Impact of highly and less job-related diversity on work group cohesion and performance: a meta-analysis

Sheila Simsarian Webbera,* Lisa M. Donahueb

aConcordia University, Department of Management, John Molson School of Business. 1455 de Maisonneuve Blvd. West, Montreal, Quebec H3G 1M8, Canada
bGeorge Mason University, Department of Psychology, MS 3175, Fairfax, VA 22030-4444, USA

Received 5 August 1999; received in revised form 5 April 2000; accepted 11 August 2000

Abstract

A meta-analysis of the data from empirical investigations of diversity in work groups was used to examine the impact of two types of diversity attributes, highly job-related and less job-related, on work group cohesion and performance. This distinction was used to test the proposition that different types of diversity will differentially impact work group cohesion and performance. In addition, type of team was examined as a possible moderator of the relationship between diversity and performance. Results showed that neither type of diversity had a relationship with cohesion or performance. Explanations and directions for future research are offered. © 2001 Elsevier Science Inc. All rights reserved.

1. Introduction

Diversity is increasing within organizations at an astronomical rate. Corporations are becoming global, building strategic alliances, and facilitating mergers and acquisitions inside and outside their primary domain of work (Jackson, May, & Whitney, 1995). Concurrently, organizations are implementing work teams with greater frequency to integrate the knowledge of workers across broad specializations, as well as geographic locations (Sundstrom, DeMeuse, & Futrell, 1990). The result is that two of the most complex phenomena in the work place, diversity and work teams, are being merged together with the goal of creating more innovative, high-performing organizations.

* Corresponding author. Tel.: +1-514-848-2902; fax: +1-514-848-4292.
E-mail addresses: swebber@vax2.concordia.ca (S.S. Webber), ldonahue@gmu.edu (L.M. Donahue).
The complexity of these phenomena present challenges to researchers in understanding how diversity influences important group outcomes. More specifically, much is still unknown regarding the nature of diversity, its impact on work group outcomes, and the intervening mechanisms by which it influences these outcomes (Pellet, Eisenhardt, & Xin, 1999). Recent narrative reviews of the effects of diversity in work groups by Jackson et al. (1995) and Milliken and Martins (1996) have contributed to our understanding of diversity in work teams. Much of this research has concluded that diversity can be a “double-edged sword.” Specifically, examinations of diversity in work groups have found that it can lead to more high-quality solutions, while also decreasing cohesion (Milliken & Martins, 1996; McLeod & Lobel, 1992; Watson, Kumar, & Michaelson, 1993; O’Reilly, Caldwell, & Barnett, 1989).

However, a recent review of the empirical research finds inconsistency among the studies; therefore, conclusive findings of the effects of diversity on work group processes and outcomes still do not exist (Cohen & Bailey, 1997).

One reason for inconsistent findings in past research may be the lack of theoretical guidance for understanding the impact of different types of diversity in work groups. Although a number of demographic attributes (herein referred to as “diversity attributes”) exist upon which a work group can demonstrate heterogeneity, until recently little consideration has been given to the idea that diversity attributes may operate differently to impact work group outcomes. As a result, earlier researchers often used a single theoretical argument to propose that all types of diversity would have a particular effect on group processes and performance. For example, researchers have used the argument of cognitive resource diversity to propose a positive relationship between all diversity attributes and performance (Wiersema & Bantel, 1992). Clearly, diversity in certain attributes offers greater cognitive resources to the group than others, and will thus differentially impact group performance. Until consideration is given to the impact of different types of diversity in work groups, further advancements in our understanding of this phenomenon will be difficult to achieve.

2. Present research

In the present research, we conduct a meta-analysis to examine the impact of diversity on work group cohesion and performance. This research imposes a theoretically grounded classification system on the study of diversity in work groups. In recent narrative reviews, Jackson et al. (1995) and Milliken & Martins (1996) have offered a distinction between readily detectable and less observable diversity, in which the former represents diversity attributes that are easily observable and likely to evoke responses that result from basic social categorization. However, this distinction was provided in an attempt to organize the existing diversity literature and does not capture the underlying mechanisms by which various diversity attributes may differentially impact work group outcomes.

More recently, Pellet (1996; Pellet et al., 1999; Simons, Pellet, & Smith, 1999) distinguished among diversity attributes as they relate to the performance of work groups. Specifically, Pellet conceptualized work group diversity in terms of highly job-related and less job-related attributes, in which job relatedness is the degree to which the attribute
captures experiences, skills, or perspectives pertinent to cognitive work tasks. Job-relatedness is theoretically important because it describes whether a type of diversity will increase the task-related knowledge, skills, and abilities (KSAs) that facilitate more effective team performance (Simons et al., 1999). Pelled argued that diversity attributes such as functional, educational, or industry background capture experiences and perspectives relevant to the tasks performed by most work groups. As such, this type of diversity is proposed by Pelled and others (Sessa & Jackson, 1995; Milliken and Martins, 1996) to have a stronger impact on the task-relevant group processes and performance.

In contrast, Pelled (1996; Pelled et al., 1999) considered diversity attributes such as age, gender, and race as less germane to the group's task. Although these attributes may reflect a broad set of experiences, they are expected to be less related to the work being performed. As suggested by Zenger and Lawrence (1989: cited in Pelled et al., 1999), "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." Instead, these attributes form the context of more general social relationships and are less directly associated with team objectives (Sessa & Jackson, 1995).

Pelled's (1996) classification has received some initial empirical support. In a recent study, Simons et al. (1999) found that more job-related diversity (education-level, company tenure, and perceived environmental uncertainty) interacted with debate to influence top management team (TMT) performance. Less job-related diversity (age diversity), however, did not. They interpreted these findings as demonstrating that debate is more likely to impact team outcomes "... when it draws on different experiences and perspectives that are relevant to a task rather than on less relevant viewpoint differences" (p. 670).

In the present research, we adopt Pelled's (1996) classification system to determine the impact of diversity attributes on cohesion and performance through a meta-analytic review of the diversity research. Our focus on these particular proximal and distal outcomes was to test characterizations of diversity as a "double-edged sword" (Milliken & Martins, 1996). Otherwise stated, our goal is to use meta-analysis to determine if diversity does indeed increase work group performance, in spite of decreasing cohesion. Utilizing meta-analysis in combination with Pelled's classification system, we expected to be able to draw sounder conclusions regarding the effects of work group diversity that heretofore have not been reached by narrative reviews of the literature.

For this research, we focus only on studies specific to work groups and teams. [Note: Although some researchers distinguish between work group and teams, we use the words "team" and "work group" interchangeably, with "work group" used more often. Both terms are used to refer to collections of individuals who are interdependent, share responsibility for outcomes, and are viewed as an intact social entity (see Cohen & Bailey, 1997).] While the literature reviews of diversity often include studies of supervisor/subordinate dyads and broad organizational groups such as departments, we believe that these groups are quite different from work groups. Therefore, our approach is to specifically examine diversity in work groups.
3. Theoretical background and hypotheses

Recently, McGrath, Berdahl, & Arrow (1995) summarized the underlying models used to study the effects of diversity in work groups. The primary model relied on by researchers is the trait model, which presumes that diversity on demographic characteristics implies diversity on underlying attributes such as task-related KSAs; values, beliefs, and attitudes; and personality.

The prevailing theoretical approach used by diversity researchers in explaining the positive effects of diversity on work group performance is one example of a trait model. Specifically, this cognitive resource perspective suggests that the group’s composition is an indicator of diversity in their knowledge bases and perspectives. Thus, according to this perspective, as demographic heterogeneity increases, so does the group’s cognitive resources and ability to engage in more complex problem-solving (Hambrick & Mason, 1984; Wiersema & Bantel, 1992; Jackson et al., 1995; McLeod, Lobel, & Cox, 1996; Watson et al., 1993). As stated by Milliken and Martins (1996), “a group that is diverse could be expected to have members who may have had significantly different experiences, and therefore, significantly different perspectives on key issues and problems” (p. 404).

Pelled (1996; Pelled et al., 1999; Simons et al., 1999) also has implied a trait model of diversity by proposing that job-related diversity attributes reflect the underlying experiences most relevant to the tasks performed by work groups. However, Pelled’s propositions regarding the impact of highly job-related diversity on group performance also find support in the expectations model of diversity (McGrath et al., 1995; to be discussed later). Moreover, the expectations model provides theoretical justification for explaining why less job-related diversity attributes can be expected to result in less social cohesion. Thus, we rely mainly on the expectations model to support hypotheses regarding the relationship of these two types of diversity with work group cohesion and performance.

3.1. Diversity attributes and expectations

Unlike the trait model, the expectations model does not propose a direct linkage between diversity in demographic characteristics and diversity in underlying attributes (e.g., task-related KSAs; values, beliefs, and attitudes; and personality). Instead, the expectations model suggests that underlying attributes are inferred by other group members on the basis of the target member’s demographic characteristics. These inferences then shape the expectations that other group members have of the target member’s behavior, often leading to their differential treatment of this member (McGrath et al., 1995).

The expectations approach directly implies social categorization processes as the mechanism by which such inferences regarding others’ underlying attributes are made. Social categorization is the process by which we place ourselves and others into different social categories. We then use these social categories as the basis for inferring others’ underlying attributes (Fiske & Taylor, 1991; Klimoski & Donahue, 2001). For instance, social category membership is an important basis for inferences regarding what others know (Nickerson, 1999). Thus, members of a TMT are likely to conclude that a particular member is a finance expert, simply because his/her educational and functional background is in this area.
The specific characteristics on which we tend to focus in categorizing others are likely to be those that are the most distinctive or salient within the social context (Nelson & Klatz, 2000). Once categorization occurs, however, we tend to think of others not as unique individuals, but as embodiments of the relevant group prototype to which they belong (Hogg & Terry, 2000). Moreover, these prototyped-based expectations will tend to become more pronounced and stereotypic as the demographic attribute in question becomes more salient (McGrath et al., 1995). All of this has direct implications for whether highly job-related or less job-related diversity attributes will have more of an impact on group processes and performance. Specifically, we argue that within the context of the task being performed by the work team, not only will highly job-related diversity attributes become more salient, but the resulting prototypes based on these attributes are more likely to lead to appropriate inferences regarding task-related KSAs. The correct identification of group members with relevant expertise and experience has been shown to be an important factor in the performance of problem-solving groups (e.g., Libby, Trotman, & Zimmer, 1987; Littlepage, Schmidt, Whisler, & Frost, 1995; Littlepage, Robinson, & Reddington, 1997).

Conversely, within the context of the group's social relationships, less job-related attributes are more likely to become salient. Thinking about group members prototypically in terms of these attributes may have a detrimental effect on cohesion. This is mainly because, according to social identity theory (Tajfel & Turner, 1979), once social categorization takes place individuals become motivated by underlying needs for self-esteem to develop more positive opinions of their own category (in-group) and negative opinions of other categories (out-group). In addition, the prototypical (stereotypical) attributes typically used to represent different social categories are those that maximize similarities within and differences between groups (Hogg & Terry, 2000). As a result, social categorization into in-group and out-group categories will further accentuate the perceived similarity to the in-group and dissimilarity to the out-group. As suggested by Tajfel and Turner (1986) "the mere perception of belonging to two distinct groups—that is, social categorization per se—is sufficient to trigger intergroup discrimination favoring the in-group" (p. 26). Such discrimination has implications for a number of group phenomena, including cohesion.

In sum, trait models of diversity, such as the cognitive resource perspective, can be used to explain why highly job-related diversity is expected to lead to improved group performance. These models, however, offer limited guidance regarding the impact of diversity attributes on the cohesion of the work team. To better understand the relationship of diversity with both cohesion and performance, we draw upon the expectations models of diversity to propose that highly and less job-related diversity attributes will differentially influence these outcomes.

3.2. Work group diversity and performance

Linking the diversity of work teams to team performance has been the focus of a number of different studies both in the TMT literature and in project team and lower-level team research (e.g., Smith, Smith, Olian, Sims, O'Bannon, & Scully, 1994; Ancona & Caldwell, 1992; Pelled et al., 1999). Using the theoretical argument of cognitive resource diversity, researchers in this area have argued that diversity on all demographic attributes facilitates
performance due to the increase in the cognitive resources offered to the team (Hambrick, Cho, & Chen, 1996; Wiersema & Bantel, 1992). However, empirical studies of this relationship have found mixed results. Specifically, Wiersema and Bantel (1992) found a negative relationship between diversity and performance. Bantel (1994), however, found a positive relationship between diversity and performance, and Ancona and Caldwell (1992) found no relationship between diversity and performance. One reason for these mixed findings is that different types of diversity may have different impacts on the performance of the work group. As indicated, Pelled (1996; Pelled et al., 1999) has proposed that because highly job-related diversity attributes are more related to the group’s task, these attributes will have a greater impact on performance than less job-related attributes. Thus, stronger conclusions may be drawn by separating out the various diversity attributes examined in previous studies in terms of their job-relatedness.

On the basis of both the trait and expectations models of diversity, we propose that highly job-related diversity will be more influential on the performance of work teams. First, consistent with Pelled (1996; Pelled et al., 1999; Simons et al., 1999), we argue that the diversity attributes which are highly job-related, encompassing experiences and knowledge pertinent to the task, will increase the underlying task-related KSAs needed for successful performance. Second, we argue that the task context is more likely to provoke categorizations in terms of these attributes. Otherwise stated, job-related attributes are expected to become more salient to group members, and therefore more likely to be used in categorizing group members. In contrast, less job-related types of diversity are expected to not only be less salient, but to also provide less informative cues for inferring others’ KSAs. More specifically, categories such as age, race, and gender are broader, and thus are expected to result in much less specific estimates regarding what others know (Nickerson, 1999). Thus, in groups where effective group process interaction depends on being able to identify group members with relevant KSAs (Libby et al., 1987; Littlepage et al., 1997), job-related diversity attributes are more likely to result in inferences that trigger relevant group processes that, in turn, lead to effective performance.

Relying on the theoretical underpinnings of both the trait and expectations models of diversity, we hypothesize the following relationship between highly job-related diversity and performance:

H1: Highly job-related diversity will have a stronger positive relationship with performance than will less job-related diversity.

3.3. Work group diversity and cohesion

Cohesion has been proposed to be an important determinant of work group performance (Evans & Dion, 1991; Gully, Devine, & Whitney, 1995; Mullen & Cooper, 1994). Shaw (1981) defines cohesion as the degree to which members of the group are attracted to each other. As such, it is proposed by O’Reilly et al. (1989) to be an essential component of the group’s overall social integration, or the “attraction to the group, satisfaction with other members of the group, and social interaction among the group members” (p. 22).

A major source of cohesiveness is interpersonal attraction (Festinger, Schachter, & Back.
1950), and diversity researchers have often drawn upon the similarity-attraction hypothesis to explain the effects of diversity on cohesion. Their basic argument is that the perception of similarity in attitudes, as inferred on the basis of similarity in demographic attributes, leads to attraction among group members. Thus, work groups that are relatively homogenous in demographic attributes will experience greater cohesion than those that are more heterogeneous (O'Reilly et al., 1989; Smith et al., 1994; Wiersema & Bantel, 1992). In suggesting that demographic attributes influence inferences regarding the similarity of other group members in terms of underlying values and beliefs, as well as the behavior toward these other members, these researchers use arguments that are consistent with the expectations model of diversity.

Empirical investigations examining the impact of diversity on the cohesion of work groups have found mixed results. For example, Harrison, Price, & Bell (1998) found a negative relationship between diversity and cohesion, while Smith et al. (1994) found no relationship between diversity and cohesion. In response to these mixed findings, researchers have suggested that the effects of diversity on cohesion may differ by the type of diversity (Milliken & Martins, 1996).

Similar arguments regarding the differential effects of different types of diversity have been made by Jehn, Northcraft, & Neale (1999) and Pelled et al. (1999) regarding emotional conflict. These researchers have suggested that within the context of social relationships, some types of diversity will become more salient and provoke social categorizations that have implications for emotional conflict. Specifically, Pelled et al. have proposed that basic impermeable attributes such as race, gender, and age (which they also classify as less job-related attributes because of their greater relevance to the group's social relationships than to their task) are more likely to lead to social categorizations that yield intercategory clashes. This is because group members cannot easily move in and out of these categories, as they can with attribute categories such as educational or functional background. According to Pelled et al., "When attributes are not easily penetrated, it is difficult for employees to 'stand in the shoes' of those in another category" (p. 5). Without the benefit of similarity in social categories, it is more difficult for group members to make inferences regarding underlying attributes, such as attitudes, beliefs, and values. Consequently, they are expected to rely on more prototyped-based representations of group members who belong to other social categories. As previously stated, this can further polarize the groups, leading to increased in-group favoritism and out-group discrimination. While Jehn et al. (1999) and Pelled et al. (1999) argue that intergroup discrimination based on social categorization processes can surface as emotional conflict, we argue that it can also lead to decreased cohesion.

Adopting expectations-based theoretical arguments from earlier diversity research that relied on the similarity-attraction hypothesis and more recent research examining the effects of diversity on group conflict, we propose that those attributes that are less job-related (more impermeable) will have a stronger impact on the cohesiveness of the work team than those aspects that are highly job-related. Thus, we hypothesize the following:

H2: Less job-related diversity will have a stronger negative relationship with cohesion than will highly job-related diversity.
3.4. Potential moderators

A variety of work groups and teams are employed in organizations today, and researchers have emphasized the distinctions among them (see Cannon-Bowers, Oser, & Flanagan, 1992; Cohen & Bailey, 1997; Sundstrom et al., 1990). For example, McGrath and colleagues (Arrow & McGrath, 1995: McGrath et al., 1995) suggest that teams can be distinguished in terms of differences in their members, tasks, and tools. Using this conceptual framework, McGrath and colleagues have defined three types of work groups—teams, task forces, and crews. Other researchers (e.g., Cohen & Bailey, 1997; Sundstrom et al., 1990) have further differentiated the team category into production teams, project teams, action/involvement teams (e.g., task forces, committees), and top management teams.

Because of their membership differences, the type of team may moderate the relationship between diversity and outcomes such as cohesion and performance. As pointed out by one reviewer, teams may differ in terms of the degree to which they exhibit heterogeneity on certain diversity attributes. We certainly agree with this assertion. Members of top management teams, project teams, and action/involvement teams, for example, are more likely to be heterogeneous on highly job-related attributes such as functional and educational background, but are less likely to be heterogeneous on less job-related attributes such as age, race, and gender. In contrast, production teams are more likely to exhibit heterogeneity on less job-related attributes and less likely to demonstrate heterogeneity on highly job-related attributes. Given these expected compositional differences and our earlier predictions, the relationship between highly job-related diversity and performance may be stronger in top management, project, and action/involvement teams—or teams that are more heterogeneous on highly job-related diversity attributes. On the other hand, the relationship between less job-related diversity and cohesion may be stronger for production teams, simply because heterogeneity on these attributes is likely to be greater in these types of teams.

While meta-analytic tests of the above propositions would serve to advance our understanding of how diversity operates in various types of teams, current limitations in the work group diversity literature do not allow for such analysis. In general, much of this research on work group diversity has focused on TMTs as a team type, mainly as a result of the interest in relational demography sparked by Hambrick and Mason’s (1984) seminal work. Only recently have researchers increased their examination of diversity’s effects in lower-level organizational groups, primarily project teams. Further, the studies involving TMTs have concentrated on the effects of diversity on performance; very few have examined cohesion as an outcome. These limitations in the work group diversity research make it difficult to investigate team type as a potential moderator of the relationship of highly and less job-related diversity with cohesion.

Examining team type as a moderator of the relationship between diversity and performance is also difficult because team type is confounded with the level of the outcome used to measure the team’s performance. In the case of TMTs, performance is measured in terms of organizational financial performance (e.g., return on investment, sales growth, etc.). This approach follows from upper echelons theory (Hambrick & Mason, 1984), which proposes that upper-level managers have an important impact on organizational outcomes. In lower-level teams, including project teams and production teams; however, the outcome measured
is team performance (e.g., quality and quantity of work) based on subjective ratings by supervisors or team leaders.

Distinguishing teams according to outcome level is important for understanding the relationship between diversity and performance. This is because the performance outcomes in question differ in how proximally or distally related they are to the team’s interaction processes. Specifically, input-process-output (IPO) models of team effectiveness (e.g., Hackman, 1987) suggest that group interaction directly influences group-level performance. However, a variety of other factors in addition to a TMT’s group processes can also have a strong impact on the organizational outcomes used to judge their performance, including competition, government regulatory interference, economic conditions, interest rate fluctuation, and currency fluctuations. Further the subjectivity of the ratings provided for lower-level teams may also be impacted by rater stereotypes and biases, causing inflated relationships. Thus, because of the proximal and subjective nature of the group-level measures used to evaluate the performance of lower-level teams, they are likely to have a more direct impact on their performance outcomes than TMTs. For these reasons, we distinguish TMTs from all lower-level teams (i.e., project and production teams), and propose the following hypothesis:

H3: The relationship between diversity and performance will be stronger for lower-level teams than for TMTs.

4. Method

4.1. Literature search

A literature search was conducted to identify both published and unpublished studies of diversity in work groups from 1980 to the present. This time span was chosen because of the surge of interest in the study of teams in the work place from the 1980s forward (Ilgen, Major, Hollenbeck, & Sego, 1993; McGrath, Arrow, & Berdahl, 2000; Sundstrom et al., 1990). Six strategies were employed to search the relevant literature. First, a computer search was done of PsycINFO (1980–1999), PsycLit (1980–1999), and ABI/INFORM (1980–1999). The specific words used to search these databases included team/group heterogeneity, team/group composition, team/group diversity, and team/group relational demography. Second, a manual search was conducted that consisted of checking the sources cited in the reference sections of literature reviews, articles, and books on this topic. Two specific sources were relied upon most heavily, a chapter by Jackson and her colleagues (Jackson et al., 1995) appearing in the text Team Effectiveness and Decision Making in Organizations, and a review chapter by Milliken and Martins (1996) published in the Academy of Management Review. Third, we manually checked five top-tiered journals to identify any relevant empirical research in this area in the past five years. We paid particular attention to the reference sections of these articles in order to identify unpublished studies. These journals included Academy of Management Journal, Journal of Applied Psychology, Administrative Science Quarterly, Personnel Psychology, and Journal of Product Innovation Management. Fourth, we manually searched the Society for Industrial and Organizational Psychology and
Academy of Management conference programs (1995–1999) and requested papers based on the title and abstract. The manual searches of journals and conference programs were conducted mainly to identify any relevant research that was unpublished or had not completed the publication process. Given that most recent research would complete the publication process in three to five years, we did not extend our manual search past these time frames. Fifth, we searched Dissertation Abstracts (1980–1999) to identify any relevant, unpublished dissertation research. Finally, we personally contacted four prominent researchers in this area to request unpublished research or papers currently under review.

Together, these six strategies yielded 76 studies. Only studies that involved groups or teams that measured one or more of the independent and dependent variables under investigation (described below) were included. In addition, we included only studies that aggregated the independent variables to the team level. We did not include larger organizational entities such as work units and departments because of our interest in investigating this phenomenon at the work group or team level. We also included laboratory studies conducted on samples (e.g., with undergraduate and graduate business students) that did not seriously limit the generalizability of the findings. Finally, we excluded studies that utilized organizational tenure, as there is less agreement among researchers regarding its relationship to task performance (see Pelled et al., 1999). The final number of studies used in the meta-analysis was 24, yielding a total of 45 correlations. Table 1 includes a list of the studies and their correlations categorized by diversity type.

4.2. Variables included in the analysis

The independent variables included diversity in age, gender, race/ethnicity, educational and functional background, and industry and occupational background. All of these variables were aggregated to the team level using an appropriate measure of heterogeneity. The two most frequently used approaches were Blau’s (1977) index of heterogeneity and Allison’s (1978) coefficient of variation. Blau’s index is used for categorical variables, \(1 - \Sigma p_i^2\), where \(p_i\) is the proportion of the group in the particular demographic category and \(i\) is the number of groups represented. The coefficient of variation is used for continuous variables and is calculated by dividing the standard deviation by the mean.

We categorized age, gender, and race/ethnicity as less job-related diversity attributes; and we categorized educational, functional, occupational, and industry background as highly job-related diversity attributes. Again, this was done to examine the effects of Pelled’s (1996) theoretically grounded classification system of diversity on work group cohesion and performance.

The dependent variables included cohesion as the proximal outcome of work group diversity. In all cases, this outcome was assessed using self-report measures administered to the team members or the team leader (e.g., CEO). Of the five studies that utilized a measure of cohesion, one study (O’Reilly et al., 1989) measured social integration, a multifaceted measure of which cohesion is one element. Two of the studies utilized the definition of social integration provided by O’Reilly et al., but measured only items from the cohesion scale. The remaining two studies proposed a similar definition as O’Reilly et al., but utilized different measures of cohesion.
Table 1
Studies and correlations used in the meta-analysis

<table>
<thead>
<tr>
<th>Empirical studies</th>
<th>Cohesion</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less job-related diversity</td>
<td>Highly job-related diversity</td>
</tr>
<tr>
<td>1. Ancona &amp; Caldwell (1992)</td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td>2. Bantel (1993)</td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td>3. Bantel (1994)</td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td>4. Campion et al. (1993)</td>
<td>Functional</td>
<td>.21</td>
</tr>
<tr>
<td>5. Campion et al. (1996)</td>
<td>Functional</td>
<td>.31</td>
</tr>
<tr>
<td>7. Hambrick et al. (1996)</td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td>8. Harrison et al. (1998)</td>
<td>Age</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>Race/Ethnic</td>
<td>.12</td>
</tr>
<tr>
<td>9. Kosnik (1990)</td>
<td>Industry</td>
<td></td>
</tr>
<tr>
<td>10. Lant et al. (1992)</td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Occupational</td>
<td></td>
</tr>
<tr>
<td>11. Mayo et al. (1996)</td>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Race/Ethnic</td>
<td>-0.18</td>
</tr>
<tr>
<td>12. Mazzeski &amp; DiStefano (1996)</td>
<td>Age</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>Race/Ethnic</td>
<td>.12</td>
</tr>
<tr>
<td>15. O'Reilly et al. (1989)</td>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>16. Pelled et al. (1999)</td>
<td>Age</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>Race/Ethnic</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td>17. Simons (1995)</td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Occupational</td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Empirical studies</th>
<th>Cohesion</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less job-related diversity</td>
<td>Highly job-related diversity</td>
</tr>
<tr>
<td></td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td>19. Smith et al. (1994)</td>
<td>- .319</td>
<td>- .224</td>
</tr>
<tr>
<td></td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td>20. Sutcliffe (1994)</td>
<td>.019</td>
<td>.186</td>
</tr>
<tr>
<td></td>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td>21. Watson et al. (1993)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Racial/Ethnic</td>
<td></td>
</tr>
<tr>
<td>22. Wiersema &amp; Bantel (1992)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>23. Wiersema &amp; Bird (1993)</td>
<td></td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Educational</td>
<td></td>
</tr>
<tr>
<td>24. Williams et al. (1995)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional</td>
<td></td>
</tr>
</tbody>
</table>

The distal outcome used in the meta-analysis was performance. In studies of top management teams, the primary performance variable reported was organizational financial performance, measured as return on investment (ROI), return on equity (ROE), average return on assets (ROA), profitability during the year, and/or increase in sales. In research on lower-level teams, the performance variable was team performance, measured in terms of the quality and/or quantity of work as evaluated by team leaders or supervisors.

Finally, in order to test for potential moderators, the type of team was coded as either TMT or lower-level.

4.3. Meta-analysis procedure

The meta-analysis performed in this study used the formulas available in MetaQuick (Stauffer, 1998), which follow the Hunter and Schmidt (1990) protocol for correcting the mean and variance of correlation coefficients across studies for artifactual variance due to sampling error, range restriction, and attenuation due to measurement error. Corrections for unreliability were made for cohesion. Only lower-level team studies, however, reported reliabilities for performance, because for TMT studies, performance was measured at the organizational level by objective indicators (e.g., return on investment). For lower-level teams, this information was used to correct for measurement error. No study reported information related to range restriction.

When correlation coefficients between the diversity categories and cohesion and perfor-
Table 2
Meta-analysis results for cohesion and performance

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Total N</th>
<th>Number Of r's</th>
<th>Obs $\rho$</th>
<th>SD$_\rho$</th>
<th>95% C.V.</th>
<th>95% C.I.</th>
<th>% of Artifact variance accounted</th>
<th>Q</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohesion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>479</td>
<td>8</td>
<td>.03</td>
<td>.04</td>
<td>.14</td>
<td>.35, .27</td>
<td>.13, .05</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Performance</td>
<td>2257</td>
<td>37</td>
<td>-.01</td>
<td>-.01</td>
<td>.08</td>
<td>.15, .18</td>
<td>.03, .05</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>Combined diversity types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohesion</td>
<td>233</td>
<td>4</td>
<td>-.02</td>
<td>-.03</td>
<td>.00</td>
<td>-.03, .03</td>
<td>.10, -.16</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>622</td>
<td>13</td>
<td>-.06</td>
<td>-.07</td>
<td>.08</td>
<td>-.08, .23</td>
<td>.00, -.15</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Less job-related diversity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Cohesion</td>
<td>246</td>
<td>4</td>
<td>.08</td>
<td>.10</td>
<td>.19</td>
<td>.51, -.30</td>
<td>.23, -.02</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>1635</td>
<td>24</td>
<td>.01</td>
<td>.02</td>
<td>.06</td>
<td>.15, -.12</td>
<td>.07, -.03</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Highly job-related diversity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

Note. Total N = total number of teams in sample; number of r's = the number of correlation coefficients on which each distribution is based; Obs $\rho$ = the uncorrected mean correlation; $\rho$ = the estimated true correlation; SD$_\rho$ = the estimated true residual standard deviation; 95% C.V. = the upper and lower bound of the 95% credibility value for each distribution based on its true correlation and SD$_\rho$ estimated; 95% C.I. = the upper and lower bound of the 95% confidence interval for each distribution; Q = homogeneity statistic.

* $p < .05$.

...mer were not provided, the conversion formulas presented by Nouri & Greenberg (1995) and Morris & DeShon (1997) were used to compute effect sizes for conversion to correlation coefficients.

To examine the proposed moderated relationships, we employed the approaches recommended by Hunter and Schmidt (1990) and Hedges and Olkin (1985; Aguinis & Pierce 1998). According to Johnson, Mullen, and Salas (1995), the Hedges and Olkin approach for assessing variability tends to yield higher indications of variability among effect sizes compared to the Hunter and Schmidt approach, resulting in a greater chance of detecting moderator effects. Further, the modified Hedges and Olkin approach proposed by Aguinis and Pierce provides a statistical test of meta-analytic moderator variable hypotheses by calculating between-subgroup homogeneity.

5. Results and interpretation

The number of correlations upon which the meta-analyses were conducted is shown in Tables 2 and 3 for the diversity categories, cohesion, and performance. Tables 2 and 3 also present the results of the separate meta-analyses. The first seven columns of the tables contain, respectively, the total number of teams in the sample, the number of correlation coefficients on which each distribution is based, the uncorrected (i.e., observed) correlation, the estimated true correlation ($\rho$), the estimated true residual standard deviation (SD$_\rho$), the lower and upper bound of the 95% credibility value for each distribution, based on its true correlation and SD$_\rho$ estimates, and the lower and upper bound of the 95% confidence interval...
Table 3
Meta-analysis results for team type

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Total N</th>
<th>Number Of r’s</th>
<th>Obs R</th>
<th>ρ</th>
<th>SDₚ</th>
<th>95% C.V.</th>
<th>95% C.I.</th>
<th>% of Artifact variance accounted</th>
<th>Q</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less job-related diversity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMTs</td>
<td>184</td>
<td>3</td>
<td>-.07</td>
<td>-.07</td>
<td>.00</td>
<td>-.07, -.07</td>
<td>.07, -.22</td>
<td>100</td>
<td>1.06</td>
<td>2</td>
</tr>
<tr>
<td>Lower-level teams</td>
<td>438</td>
<td>10</td>
<td>-.06</td>
<td>-.07</td>
<td>.12</td>
<td>.17, -.31</td>
<td>.02, -.16</td>
<td>62</td>
<td>19.71*</td>
<td>9</td>
</tr>
<tr>
<td>Highly job-related diversity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMTs</td>
<td>1405</td>
<td>20</td>
<td>.03</td>
<td>.03</td>
<td>.07</td>
<td>.17, -.11</td>
<td>.08, -.02</td>
<td>74</td>
<td>27.59</td>
<td>19</td>
</tr>
<tr>
<td>Lower-level teams</td>
<td>232</td>
<td>4</td>
<td>-.07</td>
<td>-.09</td>
<td>.00</td>
<td>-.09, -.09</td>
<td>.04, -.22</td>
<td>100</td>
<td>2.10</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. Total N = total number of teams in sample; number of r’s = the number of correlation coefficients on which each distribution is based; Obs R = the uncorrected mean correlation; ρ = the estimated true correlation; SDₚ = the estimated true residual standard deviation; 95% C.V. = the upper and lower bound of the 95% credibility value for each distribution based on its true correlation and SDₚ estimated; 95% C.I. = the upper and lower bound of the 95% confidence interval for each distribution; Q = homogeneity statistic.

*p < .05.

for each distribution. According to Whitener (1990), the credibility value has implications for whether the validity can be generalized. If this value is large and contains zero, it is likely that moderators exist. The confidence interval, on the other hand, is generated using the standard error and offers information about the accuracy of the effect size estimated from the meta-analysis. When this interval contains zero, researchers cannot conclude that the obtained correlation is significantly different from zero (Whitener, 1990).

Tables 2 and 3 also contain the statistics used to test the proposed moderated relationships. Column 8 provides the percentage of artifact variance accounted for, which is used to test for the presence of moderators following the Hunter & Schmidt (1990) 75% rule. Columns 9 and 10 provide the homogeneity Q statistic and its degrees of freedom, respectively. The Q statistic (Aguinis & Pierce, 1998; Hedges & Olkin, 1985) allows for a significance test of overall study-level effect size variability and formal testing for the presence of hypothesized moderator variables. These statistics are interpreted in the sections below.

As Table 2 indicates, there is no overall relationship between work group diversity and cohesion (ρ = 0.04). Further, although the relationship between less job-related diversity and cohesion is in the negative direction, while highly job-related diversity is in the positive direction (ρ = −0.03, ρ = 0.10, respectively), there is no relationship between either type of diversity and cohesion, and the confidence interval for these relationships includes zero. Therefore, we cannot conclude that either type of diversity impacts cohesion. The results do not support H1.

Second, the results show that diversity also has no overall relationship with performance (ρ = −0.01). Examining H2, the results indicate that although the correlations of less job-related and highly job-related diversity with performance are slightly different (ρ = −0.07, ρ = 0.02; respectively), we failed to find a relationship between either type of diversity and performance and the confidence interval for these relationships includes zero;
therefore, we cannot conclude that either type of diversity impacts performance. Thus, the results do not support H2.

Examining the moderating impact of team type on the relationship between diversity and performance, the results in Table 3 show that for less job-related diversity, TMTs using organizational performance as the outcome measure had no relationship with performance, and lower-level teams that used group-level outcomes also had no relationship with performance ($\rho = -0.07$, $\rho = -0.07$; respectively). For highly job-related diversity, TMTs had no relationship with performance, and lower-level teams also had no relationship with performance ($\rho = 0.03$, $\rho = -0.09$; respectively). These results do not support H3; therefore, we cannot conclude that the relationship between diversity and performance is stronger for lower-level teams than for TMTs.

5.1. Further examination of moderator analyses

It should be noted that with the exception of the correlation between highly job-related diversity and cohesion, more than 60% of the variance in the observed correlations across studies was explained by statistical artifacts (see Table 2 and 3). Pearlman, Schmidt, and Hunter (1980) have suggested that if 75% or more of the variance in the observed correlations across studies is explained by artifacts (sampling error, measurement unreliability, and range restriction), then the presence of a moderated relationship is unlikely. As noted by Mathieu and Zajac (1990), however, the 75% rule was proposed for meta-analyses that corrected for all three types of artifactual variance. In cases where it is only possible to correct for one or two sources of variance, as is the case in the present study, a lower percentage of artifact variance such as 50–60% across studies demonstrates that moderated relationships are unlikely. Further, examination of the credibility value shows a small range for all relations except the relationship between highly job-related diversity and cohesion, again demonstrating that moderators are unlikely. However, the credibility value does contain zero, thus we cannot be confident in concluding that moderators are not present (Whitener, 1990).

Further analysis of the moderator hypotheses were conducted utilizing the $Q$ statistic ($Q_B$ and $Q_w$) proposed by Hedges and Olkin (1985) and modified by Aguinis and Pierce (1998), (see Table 2 and 3). To test for the presence of moderators, we followed the three-step procedure proposed by Aguinis and Pierce to test moderator variable hypotheses meta-analytically. Their procedure is based on Hedges and Olkin's meta-analytic approach, but it incorporates study level corrections for methodological and statistical artifacts. Using this approach, we (1) corrected study-level effect size estimates for artifactual variance; (2) tested the overall homogeneity of study-level effect size estimates after the artifactual sources of variance have been removed; and (3) tested the effects of hypothesized moderator variables.

To test the overall variability of study-level effect sizes, we computed the homogeneity $Q$ statistic. A significant $Q$ indicates that study-level $d$s do not estimate a common population effect size, and thus a subsequent search for moderators is warranted (Aguinis & Pierce, 1998). To test our hypothesized moderator variables, we assessed the difference between mean within-subgroup effect sizes for each of the categories of our hypothesized moderator variable by computing a between-subgroup homogeneity statistic, $Q_B$, and we examined the
homogeneity of effect sizes for each of these subgroups by computing a within-subgroup homogeneity statistic $Q_w$ (see Aguinis & Pierce, 1998 for the formulas). According to Aguinis and Pierce, the presence of the predicted moderator is supported by a significant $Q_B$, which indicates a difference between the mean effect size estimates across subgroups, and a nonsignificant $Q_w$, which suggests that all the studies within each subgroup estimate a common population effect size.

The results of the hypothesized moderated analyses showed a nonsignificant $Q_B$ for the proposed moderator of job-relatedness. Specifically, with regard to the relationship between all types of diversity and cohesion with job-relatedness as the moderator, between-subgroup homogeneity was found, $Q_B(1) = 3.58, p > .05$. Similarly, with regard to the relationship between all types of diversity and performance with job-relatedness as the moderator, between-subgroup homogeneity was found, $Q_B(1) = 2.06, p > .05$. These findings suggest, according to Aguinis and Pierce (1998), that our hypothesized moderator does not moderate the relationship between diversity and cohesion or performance. Additional analysis of within-subgroup homogeneity was conducted by computing the $Q_w$ statistic (see Table 2). The results of this analysis show subgroup homogeneity for all relationships except the one between highly job-related diversity and cohesion. These findings are consistent with the results obtained using the Hunter and Schmidt (1990) 75% rule.

The results of the analysis of team type as a moderator also showed a nonsignificant between-subgroup $Q_B$. Specifically, between-group homogeneity was found for TMTs and lower-level teams with regard to the relationship between both less job-related diversity and performance, $Q_B(1) = 0.0207, p > .05$, and highly job-related diversity and performance, $Q_B(1) = 1.84, p > .05$. These results indicate that the type of team, and the corresponding level at which the team’s performance is measured, does not operate as a moderator of the relationship between either type of diversity and performance. Additional analysis of within-subgroup homogeneity was conducted by computing the $Q_w$ statistic (see Table 3). The results of this analysis show subgroup homogeneity for all relationships except the one between less job-related diversity and performance for lower-level teams. These findings are consistent with the results obtained using the Hunter and Schmidt (1990) 75% rule.

In summary, employing both the Hunter and Schmidt (1990) 75% rule and the Hedges and Olkin (1985) (modified by Aguinis & Pierce, 1998) $Q$ statistic, we find no support for our proposed moderators. Although both approaches produced consistent findings, the $Q_B$ statistic offered by Aguinis and Pierce provides a statistical test of between-subgroup homogeneity for the proposed moderator which is not possible using the Hunter and Schmidt 75% rule. Therefore, we recommend future research in the area of meta-analytic moderated analysis consider utilizing the $Q_B$ statistic.

6. Discussion

Diversity in work groups has been generalized as the “double-edged” sword (Milliken & Martins, 1996). Specifically, researchers have summarized the literature on diversity in work groups as demonstrating that increased diversity leads to less cohesion while increasing innovation and performance (Milliken & Martins, 1996; Jackson et al., 1995). However,
more recently, Pelled (1996) pointed out that to better understand the impact of diversity in work groups, researchers need to differentiate between the types of diversity. To accomplish this, Pelled and colleagues (Pelled, 1996; Pelled et al., 1999) proposed a theoretical distinction of two different types of diversity, highly job-related and less job-related diversity. Using meta-analytic methodology and theoretical arguments based on the prevailing models of diversity, we examined the “double-edged” sword of diversity by investigating the relationship between these two types of diversity and work group cohesion and performance across multiple studies.

The results of our meta-analysis demonstrate that the different types of diversity do not have differential impacts on work group cohesion and performance. Further, our results also show no relationship between both types of diversity and either group cohesion or performance. Thus, our research provides little evidence to sustain the conclusions drawn from recent literature reviews (e.g., Jackson et al., 1995; Milliken & Martins, 1996). Specifically, we found no evidence to support the idea that either type of diversity examined had much of an influence on group cohesion or performance.

Because our results support a null relationship and the confidence intervals include zero, we suggest interpreting these conclusions with caution. The impact of diversity in work groups is a growing area of research; therefore, it is yet to be thoroughly understood. This meta-analysis offers a preliminary empirical examination of a theoretical distinction based on job-relatedness. In the next section, we offer some—albeit not exhaustive—explanations for our lack of findings. Because many of these explanations relate to unmeasured variables or unexamined relationships in earlier diversity studies, they unfortunately could not be investigated in the present research.

6.1. Explanations

Because these results are inconsistent with the arguments proposed by a variety of researchers, possible explanations are warranted. One overarching possibility is that the impact of diversity on work group cohesion and performance has been overstated. Simply put, most research in this area has not discussed the magnitude of these relationships, which turn out to be quite small in size (see Table 1 for a list of correlations).

With regard to the lack of relationship between diversity and cohesion, a possible explanation relates to the moderating influence of time. More specifically, the work group's longevity may moderate the relationship between less job-related diversity and cohesion. Harrison et al. (1998) and Watson et al. (1993) have found that over time the negative affective outcomes associated with diversity decrease because people get to know each other and have a greater appreciation for and understanding of the differences in the group. Unfortunately, the studies examined in this research did not consistently report the tenure of the team; therefore, we were unable to include this variable in our analyses.

Another explanation for the lack of association between diversity and cohesion, recently highlighted in the work teams literature (e.g., Cohen & Bailey, 1997), is the impact of the organizational context. Specifically, one possible reason for the lack of relationship between diversity and cohesion is that aspects of the context, such as the organization's culture or
climate, may promote the valuing of diversity, which results in fewer problems in the area of cohesion.

Yet another explanation for our inability to find a relationship between diversity and cohesion may be a possible curvilinear relationship between these two variables. As argued by Lau and Murnighan (1998), the linear relationship examined by numerous researchers may not, in fact, be linear. Specifically, these authors have suggested that in moderately heterogeneously diverse work groups, subgroup formation is likely to occur resulting in a worse case scenario for the processes of the work team. And, in very high or very low diversity work groups, such subgrouping is less likely and therefore has less impact on group processes. Based on this argument, one possibility is that those groups with moderate diversity have poor cohesion, while groups with high and low diversity have better cohesion.

A final explanation relates to the measurement of cohesion. Some researchers have used the broader construct of social integration, of which cohesion is one element. Unfortunately, because of the small number of studies that examined either social integration or cohesion, we were unable to determine if the measure used impacted the relationships between either highly or less job-related diversity and cohesion.

To explain the lack of relationship between highly job-related diversity and performance, we suggest that just bringing together individuals with diverse backgrounds in terms of their task-related KSAs does not automatically lead to more complex problem-solving and effective group performance. In fact, we argue that bringing together these team members is the easy part, the next step is facilitating their ability to recognize different perspectives and integrate these into a more creative solution to a problem. Thus, to realize the potential value of this type of diversity on performance, researchers and practitioners should examine possible interventions including leadership and team member development that would foster cognitive elaboration and information exchange within work groups, drawing out the different knowledge and skills represented. This is especially important given the robust findings that group members do not always recognize when other members have relevant expertise (e.g., Littlepage et al., 1995; Littlepage et al., 1997), and they tend to discuss the information they have in common, leaving out important details not shared (e.g., Gigone & Hastic, 1993; Stasser & Titus, 1985).

6.2. Conclusions and future directions

We find the lack of a relationship of work group diversity with cohesion and performance surprising in light of the conclusions reached by researchers utilizing similar studies (e.g., Milliken & Martins, 1996; Jackson et al., 1995). In addition, these findings are surprising given the activities of practitioners who are forming diverse work teams with the expectation of producing more creative thinking and innovation (Jackson et al., 1995), as well as promoting and offering necessary resources and training to better “manage diversity” within organizations. Based on our findings, we propose that the vast amount of time and resources allocated to “managing” diversity be refocused on trying to better understand the effects of diversity on work groups. We do caution, however, that our findings should be interpreted as a first step in more fully understanding the complex relationship of diversity with key group processes, as well as with overall performance. But, what our results seem to suggest
is that the relationship of diversity with group processes and performance is less straightforward than researchers originally thought. Thus, these results point to the need for more sophistication in both the theories and methods used to study work group diversity. Specifically, we offer the following recommendations for future research in this area.

First, researchers should stop treating diversity as a generic concept by suggesting that all types of diversity will have a positive outcome on group performance. Instead, they should theoretically demonstrate that diversity in demographic attributes is linked to differences in underlying task-related KSAs; values, beliefs, and attitudes; and personality. As stated, the work by Pelled (1996; Pelled et al., 1999; Simons et al., 1999) which distinguishes demographic attributes in terms of job-relatedness is a much-needed move in this direction. We encourage researchers to continue to theoretically demonstrate the relationship between diversity attributes and purported underlying differences. Second, diversity researchers should stop treating team as a generic concept and should instead examine the type of team as a moderator of the relationship between diversity and important group outcomes. As we pointed out, expected membership differences in various types of teams may correspond to the degree of heterogeneity exhibited on highly and less job-related diversity attributes.

Third, researchers should move beyond “black box” studies (Lawrence, 1997) in which diversity is expected to have a direct effect on work group outcomes and intervening processes either go unmeasured or ignored. Instead, and consistent with theoretical frameworks of group performance (e.g., Hackman, 1987), researchers should examine the mediating and moderating role of group processes. The results of recent research supports the presence of such relationships (e.g., Pelled et al. 1999; Simons et al., 1999). Fourth, and related to the above, researchers should examine the potential moderating influence of time and the presence of a curvilinear relationship between diversity and performance, as suggested by recent authors (Harrison et al., 1998; Lau & Murnighan, 1998). Finally, researchers should investigate the impact of various training interventions to assist work teams in understanding and utilizing the different knowledge and skills represented in diverse work groups.

In conclusion, the results of this meta-analysis suggest that diversity in work groups is a complex phenomenon. We recommend that future investigations continue to try to understand the relationships between diversity and important group outcomes. Further, we propose that researchers and practitioners begin to refocus diversity issues to determine when, where, and how diversity makes a difference in terms of work group processes and performance.

Acknowledgments

The authors contributed equally to this paper. Portions of this article were accepted for presentation at the 1999 Society for Industrial and Organizational Psychology Conference, Atlanta, GA. The authors would like to thank Jose Cortina and three anonymous reviewers for their helpful comments on earlier versions of this manuscript.
References


