diminish their focus on demographic categories. Group longevity may therefore moderate any tendency of diversity to yield emotional conflict in work groups:

Hypothesis 6 (H6): Group longevity will diminish the positive associations between diversity variables and conflict in work groups.

Task and Emotional Conflict as Mediators: Their Links to Performance

The task and emotional conflict triggered by a group's diversity may, in turn, affect the cognitive task performance of the group, although the mediating roles of the two types of conflict are apt to differ. The task conflict that diversity yields is likely to enhance group performance on cognitive tasks. Exposure to opposing points of view encourages group members to gather new data, delve into issues more deeply, and develop a more complete understanding of problems and alternative solutions (Tjosvold, 1986). Also, the constructive criticism associated with task conflict can facilitate vigilant problem solving, an approach that Janis (1989) recommended for making important decisions. If group members fail to criticize each other's ideas because they are too concerned about maintaining unanimity, they may overlook important details, succumbing to "groupthink" (Janis, 1982).

Previous case study and empirical findings support the notion that task conflict enhances cognitive task performance. Hoffman and Maier (1961) found in a lab study that groups with conflicting opinions produced better solutions to standardized sets of problems. Later, using a sample of student groups, Jehn (1994) showed that task

conflict was positively associated with group performance on a class project. In a multiple case study, Eisenhardt, Kahwajy, and Bourgeois (1997b) found that top management teams in high-performing firms had higher task conflict than teams in low-performing firms. Amason's (1996) large-scale study of top management teams was consistent with this pattern, revealing that task conflict was positively associated with decision quality. Thus,

Hypothesis 7 (H7): Task conflict will have a positive association with the cognitive task performance of work groups.

In contrast to task conflict, the emotional conflict yielded by diversity is likely to impair the cognitive task performance of work groups. First, since anxiety (an emotion that characterizes emotional conflict) often leads to cognitive interference (Sarason, 1984), group members may not take relevant information into consideration when solving problems. Second, the hostility that characterizes affective conflict may make group members reluctant to share or listen to each other's potentially useful ideas or information. Third, when there is emotional conflict, group members are likely to consume time and energy making—or defending themselves against—personal attacks; as a result, they may have little remaining time and energy to devote to critical task-related matters (Evan, 1965; Jehn, 1994).

The results of prior research are consistent with the notion that emotional conflict impairs cognitive task performance. Evan's (1965) study of research and development teams suggested that interpersonal attacks diminished team productivity. Also, Jehn's (1994) above-mentioned study of student groups showed that intragroup emotional conflict was negatively associated with group performance on a class project. Similarly, Amason's (1996) study of top management teams revealed that emotional conflict, which he called affective conflict, was negatively associated with decision quality. Eisenhardt, Kahwajy, and Bourgeois (1997b), too, found that emotional conflict impaired team process and firm performance. Thus,

Hypothesis 8 (H8): Emotional conflict will have a negative association with the cognitive task performance of work groups.

The above eight hypotheses constitute an intervening process theory of work group diversity, conflict, and performance. Below we describe the field study conducted to test our hypotheses.

METHODS

Sample

Participants in this study included the members of 45 teams from the electronics divisions of three major corporations. The teams were involved in monitoring and modifying work processes with the objective of improving those processes, and often they were also involved in the design of new products. All teams were assembled to complete lengthy but time-limited projects, and many of the teams were cross-functional, including research and development (R&D) and manufacturing representatives within their functional mixtures. They also were engaged in cognitive tasks that,

according to team members' assessments, ranged from moderate to high complexity.

We obtained team performance ratings from 41 out of 45 team managers and received completed questionnaires from members of all 45 teams. Although the number of teams is modest by some standards, it compares well with other field studies and, as a field study, yields valuable insights not attainable in the laboratory. A total of 443 team-member questionnaires were distributed, and 317 were returned. On average, 73 percent of the members of a team returned completed questionnaires. The teams in our sample had an average size of approximately 10 members (s.d. = 3.2). The average team had a mean age of 38.5 years (s.d. = 5.5) and a mean tenure of 10.6 years (s.d. = 4.9). Also, the average percentage of nonwhites on a team was 19 percent (s.d. = 21 percent); the average percentage of women on a team was 25 percent (s.d. = 18 percent); and the average R&D representation on a team was 44 percent (s.d. = 33 percent).

Measures

Diversity. This study included two types of work group diversity indices, one for numeric demographic data and another for categorical demographic data. The team-member questionnaire was the source of demographic data used to compute these indices. Following an approach recommended by Allison (1978) for numeric variables, we used the coefficient of variation (standard deviation divided by the mean) to measure age diversity and tenure diversity. Thus, to assess age diversity within teams, we divided each team's standard deviation of age by the team's mean age. Similarly, we assessed company tenure diversity within teams by dividing each team's standard deviation of tenure by the team's mean tenure.

To measure team diversity with respect to categorical variables (functional background, gender, and race), we used an index recommended by Teachman (1980):

$$H = -\sum_{I=1}^{I} P_i \, (\ln P_i).$$

The index takes into account how work group members are distributed among the possible categories of a variable. The total number of categories of a variable equals I, and P_i is the fraction of team members falling into category I. For example, the gender variable has two possible categories (I = 2): 1 corresponds to a woman and 2 to a man. If a given team of ten members has three women and seven men, then P_1 equals .3, P_2 equals .7, and H equals .61. If a team of ten members has one woman and nine men, then P_1 equals .1, P_2 equals .9, and H equals .32. As Ancona and Caldwell (1992: 328) noted, "The only exception occurs when [a category] is not represented." In such a case, one cannot set P_i equal to zero, for the natural logarithm of zero does not exist; thus, one would only use the P_i values for the other categories to compute H.

Conflict. The task conflict scale (α = .78) comprised four questionnaire items based on Jehn's (1994) measure of task conflict, and the emotional conflict scale (α = .83) comprised

four questionnaire items adapted from Jehn's (1994) emotional conflict scale. Each item measured conflict on a 5-point Likert scale. Group-level indices of task and emotional conflict were formed by averaging individual-level indices of task and emotional conflict. Aggregation to the group level was justified by an eta square of .33 for task conflict ($F = 3.03$, $p < .001$) and an eta square of .32 for emotional conflict ($F = 2.92$, $p < .001$). In general, an eta square greater than .20 indicates that any two people within the same group are more similar in their responses than two people who are members of different groups (Georgopoulos, 1986; Florin et al., 1990).

Moderators. The task routineness measure ($\alpha = .62$) comprised three items drawn from Gladstein's (1984) task complexity measure, reverse-scored on a 5-point Likert scale. Aggregation to the group level was justified by an eta square of .31 ($F = 2.72$, $p < .001$). To ensure the discriminant validity of the task conflict, emotional conflict, and task routineness items, we conducted a confirmatory factor analysis (CFA). Table 1 shows the results of this analysis. The CFA, which used the maximum likelihood method, produced a chi square of 46.90 with 41 degrees of freedom (i.e., a chi square of 1.14 per degree of freedom). The goodness-of-fit index is .98, and the root-mean-square residual is .035. These figures indicate that there is a good fit between the data and the theoretical factor structure (Wheaton, Alwing, and Summers, 1977).

Our group longevity measure, following previous studies (Katz, 1982; Smith et al., 1994), was the average length of time the members of a team had belonged to that team. A team with a higher average has a longer history of working together.

Performance. Using 5-point Likert scales, each team's manager rated the team on two dimensions from Ancona and Caldwell's (1992) measure of manager-rated team

Table 1

Results of Confirmatory Factor Analysis of Task Routineness and Intragroup Conflict Items

Item	Factor 1 Task conflict	Factor 2 Relationship conflict	Factor 3 Task routineness
1. To what extent are there differences of opinion in your team?	.68		
2. How often do the members of your team disagree about how things should be done?	.83		
3. How often do the members of your team disagree about which procedure should be used to do your work?	.76		
4. To what extent are the arguments in your team task-related?	.57		
5. How much are personality clashes evident in your team?		.80	
6. How much tension is there among the members of your team?		.83	
7. How often do people get angry while working in your team?		.62	
8. How much jealousy or rivalry is there among the members of your team?		.70	
9. The technology, required skills, and information needed by the team are constantly changing. [reverse-scored]			.67
10. During a normal work week, exceptions frequently arise that require substantially different methods or procedures for the team. [reverse-scored]			.79
11. Frequent interaction between team members is needed to do our work effectively. [reverse-scored]			.47

performance: efficiency of team operations and number of innovations or new ideas introduced by the team. The group performance measure comprised these items (α = .61). Principal components analysis revealed that the items constituted a single factor accounting for 74 percent of the variance in responses. Each item in the scale had a factor loading of .86.

When forming each multi-item measure in this study (i.e., task conflict, emotional conflict, task routineness, and performance), we combined items by using factor scores as weights and computed a weighted sum of scale items, as described by Pedhazur and Schmelkin (1991: 125). Using factor scores in this manner creates a more accurate measure than simply computing a mean, which assigns equal weights to items. This procedure was feasible because respondents answered all of the items in a scale.

Controls. Group size was a control variable in our study because the literature on groups has noted that size is a key variable influencing group dynamics and performance (Brewer and Kramer, 1986) and because larger teams have more potential for heterogeneity (Bantel and Jackson, 1989; Jackson et al., 1991). We also controlled for site differences in our regressions. In particular, Site B had a distinct culture. It was more conservative and bureaucratic than the other two sites, offered less training in team skills, and was heavily connected to the declining defense industry. We controlled for group longevity and task routineness (the moderators described above) because previous research has found that the average tenure of group members and the nature of group tasks often influence group interactions and performance (e.g., Katz, 1982; Weingart, 1992). Also, when testing for moderating effects, it is necessary to control for the main effect of the moderating variable (Baron and Kenny, 1986).

Data Analysis

We tested the hypotheses using seemingly unrelated regression (SURE) and ordinary least squares regression. When evaluating the significance of most predicted effects (all except those predicted in competing hypotheses), we used one-tailed tests, which are suitable for directional hypotheses (Erickson and Nosanchuk, 1977; Wonnacott and Wonnacott, 1984). LISREL was not an alternative for this study because it was inappropriate for our sample size (Breckler, 1990).

Hypotheses 1 through 6 (with the exception of hypothesis 3) were tested with SURE analyses. SURE is a statistical technique that solves a set of regression equations simultaneously and allows for error covariances among the equations (Zellner, 1962; Parker and Dolich, 1986; Ghosh, 1991). It is appropriate to use the technique in this study because the predictors in the equations with emotional conflict as a dependent variable have considerable overlap with the predictors in the equations with task conflict as a dependent variable. Also, we expected that task and emotional conflict would be significantly correlated. Each SURE analysis involved two equations, one that had task

conflict as the dependent variable and another that had emotional conflict as the dependent variable.

Our first SURE analysis only included control variables as predictors. In the second SURE analysis, we included both diversity and control variables as predictors. In the third SURE analysis, we added interaction terms, having mean centered the interaction term variables to reduce potential multicollinearity effects. We used the likelihood ratio test (Kennedy, 1979; Ghosh, 1991) to determine the goodness of fit of the models. For the likelihood ratio test, the criterion is as follows:

$$-2[\ln(\text{likelihood function}_{\text{constrained model}}) - \ln(\text{likelihood function}_{\text{unconstrained model}})].$$

The criterion has a χ^2 distribution with degrees of freedom equal to the number of constraints.

To test H3, which suggested that task and emotional conflict would be positively associated, we examined the zero-order correlation between the two conflict measures. We also performed an OLS regression (not shown but available from the authors), using the approach that Smith et al. (1994) used in their study of demography, group process, and performance when assessing the effect of one intervening process variable (informality of communication) on another intervening process variable (social integration). They treated social integration as an "intervening dependent variable," while "all team demography and control variables, plus informality of communication were treated as independent variables" (1994: 429). Here, we assessed the effect of emotional conflict on task conflict by regressing task conflict on the diversity variables, controls, and emotional conflict. Similarly, we assessed the effect of task conflict on emotional conflict by regressing emotional conflict on the diversity variables, controls, and task conflict.

When testing H7 and H8, which posited that task and emotional conflict would act as mediators between diversity and group performance, we followed a standard procedure used to test for mediating effects, "first, regressing the mediator on the independent variable; second, regressing the dependent variable on the independent variable; and third, regressing the dependent variable on both the independent variable and on the mediator . . . " (Baron and Kenny, 1986: 1177). The independent variables were the diversity variables in this study, and the proposed mediating variables were task and emotional conflict. The dependent variable was group performance. After running the regressions, we examined the results to see if a mediating effect was present. Pure mediation calls for the following conditions (Baron and Kenny, 1986): First, the independent (diversity) variables should affect the mediator (task or emotional conflict) in the first equation. Second, the independent variables should affect the dependent variable (performance) in the second equation. Third, the mediator (task or emotional conflict) should affect the dependent variable in the third equation, and the effect of the independent variables on the dependent variable should be less in the third equation than in the second equation.

RESULTS

Table 2 shows the means, standard deviations, and correlations among all predictor, outcome, and control variables.

We performed several checks on the correlational properties of the data before testing our hypotheses. First, we reviewed the correlations among the independent variables shown in table 2. The median correlation magnitude (absolute value) was .19, and the correlation with the greatest magnitude was .48. As noted by Tsui et al. (1995: 1531), "There is no definitive criterion for the level of correlation that constitutes a serious multicollinearity problem. The general rule of thumb is that it should not exceed .75." Similarly, Kennedy (1979) indicated that correlations of .8 or higher are problematic. As a second check, we examined the variance inflation factor (VIF) of each independent variable. The largest VIF in our regressions was less than 3, a sign that multicollinearity was not a problem (Guo, Chumlea, and Cockram, 1996).

As described earlier, H1, 2a, and 2b were tested with SURE analysis involving two equations. Using the likelihood-ratio test, we determined that the goodness of fit of this two-equation model was $\chi^2 = 24.81$ (d.f. = 10; $p < .01$). Table 3 presents the SURE equations with task conflict as the dependent variable. Model 2 tested H1, that functional background and tenure diversity would have positive associations with task conflict in work groups. This hypothesis was supported for functional background diversity, which had a significant positive relationship with task conflict (beta = 1.08, $p < .05$), while, consistent with the same hypothesis, diversity in race, gender, and age had nonsignificant associations with task conflict. Unexpectedly, the relationship between tenure diversity and task conflict was also nonsignificant. This suggests that functional background diversity is the key demographic driver of task conflict.

Table 4 presents the SURE equations with emotional conflict as the dependent variable. Model 2 tested H2a and 2b,

Table 2

Means, Standard Deviations, and Intercorrelations among Study Variables*

	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11
1. Group size	9.84	3.17	–										
2. Task routineness	5.31	.94	−.10	–									
3. Group longevity	.89	.71	.43	−.03	–								
4. Site B	–	–	−.04	.22	.08	–							
5. Emotional conflict	4.58	1.04	.35	−.43	.03	.09	–						
6. Task conflict	6.93	1.32	.28	−.23	.01	.32	.48	–					
7. Race diversity	.45	.42	−.18	.09	−.17	−.46	−.04	−.24	–				
8. Gender diversity	.48	.20	−.20	.26	−.28	.31	−.10	.03	.25	–			
9. Age diversity	.24	.08	−.14	.30	−.14	−.06	−.45	−.27	.19	.18	–		
10. Company tenure diversity	.72	.20	.18	−.06	.37	−.25	.27	−.01	−.08	−.35	−.18	–	
11. Functional background diversity	.64	.37	−.07	.22	−.24	−.07	−.21	.14	.26	.19	.33	−.16	–
12. Group performance	6.26	1.30	.16	.00	.40	−.30	−.07	.05	−.02	−.19	.09	.36	−.13

* Because the emotional conflict, task conflict, task routineness, and group performance measures were each formed by using factor scores to compute weighted sums of scale items, their means may exceed 5, even though their items had only 5 response anchors. All correlations above .19 are significant at $p < .10$, one-tailed tests. All correlations above .25 are significant at $p < .05$, one-tailed tests.

Table 3

Seemingly Unrelated Regression (SURE) Results: Equations with Task Conflict as Dependent Variable ($N = 45$)*

Independent variables	Model 1	2	3	4
Controls				
Group size	.14••	.12••	.11••	.08
Site B	1.25•••	1.34••	1.41•••	1.28••
Task routineness	−.39••	−.40••	−.42••	−.52•••
Group longevity	−.33	−.30	−.18	−.19
Predictors				
Race diversity (RD)		−.02	.15	.11
Gender diversity (GD)		−.04	.17	.50
Age diversity (AD)		−3.64	−3.46	−4.17•
Company tenure diversity (TD)		.64	.58	.61
Functional background diversity (FD)		.108••	.98••	1.01••
Interactions				
FD × Task routineness			.74•	
FD × Group longevity				−.153••
R-squared	.29	.39	.44	.42

• $p < .10$; •• $p < .05$; ••• $p < .01$.
* Entries are SURE coefficients. One-tailed tests were used for effects predicted in directional hypotheses.

predictions about diversity and emotional conflict based on categorization theory and social comparison theory, respectively. Consistent with H2a, both tenure diversity (beta = 1.76, $p < .01$) and race diversity (beta = .81, $p < .05$) had significant positive associations with emotional conflict, while functional background diversity did not. Consistent with H2b, age diversity had a significant negative association with emotional conflict (beta = −3.43, $p < .05$). Surprisingly, gender diversity was not significantly related to emotional conflict in either direction.

As a sensitivity analysis, we replaced each of our diversity measures for categorical variables (race, functional background, and gender) with an alternative measure. Instead of Teachman's index, we used Blau's (1977) heterogeneity index:

$$H = 1 - \sum_{i=1}^{I} (P_i)^2,$$

where P_i is the proportion of group members in category i, and I is the number of possible categories. For example, if a given team of ten members has three women and seven men, then P_1 equals .3, P_2 equals .7, and H equals .42. With this alternative measure, we still obtained the pattern of results described above.

Since the effect of gender diversity was nonsignificant, we conducted several additional sensitivity analyses to test the robustness of this result. We replaced our gender diversity measure with the simpler measure that South et al. (1982) used in their study of intergender relations: the proportion of women in a work group. This measure, like Teachman's, revealed no significant effect of gender diversity. Next, we

Table 4

Seemingly Unrelated Regression (SURE) Results: Equations with Emotional Conflict as Dependent Variable (*N* = 45)*

	Model			
Independent variables	1	2	3	4
Controls				
Group size	.12••	.13•••	.12•••	.13•••
Site B	.49•	1.04•••	1.20•••	1.06•••
Task routineness	−.47•••	−.41•••	−.48•••	−.43•••
Group longevity	−.22	−.49••	−.49•••	−.52•••
Predictors				
Race diversity (RD)		.81••	.89•••	.76••
Gender diversity (GD)		−.27	−.19	−.36
Age diversity (AD)		−3.43••	−2.63•	−3.52••
Company tenure diversity (TD)		1.76•••	1.74•••	1.64•••
Functional background diversity (FD)		−.23	−.38	−.12
Interactions				
RD × Task routineness			−1.12••	
TD × Task routineness			−.59••	
RD × Group longevity				−.70••
TD × Group longevity				−1.44•
R-squared	.32	.54	.60	.59

• $p < .10$; •• $p < .05$; ••• $p < .01$.
* Entries are SURE coefficients. One-tailed tests were used for effects predicted in directional hypotheses.

used a dummy variable to indicate whether or not a group was skewed, with women constituting less than 20 percent of its membership or men constituting less than 20 percent of its membership (Kanter, 1977). We looked at the effects of skewness in favor of men (i.e., larger proportion of men), skewness in favor of women, and skewness in general, and still, we found no significant effects of gender composition on conflict in work groups.

Our initial test of H3, which predicted a positive association between task conflict and emotional conflict, was the zero-order correlation between the two variables. The two types of conflict were positively correlated ($r = .48$, $p < .01$), consistent with H3. Several OLS regressions (not shown but available from the authors) corroborated this finding. A regression of task conflict on diversity variables, controls, and emotional conflict showed that emotional conflict was a significant predictor of task conflict (beta = .35, $p < .05$). A regression of emotional conflict on diversity variables, controls, and task conflict showed that task conflict was a significant predictor of emotional conflict (beta = .26, $p < .05$).

Like H1 and 2, H4a, 4b, and 5 were tested with a SURE analysis involving two equations. Using the likelihood ratio test, we determined that the goodness of fit of this two-equation model was $\chi^2 = 10.36$ (d.f. = 3; $p < .05$). Hypothesis 4a proposed that task routineness would reduce the positive relationships between diversity and task conflict, while competing hypothesis 4b suggested that task routineness would enhance the positive relationships between diversity and task conflict. In table 3 (above), model 3 presents the equation that tested these predictions.

Consistent with H4b, the interaction of task routineness and functional background diversity had a significant positive association with task conflict (beta = .74, $p < .10$), suggesting that functional background differences were more likely to trigger task conflict when tasks were routine than when tasks were nonroutine. A partial derivative analysis revealed that the effect of functional background diversity on task conflict was monotonic over the range of task routineness observed in our sample.[3] A graphical display showed that the effect was stronger for higher levels of task routineness.

H5 proposed that task routineness would diminish the positive relationships between diversity and emotional conflict. In table 4, model 3 presents the equation that tested this hypothesis. Consistent with H5, the interaction of task routineness and race diversity had a significant negative association with emotional conflict (beta = −1.12, $p < .05$), and the interaction of task routineness and tenure diversity had a significant negative association with emotional conflict (beta = −.59, $p < .05$). Partial derivative analyses revealed that the effects of race diversity and tenure diversity were monotonic over the range of task routineness observed in our sample. Graphical displays showed that these effects were weaker for higher levels of task routineness.

H6 proposed that group longevity would diminish the positive relationships between diversity and conflict. This hypothesis was tested with a SURE analysis involving two equations, one for task conflict and one for emotional conflict. The goodness of fit of this two-equation model was $\chi^2 = 9.486$ (d.f. = 3; $p < .05$). As revealed in model 4 of tables 3 and 4, all moderating effects of longevity supported H6. The interaction of group longevity and functional background diversity had a significant negative association with task conflict (beta = −1.53, $p < .05$). The interaction of group longevity and race diversity had a significant negative association with emotional conflict (beta = −.70, $p < .05$), as did the interaction of group longevity and tenure diversity (beta = −1.44, $p < .10$). Partial derivative analyses revealed that the effects of the diversity variables were nonmonotonic over the range of longevity in our sample. Graphical displays showed that longevity had to reach a certain threshold to diminish the diversity-conflict relationships. For functional background diversity, this threshold was .66 years. For race diversity, the threshold was 1.09 years, and for tenure diversity, the threshold was 1.14 years.

The above-mentioned regressions of conflict against diversity variables constituted the first portion of our tests of mediating effects. Table 5 displays the remaining portions required by the Baron and Kenny (1986) method. Model 2 of table 5 shows the regression of our dependent variable (performance) against the independent variables (diversity measures) and controls. Model 3 shows the regression of our dependent variable against one process variable (task conflict), the independent variables, and controls. Model 4 shows the regression of our dependent variable against another process variable (emotional conflict), the independent variables, and controls. Finally, model 5 shows the regression of our dependent variable against both

3

To interpret each significant interaction further, we took an additional analytical step not shown in this paper (but available from the authors). Specifically, we examined the functional form of the interactions. This procedure, which Schoonhoven (1981) described in detail, is appropriate for interactions involving two continuous variables and avoids the information loss associated with median split procedures. First, we took a partial derivative to determine mathematically whether the moderated relationship (i.e., the relationship between the diversity variable and conflict, moderated by either task routineness or group longevity) was monotonic or nonmonotonic. We then plotted the partial derivative over the range of the mean-centered moderating variable (either task routineness or group longevity); this plot illustrated how the relationship between diversity and conflict changed over the range of the moderator's values (i.e., the values of task routineness or group longevity).

process variables (task and emotional conflict), the independent variables, and controls. The results do not strictly satisfy all requirements for a mediating effect, for we did not find a significant relationship between diversity and performance in model 2, although we did find evidence (model 3) that task conflict is positively associated with cognitive task performance (beta = .30, $p < .05$), as H7 predicted. This positive relationship remained when diversity variable predictors were removed from the equation.

Thus far, an implicit assumption in our hypotheses and analyses has been that the effects of diversity variables on conflict are independent of one another. It is conceivable, however, that this is not the case. Interactions among different types of diversity may also shape work unit dynamics (Alexander et al., 1995). Therefore, we performed additional, exploratory analyses to examine how the joint effects of diversity variables influenced conflict in these data. Given the exploratory nature of this analysis, we used two-tailed tests. Since there were five diversity variables, the number of possible combinations of two diversity variables was 10. We ran 10 SURE analyses to assess the effects of these 10 diversity variable interaction pairs on each of two conflict variables. Out of 20 possible interaction effects, four were significant. Gender diversity and age diversity had a positive interaction effect on emotional conflict (beta = 17.40, $p < .05$), as did gender diversity and functional background diversity (beta = 4.00, $p < .05$). Age diversity and tenure diversity had a negative interaction effect on emotional conflict (beta = −15.39, $p < .05$), and the interaction between race diversity and functional background diversity had a negative effect on task conflict (beta = −2.50, $p < .05$).

Table 5

Regression of Group Performance on Conflict, Diversity, and Control Variables ($N = 41$)*

	Model				
Independent variables	1	2	3	4	5
Controls					
Group size	.02	.03	−.07	.00	−.05
Task routineness	.16	.17	.22	.19	.19
Group longevity	.44••	.35••	.40••	.37•	.37•
Site B	−.39••	−.40••	−.51•••	−.42•	−.48••
Diversity variables					
Race diversity		−.12	−.11	−.14	−.09
Age diversity		.13	.20	.14	.18
Company tenure diversity		.14	.11	.12	.14
Functional background diversity		−.10	−.19	−.10	−.21
Conflict variables					
Task conflict			.30••		.32••
Emotional conflict				.06	−.09
R-squared	.29	.34	.40	.34	.40
F	3.70••	2.06•	2.27••	1.79	2.00•

• $p < .10$; •• $p < .05$; ••• $p < .01$.

* Entries are standardized regression coefficients. One-tailed tests were used for effects predicted in directional hypotheses.

DISCUSSION

Researchers of work group demography and top management team composition have often relied on the argument that diversity increases conflict, which, in turn, influences group performance. This study assesses the validity of that argument and reveals that the black box between diversity and performance is complex. Our results suggest that different types of diversity have distinctive effects. The diversity variables that drive task conflict differ from those that drive emotional conflict, and task conflict, in turn, tends to have more favorable performance consequences than emotional conflict. Additionally, our findings indicate that diversity can both increase and decrease conflict. The results also suggest that the combination of diversity types present and several contextual moderators influence the strength of the relationship between a particular diversity variable and conflict. In short, our findings point to multiple interrelated factors that must be considered when determining how work group composition will shape conflict and, ultimately, performance.

Given that a preponderance of group research has been confined to laboratory settings, the current study is noteworthy in its use of field data, including team members' questionnaires as well as managers' ratings of team performance, but the data are not without limitations. In particular, they are cross-sectional, a feature that renders causal interpretations difficult. Also, the sample size of 45 teams, while substantial for a field investigation, limits the power of statistical tests. In addition, several of our predictor variables (e.g., age diversity) had only modest variation across groups, so their effects may be underestimated. Nonetheless, this study begins to develop a multifaceted model of demographics, process variables, and outcomes, and the use of workplace data helps ensure that our findings have external validity. The findings themselves reveal some telling relationships among diversity, conflict, and performance.

Diversity and Conflict

The findings suggest that task conflict is a relatively straightforward phenomenon driven by functional background differences, a highly job-related type of diversity. Apparently, because functional background is so related to work, people are particularly likely to draw on belief structures based on functional background when addressing workplace issues; hence, functional background differences are the key source of task conflict in work groups. This result substantiates managers' use of cross-functional teams to create difference of opinion.

While task conflict is a relatively simple phenomenon driven by functional background differences, our findings suggest that emotional conflict is more complicated. On one hand, emotional conflict is increased by dissimilarity in race and tenure. It appears that, because race and tenure attributes are relatively impermeable, people find it difficult to identify with (and easy to stereotype) those of a different race or tenure. Race and tenure differences therefore tend to

encourage heated interactions in work groups. Given this tendency, managers may want to pay particular attention to group process in multi-race and mixed-tenure settings. On the other hand, emotional conflict is increased by similarity in age. Any tendency for age differences to trigger emotional conflict appears to be overshadowed by the tendency for age similarity to trigger social comparison and, ultimately, emotional conflict. Age is a career-related attribute, so employees tend to measure their own career progress by looking at that of coworkers in their age cohort (Lawrence, 1988). When age similarity in a group increases, these career progress comparisons, which prompt jealous rivalry, often increase. As Hambrick (1994: 202) noted, "if group members are extremely similar, they face the prospect of head-on rivalries that could drive them apart. . . . This might particularly occur if several group members were the same age and vying to be the next group leader. . . . " Overall, these findings suggest that diverse groups face countervailing forces, such that some forms of diversity increase conflict and other forms do the reverse. Hence, managers must be prepared to meet challenges presented by heterogeneity as well as homogeneity in their work groups.

The lack of a gender diversity effect in this study is intriguing, given that other studies have found important effects of gender heterogeneity on work group outcomes, including reduced performance on cognitive tasks (Kent and McGrath, 1969; Murnighan and Conlon, 1991), reduced cross-gender support (South et al., 1982), and increased within-gender support (Ely, 1994). One possible interpretation is that gender composition, by itself, is unrelated to conflict. Alternatively, gender diversity may trigger categorization and social comparison processes that cancel each other's effects. Yet another possibility is that the absence of gender diversity effects on conflict stems from the bicategorical nature of gender. If a demographic attribute has only two possible categories, then an increase in group diversity means that the distribution in the group becomes more balanced. When this occurs, members of the category previously holding the numerical majority have less opportunity to interact with similar others; consequently, they may face more conflicts. Members of the category previously in the numerical minority, however, have a greater opportunity to interact with similar others; consequently, they may face fewer conflicts. Hence, the net level of the conflict in the work group may change very little. Future research could profitably examine gender dynamics further.

Our exploratory analysis of diversity variable interactions yielded four significant findings. Some of these lend credence to recent arguments by Alexander et al. (1995). For example, we found that age diversity and tenure diversity, which were negatively correlated, had a negative interaction effect on emotional conflict. This finding is consistent with Alexander et al.'s (1995) assertion that when two types of diversity are negatively correlated, categorization may become more difficult, and intercategory tensions may be less likely. Conversely, the positive interaction effect of gender diversity and age diversity supports Alexander et al.'s

claim that categorization is easier—and intercategory tensions more likely—when two types of diversity are positively correlated. The positive interaction effect of gender diversity and functional background diversity also supports this claim.

Another exploratory finding, the negative interaction effect of race diversity and functional background diversity on task conflict, may reflect subtle racism, a form of prejudice that is more indirect and difficult to detect than old-fashioned direct racial prejudice (Dovidio and Gaertner, 1991). Due to societal norms discouraging the blatant expression of racial prejudice, employees may have reverted to a less obvious form of discrimination. That is, they may have ignored the opinions of persons from different functional areas when those persons were of a different race as well. Any interpretations of the observed diversity variable interactions must be treated with caution, however, given the exploratory nature of these analyses. Nonetheless, they do represent an avenue for future research.

Moderator Effects

Our work also extends traditional main-effects research on diversity by examining conditional effects. We found that task routineness and diversity interacted to influence conflict, although the direction of this interaction effect was not the same for both types of conflict. While task routineness reduced the positive associations between diversity and emotional conflict, it enhanced the positive associations between diversity and task conflict. The effect on emotional conflict supports our reasoning that, because routine tasks create less frustration than complex tasks, people in groups performing routine tasks are less likely to displace frustration onto dissimilar others. The effect on task conflict supports the notion that group members performing routine tasks seek task debates with dissimilar others to make their work more interesting. Consistent with previous sensation-seeking research, our study suggests that when a task is understimulating, people seek arousing experiences (e.g., task conflict based on background differences) but not unpleasant arousing experiences (e.g., emotional conflict based on background differences).

The moderating effects of group longevity were consistent with our hypotheses as well. The positive associations between diversity variables and both types of conflict were weaker in groups with more longevity. Apparently, after working together for a period of time, group members of different backgrounds either develop a shared understanding of tasks or learn to anticipate and deflect opposition to their ideas. Additionally, during that period, the boundaries of social categories may become blurred, so that individuals who were once considered out-group members become in-group members.

Conflict and Performance

Another significant finding in our sample of work groups was that task conflict had a positive association with cognitive task performance. Such conflict evidently fosters a deeper understanding of task issues and an exchange of information

that facilitates problem solving, decision making, and the generation of ideas. Although we found that functional background diversity was related to task conflict and task conflict was related to performance (both with and without controlling for diversity variables), a third sign was needed to verify a mediating effect: a positive relationship between diversity and performance, without task conflict included in the equation. We did not see that third sign, perhaps because we were unable to use a full structural equation model. A LISREL analysis with a larger sample may ultimately confirm mediation.

Surprisingly, we found no evidence that emotional conflict impaired performance. In a recent study, Jehn (1995) found no relationship between emotional conflict and performance, despite having found one before (Jehn, 1994). She explained that "while relationship troubles cause great dissatisfaction, the conflicts may not influence work as much as expected, because the members involved in the conflicts choose to avoid working with those with whom they experience [emotional] conflict" (Jehn, 1995: 276). In the present study, subjects may also have found ways to cope with those with whom they had emotional conflicts.

Although there are clear theoretical and empirical distinctions between task and emotional conflict, our results show that each type of conflict tends to accompany the other. Task conflicts may be taken personally by group members and generate emotional conflict, or emotional conflict may prompt group members to criticize each other's ideas, thereby fostering task conflict.

Taken together, our findings suggest that diversity variables can influence conflict and yet, with the exception of functional background diversity, do not necessarily have much bearing on work group performance. That is, while race, tenure, and age diversity influenced emotional conflict, they lacked substantial ties to performance. Groups (at least these groups) were apparently able to manage their negative effects. At the same time, except in the case of functional background diversity, groups did not achieve sizable gains from background differences.

Directions for Future Research

This investigation opens a number of avenues for related research. One avenue is to continue exploring the role of process variables in the relationship between work group diversity and performance. Although we have measured task and emotional conflict in this study, it would also be helpful to measure some of the other process variables, such as social comparison, categorization, and sensation seeking, that were unmeasured components of our theory. Because these processes provide explanations for competing predictions, it would be useful to investigate them further. Additionally, a next logical step for researchers is to conduct a study that assesses the linkages among conflict and the process variables that other demography studies have examined—e.g., external communication (Ancona and Caldwell, 1992), social integration (O'Reilly, Caldwell, and Barnett, 1989), and informal communication (Smith et al., 1994). Another opportunity is to explore further the

dynamics of gender diversity in work groups. The effects of such diversity may be difficult to capture in a deductive empirical study such as this one. Case research may be more revealing, allowing a more microscopic view of member exchanges in groups with moderate, low, and high levels of gender diversity. Finally, it would be useful to assess the robustness of the diversity variable interaction effects found here. Further study would be useful in probing the generalizability of those findings. Also, investigations with different dependent variables may determine whether diversity variable interactions influence processes other than conflict. Subtle racism, which one of our diversity variable interactions may reflect, is a particularly provocative topic for future research.

Conclusion

Demography scholars (e.g., Wagner, Pfeffer, and O'Reilly, 1984; Ancona and Caldwell, 1992) have frequently implied that conflict plays a central role in the relationship between diversity and work group outcomes. The present study has built on this notion, examining a complex set of linkages among work group diversity, conflict, and performance. The findings suggest that different types of diversity have distinct relationships with task and emotional conflict because of the specific properties of each type of diversity, because of interactions among the diversity types that are present, and because of the group's longevity and task routineness. Further, the findings suggest that task conflict has more favorable performance consequences than does emotional conflict. Overall, these results offer researchers a clearer view of the black box between work group diversity and performance.

This study also sheds light on patterns that practitioners can expect in diverse work groups. In particular, managers and members of cross-functional teams can take comfort in knowing that task conflict is likely in those teams and that such conflict may enhance performance. There is also a basis for expecting age variation to diminish emotional conflict. At the same time, race and tenure diversity may increase emotional conflict, especially in new groups with nonroutine tasks. Anticipating such a possibility may be critical if organizations hope to manage employees' background differences successfully.

REFERENCES

Alexander, Jeffrey, Beverly Nuchols, Joan Bloom, and Shoou-Yih Lee
1995 "Organizational demography and turnover: An examination of multiform and nonlinear heterogeneity." Human Relations, 48: 1455–1480.

Allison, Paul D.
1978 "Measures of inequality." American Sociological Review, 43: 865–880.

Amason, Allen C.
1996 "Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams." Academy of Management Journal, 39: 123–148.

Ancona, Deborah G., and David F. Caldwell
1992 "Demography and design: Predictors of new product team performance." Organization Science, 3: 321–341.

Bantel, Karen A., and Susan E. Jackson
1989 "Top management and innovations in banking: Does the composition of the top team make a difference?" Strategic Management Journal, 10: 107–124.

Baron, Reuben M., and David A. Kenny
1986 "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." Journal of Personality and Social Psychology, 51: 1173–1182.

Berlyne, Daniel E.
1967 "Arousal and reinforcement." In D. Levine (ed.), Nebraska Symposium on Motivation, 1967: 1–110. Lincoln, NE: University of Nebraska Press.

Blau, Peter M.
1977 Inequality and Heterogeneity. New York: Free Press.

Breckler, Steven J.
1990 "Applications of covariance structure modeling in psychology: Cause for concern?" Psychological Bulletin, 107: 260–273.

Brewer, Marilynn B., and Roderick M. Kramer
1986 "Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing." Journal of Personality and Social Psychology, 50: 543–549.

Buhler, Patricia
1997 "Scanning the environment: Environmental trends affecting the workplace." Supervision, 58: 24–26.

Chen, Peter Y., and Paul E. Spector
1992 "Relationships of work stressors with aggression, withdrawal, theft, and substance use: An exploratory study." Journal of Occupational and Organizational Psychology, 65: 177–184.

Dean, James W., Jr., and Scott A. Snell
1991 "Integrated manufacturing and job design: Moderating effects of organizational inertia." Academy of Management Journal, 34: 776–804.

Dearborn, DeWitt C., and Herbert A. Simon
1958 "Selective perception: A note on the department identifications of executives." Sociometry, 21: 140–144.

Donnellon, Anne
1993 "Crossfunctional teams in product development: Accommodating the structure to the process." Journal of Product Innovation Management, 10: 377–392.

Dovidio, John F., and Samuel L. Gaertner
1991 "Changes in the expression and assessment of racial prejudice." In Harry J. Knopke, Robert J. Norrell, and Ronald W. Rogers (eds.), Opening Doors: Perspectives on Race Relations in Contemporary America: 119–225. Tuscaloosa, AL: University of Alabama Press.

Eisenhardt, Kathleen M., and L. J. Bourgeois
1988 "Politics of strategic decision making in high-velocity environments." Academy of Management Journal, 32: 543–576.

Eisenhardt, Kathleen M., Jean L. Kahwajy, and L.J. Bourgeois
1997a "How management teams can have a good fight." Harvard Business Review, 75 (4): 77–85.
1997b "Conflict and strategic choice: How top management teams disagree." California Management Review, 39 (2): 42–62.

Eisenhardt, Kathleen M., and Claudia B. Schoonhoven
1990 "Organizational growth: Linking founding team strategy, environment, and growth among U.S. semiconductor ventures, 1978–1988." Administrative Science Quarterly, 35: 504–529.

Ely, Robin J.
1994 "The effects of organizational demographics and social identity on relationships among professional women." Administrative Science Quarterly, 39: 203–238.

Erickson, Bonnie H., and T. A. Nosanchuk
1977 Understanding Data. Toronto: McGraw-Hill Ryerson.

Evan, William
1965 "Conflict and performance in R&D organizations." Industrial Management Review, 7: 37–46.

Festinger, Leon
1954 "A theory of social comparison processes." Human Relations, 7: 117–140.

Fiske, Donald W., and Salvatore R. Maddi
1961 Functions of Varied Experience. Homewood, IL: Dorsey.

Fiske, Susan T., and Shelley E. Taylor
1991 Social Cognition. New York: McGraw-Hill.

Florin, Paul, Gary A. Giamartino, David A. Kenny, and Abraham Wandersman
1990 "Levels of analysis and effects: Clarifying group influence and climate by separating individual and group effects." Journal of Applied Social Psychology, 20: 881–900.

Gallagher, Dennis, Ed Diener, and Randy J. Larsen
1989 "Individual differences in affect intensity: A moderator of the relation between emotion and behavior." Unpublished manuscript, Department of Psychology, University of Illinois, Urbana-Champaign.

Georgopoulos, Basil S.
1986 Organizational Structure, Problem Solving, and Effectiveness. San Francisco: Jossey-Bass.

Ghosh, Sukesh K.
1991 Econometrics: Theory and Applications. Englewood Cliffs, NJ: Prentice-Hall.

Gladstein, Deborah L.
1984 "Groups in context: A model of task group effectiveness." Administrative Science Quarterly, 29: 499–517.

Guetzkow, Harold, and John Gyr
1954 "An analysis of conflict in decision-making groups." Human Relations, 7: 367–381.

Guo, Shumei S., William C. Chumlea, and David B. Cockram
1996 "Use of statistical methods to estimate body composition." American Journal of Clinical Nutrition, 64: 428S–435S.

Hambrick, Donald C.
1994 "Top management groups: A conceptual integration and reconsideration of the 'team' label." In B. M. Staw and L. L. Cummings (eds.), Research in Organizational Behavior, 16: 171–213. Greenwich, CT: JAI Press.

Hambrick, Donald C., and Phyllis A. Mason
1984 "Upper echelons: The organization as a reflection of its top managers." Academy of Management Review, 9: 193–206.

Hebb, D. O.
1955 "Drives and the CNS (conceptual nervous system)." Psychological Review, 62: 243–254.

Heriot, Gail
1996 "Diversity is more than skin deep." Los Angeles Times, July 14: M5.

Hoffman, L. Richard, and Norman R. F. Maier
1961 "Quality and acceptance of problem solutions by members of homogeneous and heterogeneous groups." Journal of Abnormal and Social Psychology, 62: 401–407.

Jackson, Susan E.
1992 "Team composition in organizational settings: Issues in managing an increasingly diverse workforce." In Stephen Worchel, Wendy Wood, and Jeffrey A. Simpson (eds.), Group Process and Productivity: 138–173. Newbury Park, CA: Sage.

Jackson, Susan E., Joan F. Brett, Valerie I. Sessa, Dawn M. Cooper, Johan A. Julin, and Karl Peyronnin
1991 "Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions, and turnover." Journal of Applied Psychology, 76: 675–689.

Janis, Irving L.
1982 Victims of Groupthink. Boston: Houghton Mifflin.
1989 Crucial Decisions: Leadership in Policymaking and Crisis Management. New York: Free Press.

Jehn, Karen A.
1994 "Enhancing effectiveness: An investigation of advantages and disadvantages of value-based intragroup conflict." International Journal of Conflict Management, 5: 223–238.
1995 "A multimethod examination of the benefits and detriments of intragroup conflict." Administrative Science Quarterly, 40: 256–282.

Kahneman, Daniel
1973 Attention and Effort. Englewood Cliffs, NJ: Prentice-Hall.

Kanter, Rosabeth M.
1977 Men and Women of the Corporation. New York: Basic Books.

Katz, Ralph
1982 "The effects of group longevity on project communication and performance." Administrative Science Quarterly, 27: 81–104.

Kennedy, Peter
1979 A Guide to Econometrics. Cambridge, MA: MIT Press.

Kent, R. N., and N. E. McGrath
1969 "Task and group characteristics as factors influencing group performance." Journal of Experimental Social Psychology, 5: 429–440.

Kirkpatrick, David L.
1986 "Performance appraisal: When two jobs are too many." Training: The Magazine of Human Resource Development, 23 (3): 65–68.

Kramer, Roderick M.
1991 "Intergroup relations and organizational dilemmas: The role of categorization processes." In L. L. Cummings and B. M. Staw (eds.), Research in Organizational Behavior, 13: 191–228. Greenwich, CT: JAI Press.

Lawrence, Barbara S.
1988 "New wrinkles in the theory of age: Demography, norms, and performance ratings." Academy of Management Journal, 31: 309–337.
1997 "The black box of organizational demography." Organization Science, 8: 1–22.

Major, B., and B. Forcey
1985 "Social comparisons and pay evaluations: Preferences for same-sex and same-job wage comparisons." Journal of Experimental Social Psychology, 21: 393–405.

Milberg, Sandra, and Margaret S. Clark
1988 "Moods and compliance." British Journal of Social Psychology, 27: 79–90.

Miller, C. T.
1982 "The role of performance-related similarity in social comparison of abilities: A test of the related attributes hypothesis." Journal of Personality and Social Psychology, 46: 1222–1228.

Milliken, Frances J., and Luis L. Martins
1996 "Searching for common threads: Understanding the multiple effects of diversity in organizational groups." Academy of Management Review, 21: 402–433.

Murnighan, J. Keith, and Donald E. Conlon
1991 "The dynamics of intense work groups: A study of British string quartets." Administrative Science Quarterly, 36: 165–186.

Murray, Alan I.
1989 "Top management group heterogeneity and firm performance." Strategic Management Journal, 10: 125–141.

Nelson, Reed E.
1989 "The strength of strong ties: Social networks and intergroup conflict in organizations." Academy of Management Journal, 32: 377–401.

O'Reilly, Charles A., David F. Caldwell, and William P. Barnett
1989 "Work group demography, social integration, and turnover." Administrative Science Quarterly, 34: 21–37.

Parker, Thomas H., and Ira J. Dolich
1986 "Toward understanding retail bank strategy: Seemingly unrelated regression applied to cross-sectional data." Journal of Retailing, 62: 298–320.

Pedhazur, Elazar J., and Liora P. Schmelkin
1991 Measurement, Design, and Analysis: An Integrated Approach. Hillsdale, NJ: Erlbaum.

Pelled, Lisa H.
1996 "Demographic diversity, conflict, and work group outcomes: An intervening process theory." Organization Science, 7: 615–631.

Pelled, Lisa H., Thomas G. Cummings, and Mark A. Kizilos
1999 "The influence of organizational demography on customer-oriented prosocial behavior: An exploratory investigation." Journal of Business Research, in press.

Pelled, Lisa H., Gerald E. Ledford, and Susan A. Mohrman
1998 "Individual demographic dissimilarity and organizational inclusion: A field investigation." Academy of Management Best Paper Proceedings, OMT Division: A1–A17.

Reardon, Kathleen K.
1995 They Don't Get It, Do They? Communication in the Workplace—Closing the Gap between Men and Women. Boston: Little, Brown.

Ross, Raymond S.
1989 Small Groups in Organizational Settings. Englewood Cliffs, NJ: Prentice-Hall.

Ryan, Alan J.
1991 "Job rotations grease long-term IS wheels." Computerworld, 25 (24): 76.

Sarason, Irwin G.
1984 "Stress, anxiety, and cognitive interference: Reactions to tests." Journal of Personality and Social Psychology, 46: 929–938.

Schoonhoven, Claudia B.
1981 "Problems with contingency theory: Testing assumptions hidden within the language of contingency 'theory'." Administrative Science Quarterly, 26: 349–377.

Sessa, Valerie I., and Susan E. Jackson
1995 "Diversity in decision-making teams: All differences are not created equal." In Martin M. Chemers, Stuart Oskamp, and Mark A. Costanzo (eds.), Diversity in Organizations: New Perspectives for a Changing Workplace: 133–156. Thousand Oaks, CA: Sage.

Smith, Ken G., Ken A. Smith, Judy D. Olian, Henry P. Sims, Douglas P. O'Bannon, and Judith A. Scully
1994 "Top management team demography and process: The role of social integration and communication." Administrative Science Quarterly, 39: 412–438.

South, Scott J., Charles M. Bonjean, William T. Markham, and Judy Corder
1982 "Social structure and intergroup interaction: Men and women of the federal bureaucracy." American Sociological Review, 47: 587–599.

Staw, Barry M., Lloyd E. Sandelands, and Jane E. Dutton
1981 "Threat-rigidity effects in organizational behavior: A multilevel analysis." Administrative Science Quarterly, 26: 501–524.

Storms, Philip L., and Paul E. Spector
1987 "Relationships of organizational frustration with reported behavioural reactions: The moderating effect of locus of control." Journal of Occupational Psychology, 60: 227–234.

Tajfel, Henri
1972 "Social categorization." In Serge Moscovici (ed.), Introduction à la Psychologie Sociale, 1: 272–302. Paris: Larousse.
1978 Differentiation between Social Groups. New York: Academic Press.
1982 Social Identity and Intergroup Relations. Cambridge: Cambridge University Press.

Tajfel, Henri, M. G. Billig, R. P. Bundy, and C. Flament
1971 "Social categorization and intergroup behavior." European Journal of Social Psychology, 1: 149–177.

Teachman, Jay D.
1980 "Analysis of population diversity." Sociological Methods and Research, 8: 341–362.

Tjosvold, Dean
1986 "Constructive controversy: A key strategy for groups." Personnel, 63: 39–44.
1991 Team Organization: An Enduring Competitive Advantage. New York: Wiley.

Tsui, Anne S., Susan J. Ashford, Lynda St. Clair, and Katherine R. Xin
1995 "Dealing with discrepant expectations: Response strategies and managerial effectiveness." Academy of Management Journal, 38: 1515–1543.

Turner, John C.
1975 "Social comparison and social identity: Some comparisons for intergroup behavior." European Journal of Social Psychology, 5: 5–34.

Van de Ven, Andrew H., Andre Delbecq, and R. Koenig
1976 "Determinants of coordination modes within organizations." American Sociological Review, 41: 322–338.

Vernon, J., and T. E. McGill
1960 "Utilization of visual stimulation during sensory deprivation." Perceptual and Motor Skills, 11: 214.

Wagner, W. Gary, Jeffrey Pfeffer, and Charles A. O'Reilly
1984 "Organizational demography and turnover in top management groups." Administrative Science Quarterly, 29: 74–92.

Waller, Mary J., George P. Huber, and William H. Glick
1995 "Functional background as a determinant of executives' selective perception." Academy of Management Journal, 38: 943–974.

Walsh, James P.
1988 "Selectivity and selective perception: An investigation of managers' belief structures and information processing." Academy of Management Journal, 31: 873–896.

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

Weingart, Laurie
1992 "Impact of group goals, task component complexity, effort, and planning on group performance." Journal of Applied Psychology, 77: 682–693.

Wheaton, B., B. Muthen, D. Alwing, and G. Summers
1977 "Assessing reliability and stability in panel models." In D. Heise (ed.), Sociological Methodology: 84–136. San Francisco: Jossey-Bass.

Wickens, Christopher D.
1989 "Attention and skilled performance." In D. Holding (ed.), Human Skills: 71–105. New York: Wiley.

Wiersema, Margarethe F., and Karen Bantel
1992 "Top management team demography and corporate strategic change." Academy of Management Journal, 35: 91–121.

Wonnacott, Thomas H., and Ronald J. Wonnacott
1984 Introductory Statistics for Business and Economics. New York: Wiley.

Zellner, Arnold
1962 "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." Journal of the American Statistical Association, 57: 348–368.

Zenger, Todd R., and Barbara S. Lawrence
1989 "Organizational demography: The differential effects of age and tenure distributions on technical communication." Academy of Management Journal, 32: 353–376.

Zimbardo, Philip G., and Michael R. Leippe
1991 The Psychology of Attitude Change and Social Influence. New York: McGraw-Hill.

Zuckerman, Marvin
1979 Sensation Seeking. Hillsdale, NJ: Erlbaum.
1984 "Sensation seeking: A comparative approach to a human trait." Behavioral and Brain Sciences, 7: 413–471.