Group Decision and Negotiation **12**: 217–241, 2003

Cracks in Diversity Research: The Effects of Diversity Faultlines on Conflict and Performance

SHERRY M.B. THATCHER
MIS Department, Eller College of Business, 430 McClelland Hall, University of Arizona, Tucson, AZ 85721, USA (E-mail: sherryt@bpa.arizona.edu)

KAREN A. JEHN[1] AND ELAINE ZANUTTO[2]
[1]Management Department, [2]Statistics Department, 3000 Steinberg Hall-Dietrich Hall, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

In this quasi-field study, we investigate the effects of diversity faultlines on the conflict experience, performance, and morale of 79 groups. This is one of the first studies to operationalize the construct of diversity-related faultlines (Lau and Murnighan 1998). One of the most important contributions of this research is that *faultlines incorporate multiple characteristics of group members simultaneously* rather than assessing just one demographic characteristic at a time as most past diversity research has done. We develop a measure to capture the complexity of the faultline construct and to examine the effects of various group diversity faultline profiles on group outcomes. Linear results with a limited range of data show that faultlines are negatively related to conflict and positively related to morale and performance. Supplemental analyses that take into account the unique characteristics of our dataset indicate curvilinear relationships between diversity faultlines and relationship conflict, process conflict, group morale, and group performance. Groups with either virtually no faultlines (very diverse members) or strong faultlines (split into 2 fairly homogeneous subgroups) had higher levels of conflict and lower levels of morale and performance than groups with medium faultlines. The results suggest a more complex relationship between diversity and group process and outcome variables than typically described in diversity research. A detailed discussion of the faultline measure we developed and the methodological issues associated with measuring and interpreting faultlines are reported.

Key words: coalitions, conflict, diversity, faultlines, performance

1. Introduction

Diversity research has generally examined demographic characteristics in groups and attempted to relate this to various group outcomes (Tsui, Egan, and O'Reilly 1992), but the results linking group diversity, group processes, and performance have been mixed, at best. We provide here a brief review of two often used diversity constructs, heterogeneity and relational demography, and then propose a new construct for examining group composition and demographic profiles of groups - group faultlines.

Some past diversity studies have shown that group heterogeneity in tenure, educational background, functional background, and ethnicity can improve group performance (e.g., Bantel and Jackson 1989; Eisenhardt, Kahwajy, and Bourgeois 1997; Hambrick, Cho, and

Chen 1996). Other studies have shown that tenure, age, and ethnic heterogeneity lead to decreased levels of performance (Michel and Hambrick 1992; Zajac, Golden, and Shortell 1991), while still other studies have shown no relationship between performance and heterogeneity on the dimensions of age, sex, and ethnicity (O'Reilly, Williams, and Barsade 1997; O'Reilly, Snyder, and Boothe 1993; Wiersema and Bantel 1992). Tenure, age, sex, and ethnic heterogeneity have also been shown to lead to higher turnover, higher absenteeism, and lower levels of satisfaction (Ely 1995; Jackson et al. 1991; O'Reilly, Caldwell, and Barnett 1989; Pfeffer and O'Reilly 1987; Wharton and Baron 1987; Wiersema and Bird 1993). In addition to the inconsistency of this work, this past research has considered diversity as a group-level variable defined as the degree to which there is dispersion of a particular attribute in a specified population (Blau 1977), assessing the effects of only one demographic characteristic at a time.

Another method used to study diversity is to look at individual level differences and to examine how one group member is different from other members in the group. This research is typically referred to as relational demography or demographic dissimilarity. Researchers concerned with the effects of relational demography on process and outcome variables have also found inconsistent results and reached mixed conclusions. For example, Jackson et al. (1991) and Wagner, Pfeffer, and O'Reilly (1984) found that the more different individuals are in age from others in their workgroup, the more likely they will be to leave the group. However, Wiersema and Bird (1993) found that for Japanese employees, age dissimilarity had no effect at all on turnover. In addition, relational demography scores on other dimensions (tenure, generational similarity, gender and cultural similarity) had no effect on turnover (Jackson et al. 1991; Kirchmeyer 1995; Wagner et al. 1984). Tsui et al. (1992) found that race dissimilarity (the degree to which a member differed from others on race) decreased psychological commitment. On the other hand, Tsui et al. (1992) also found that psychological commitment was not affected by tenure, age, gender and education dissimilarity. Riordan and Shore (1997) also looked at race, gender and tenure similarity when compared to other group members. Consistent with Kirchmeyer (1995) and Tsui et al. (1992) they found that gender and tenure similarity did not influence group cohesiveness which has been found to lead to commitment (Allen and Meyer 1990). However, whereas Tsui et al. (1992) found that race dissimilarity decreased psychological commitment, Riordan and Shore (1997) found no effects of race similarity on group cohesiveness. Finally, Tsui et al. (1992) did not find any significant relationships between education-level similarity and organizational commitment, frequency of absences or intent to stay with the organization. Jackson et al. (1991) on the other hand, found that education-level similarity reduced individual-level turnover decisions. In sum, the results of research using the relational demography construct to examine individual-level differences within groups are also mixed across study, dependent variable examined, and characteristic(s) of similarity/dissimilarity investigated.

Like the group diversity (heterogeneity) literature, we propose that one downside of the relational demography literature is that it is been operationalized examining only one demographic characteristic at a time (Thatcher 1999). Thus, in addition to a complicated and inconsistent set of findings, neither the diversity nor the relational demography research has been able to explain how a *combination* of demographic characteristics influences a group or an individual within a group simultaneously. In sum, the study of both group diversity and relational demography have provided a number of inconclusive results and suggests that alternative ways of studying diversity in organizations is needed (Bacharach, Bamberger, and Mundell 1993; Frey 2000; Williams and O'Reilly 1998).

Lau and Murnighan's (1998) faultline concept, by drawing attention to the dispersion of multiple demographic characteristics among group members, may help diversity researchers untangle some of the seemingly contradictory findings relating diversity and group outcomes by focusing on member alignment across multiple characteristics rather than looking at one characteristic or one individual within a group at a time. The construct we present in this paper, diversity faultlines, incorporates multiple dimensions of individual differences in comparison to other group members *simultaneously*. In addition, we develop and present a measure that reflects the complexity of groups with multiple members measured on multiple demographic characteristics. The creation of a tool that measures multiple demographic dimensions simultaneously is one of the major contributions of this paper. Ideally, the use of theory based on member alignment, subgroup dynamics, and coalition formation within groups to predict conflict and group outcomes, and the empirical results of this study, will also further knowledge regarding group composition and diversity within groups.

2. Theory and Hypotheses

2.1. Defining Faultlines

Faultlines can be conceived of as hypothetical dividing lines that split a group into subgroups based on one or more attributes (Bezrukova, Thatcher, and Jehn 2001; Lau and Murnighan 1998). In other words, not only must the various attributes of a group be considered, but also the alignment of those attributes among the members, and the number of potentially homogeneous subgroups (Lau and Murnighan 1998). An example of a group with a strong faultline would be a group of four people that consists of two African-American men around the age of 50 and two Caucasian females around the age of 25. In this case, the faultline is strong because within the group are two homogenous subgroups based on their age, gender, and race. An example of a group with weak faultlines would be a group of four people consisting of one Asian female in her 20s, one Caucasian female in her 40s, one African-American man in his 40s, and one Caucasian man in his 20s. In this particular group, subgroups may form around the categories of gender or age but the subgroups would not be as strong as in the first example because any subgroup that forms is only similar on one dimension. Given that workgroups are often made up of more than four people and may form subgroups based on a variety and combination of demographic dimensions (e.g., ethnicity, age, race, gender, tenure, educational level, educational background, functional area), it is important to explore new constructs that can help us understand the effects of diversity characteristic distributions within groups. The concept of diversity faultlines is a first step in this direction.

In their seminal piece on faultlines, Lau and Murnighan (1998) use a comparison table to illustrate the differences between faultlines and traditional diversity conceptualizations (see Table 1). Four-person teams with a variety of demographic distributions are used to explain the differences. The table clearly shows that teams with very high diversity scores may have very weak faultlines illustrating that faultlines and diversity are different constructs. For example, Group #6 in Table 1 shows how a team consisting of a 20-year-old Native American female who is an unskilled worker, a 30-year-old White male supervisor, a 65-year-old Black female executive, and a 50-year-old Asian male machinist includes

Table 1. A Comparison of Four-Person Group Diversity Scores and Faultline Strength Across Four Demographic Characteristics: Race, Sex, Age, and Occupational Roles

Group #	Member A	Member B	Member C	Member D	Diversity[a]	Faultline Strength[b,c]
1	White Male 20 Sales	White Male 20 Sales	White Male 20 Sales	White Male 20 Sales	None	None
2	White Male 50 Plant Manager	White Male 55 Plant Manager	Black Female 31 Clerical Staff	Black Female 35 Clerical Staff	Low (4 diff; 2 levels; 0.23)	Very Strong (4 align, 1 way)
3	White Male 50 Plant Manager	White Female 31 Clerical Staff	Black Male 55 Clerical Staff	Black Female 35 Plant Manager	Low (4 diff; 2 levels; 0.23)	Medium (2 align, 1 way; 1 align, 2 ways)
4	White Male 60 Plant Manager	Asian Female 30 HR Manager	White Female 58 Plant Manager	Black Male 35 HR Manager	Moderate (4 diff; 2-3 levels; 0.29)	Strong (3 align; 1 way)
5	White Male 60 Plant Manager	Black Female 20 Plant Manager	Black Male 40 Secretary	Asian Female 30 Sales	High (4 diff; 2-3 levels; 0.39)	Medium (1 align; 3 ways)
6	Native American Female 20 Unskilled	White Male 30 Supervisor	Black Female 65 Executive	Asian Male 50 Machinist	Maximum (4 diff; 2-4 levels; 0.45)	Very Weak (1 align; 1 way)

(Adopted from Lau and Murnighan, 1998)

[a]Diversity is described in terms of three elements: (1) the total number of attributes in which at least two members are different (denoted as "diff"); (2) how these attributes might be organized into similar categories or levels; and (3) the variance of each attribute, as calculated using the O'Reilly, Caldwell, and Barnett (1989) group coefficient of variance (the standard deviation of the selected attribute divided by its mean).
[b]With the number of identified attributes fixed at four, faultline strength is determined here by the number of demographic attributes that align (denoted as "align") and the possible ways to subdivide the group on the basis of these attributes (denoted as "ways").
[c]We slightly altered Lau and Murnighan's (1998) classification of faultline strength to reflect the complexity found in these groups. We changed the "weak" classification of Lau and Murnighan's table to "medium" to clearly differentiate the subgroups who aligned on some characteristic other than sex and those who could only form subgrops based on sex (very weak groups). Groups would have "weak" faultlines if there were only 2 alignments, 1 way (example not shown in table).

members who are different on all demographic characteristics except sex. This high level of diversity contributes to a very weak faultline score as the only possible alignment across these four workers is on the category of sex. Group #2 in Table 1 shows the reverse effect. A team consisting of two 50-ish White male plant managers and two 30-ish Black female clerical staff is not very diverse but would have the potential for a *very strong* faultline score as the distribution of these four demographic characteristics creates two homogeneous subgroups. The only situation in which the diverstity and faultline scores are in agreement is when a team is completely homogeneous (Group #1; Table 1). See Table 1 for further group-level examples of the differences between the concepts of diversity and faultlines.

We base our theorizing regarding the concept of faultlines on self-categorization theory (Turner 1982), social identity theory (Tajfel 1978), the similarity-attraction paradigm (Byrne 1971), and coalition theory (Murnighan and Brass 1991). Self-categorization theory posits that individuals classify themselves and others into familiar categories in order to make predictions about subsequent interactions. These categorization groupings (e.g., female, engineer) are also used in defining an individual's social identity (Turner 1987). Thus, these first two mechanisms are highly related: social identity refers to an individual's knowledge that she belongs to certain social groups and has the emotions and values that are linked to that social group or category (self-categorization; Turner 1982). Subgroups (and therefore faultlines) can only exist when individuals identify with certain groupings and categorize themselves as belonging to particular groups.

Byrne's (1971) similarity-attraction paradigm suggests that people are attracted to others who are similar to themselves. This reasoning has led past researchers to suggest that demographic characteristics provide a means of determining similarity. Lincoln and Miller (1979) have shown that similarity between individuals may lead to more frequent communication and a desire to remain in the group. Byrne's (1971) theory suggests that individuals tend to apply negative assumptions to those with whom they are dissimilar. Possible effects of the reduction in attraction are that individuals withdraw from the organization or reduce their input (Cummings, Zhou, and Oldham 1993). These theories help explain how demographic characteristics are used to create subgroups – a key component of faultlines.

Although the faultline concept is based on some of the same theoretical foundations as the diversity literature, the measurement of faultlines and its conceptual emphasis on the creation of subgroups (or coalitions) within a group enables us to better understand how the context of the group influences group process and outcomes. Coalitions are subsets of members who often pool their resources and unite as a single voice (Murnighan and Brass 1991). Groups with very weak faultlines will have little to no subgroups either because the group is completely homogeneous or because it is as completely heterogenerous as a group can get (Groups #1 and 6, respectively, in Table 1). Groups with strong faultlines have two fairly homogeneous subgroups which may act as two distinct groups in the completion of a group task. These groups are also likely to have relatively low to moderate diversity scores (Groups #2 and 4, respectively, in Table 1).

In order to examine how diversity faultlines can help us understand the effects of diversity on groups we explore the relationship between diversity faultlines and three possible outcomes (conflict type, performance, and morale). First, the various types of conflict (task, relationship and process) have been widely studied in the diversity literature, and have been shown to partially mediate the relationship between group composition and group outcomes (Jehn, Chadwick, and Thatcher 1997; Jehn, Northcraft, and Neale, 1999; Pelled 1996b). Lau and Murnighan (1998) also view conflict as a potential outcome of strong faultlines. The relationship between diversity and performance has also been looked at extensively with mixed results therefore, we believe it is important to investigate the effects of diversity faultlines on performance. Finally, satisfaction (or morale) is an important component for ongoing teams. Many diversity and relational demography studies have investigated the effects of differences on satisfaction; we will extend this literature by examining the effect that diversity faultlines have on group morale.

2.2. Linking Faultlines to Conflict

Research in organizational conflict has focused on three dimensions of conflict, relationship, task, and process conflict (Amason 1996; Jehn 1995, 1997; Jehn et al. 1997; Pelled 1996a). Relationship conflicts are disagreements and incompatibilities among group members about personal issues that are not task related, such as social events, gossip, and world news. Task conflicts are disagreements among group members' ideas and opinions about the task being performed, such as disagreement regarding an organization's current strategic position or determining the correct data to include in a report. Process conflicts are disagreements about how a task should be accomplished (Jehn 1997). Although task and process conflicts tend to be related, Jehn (1997) delineated between task and process conflict based on findings of an ethnographic study of work groups. Weingart (1992) also found that process issues are seen as different than issues concerning tasks by workgroup members in that process issues concern making plans and task delegation proposals while task goals focus on the content or outcome of the task.

The similarity-attraction paradigm and literature from coalition theory suggest that individuals are likely to form coalitions when they have numerous similarities across a variety of demographic characteristics because they will tend to have pleasant interactions with each other (Byrne 1971; Pool 1976; Stevenson, Pearce, and Porter 1985). The similarity inherent in this type of coalition formation is likely to lead to less conflict within subgroups. However, coalitions tend to perpetuate the notion of in-groups and out-groups leading to increases of conflict *between* or across the subgroups (Hogg, Turner, and Davidson 1990).

Lau and Murnighan (1998) proposed in their theoretical model that demographic faultlines may lead to interpersonal conflicts as members break into subgroups. Diversity research has shown that groups with differences among members experience more relationship conflict than homogeneous groups (Jehn et al. 1997; Pelled 1996b). The conflict is often based on interpersonal biases and prejudices based on demographic differences and negative stereotypes of outgroup members (Abrams et al. 1990). A group where half the members are White, male accountants, for example, and the other half are Black, female engineers may have negative stereotypes of each other leading to interpersonal conflicts. Therefore, we propose that:

Hypothesis 1: The greater the faultline strength the more groups will experience relationship conflict.

Individuals with very different life experiences approach problems from different viewpoints and are often trained to identify and solve problems using methods particular to their functional or disciplinary background (Bantel and Jackson 1989; Pelled 1996). Pelled (1996) found that functional background and educational diversity were related to conflicts that focused on the task or content of ideas. Research suggests that debates regarding the group's specific task-related goals often stem from differences in educational or functional backgrounds and in previous work experience (Ancona and Caldwell 1992; Zenger and Lawrence 1989). Coalitions formed around age, sex, and race may also contribute to a common view of task-related goals. According to Lau and Murnighan (1998), subgroupings based on alignment of demographic characteristics may further increase discussion and conflict about the work task. Therefore, we expect that diversity faultlines will cause increased levels of task conflict.

Hypothesis 2: The greater the faultline strength the more groups will experience task conflict.

Diversity faultlines are also proposed to affect the amount of process conflict experienced in a group. Kabanoff (1991) suggests that contention about the rules which dictate the allocation of material interests drives distributive conflict. Subgroups with different work experiences have different views and may have different rules about how work should be done and who should be responsible for certain aspects of the work. The norms of the location where one grew up, for example, may affect the strategy one uses to approach a task. If a subset of group members come from countries where open and critical discussion about ideas is the norm (e.g., the United States, Canada) and another subset comes from countries where there is a history of consensus building (e.g., Japan, East Asia), we propose that they will experience high levels of process conflict since there are different approaches for fulfilling the task. This may be especially true if alignment on additional factors such as age and educational background exist leading to strong, opposing coaltions. That is, should these two subsets of group members also differ on their educational background (e.g., the sub-group from North America has business backgrounds and the sub-group from East Asia has engineering backgrounds) the level of process conflict will be exacerbated as individuals from different educational backgrounds also have different approaches to task completion. This leads us to hypothesize that:

Hypothesis 3: The greater the faultline strength the more groups will experience process conflict.

2.3. Linking Faultlines to Group Outcomes

It is also important to understand the relationship between diversity faultlines, group performance, and group morale. Objective performance is the quality of a group's work measured by objective standards (e.g., third-party ratings). Morale is the degree to which group members feel satisfied and committed about the group interactions.

For instance, Lau and Murnighan (1998) propose that faultlines will have a negative effect on group and individual member outcomes based on the potential conflict, communication problems, and distancing across the divide. This is similar to the findings in the coalition and group negotiation research. When coalitions form to divide a workgroup, they often enter into competition with one another for resources (Polzer, Mannix, and Neale 1998). The processes that members typically go through to negotiate common agreements will be hindered (Clark, Anand, and Roberson 2000) as communication and task interdependence is damaged. The competition that ensues disrupts the common group goal and decreases the overall morale of group members, as well (Murnighan 1978). Competition and adverse coalitions often cause negative attitudes toward the opposing side. The negative attitudes across subgroups will leave members feeling dissatisfied with the group process and experience and thus lower group morale. Therefore, we propose that faultlines will have a linear effect on group performance and morale:

Hypothesis 4: The greater the faultline strength the lower the objective performance.

Hypothesis 5: The greater the faultline strength the lower the level of morale.

3. Method

3.1. Sample and Procedure

This quasi-field study was conducted on 742 individuals in 144 teams of five or six participants, performing comparable tasks. Members were full-time MBA students at a large northeastern U.S. university. The average age of the members was 27, and 72% of the participants were male. Participants came from 55 different countries. Our final sample size was 79 groups since any group that did not have complete demographic information for all members was dropped from the sample. For a complete descriptions of the demographic characteristics of the team members, see Table 2.

Teams were randomly formed from the population of first-year MBA students to work on a variety of tasks, including group papers, group projects, and class assignments done as a group. The teams worked together an average of 15 hours a week performing various tasks over a 14-week period. A questionnaire was administered to the team members at the end of the period (at the end of the semester), but before they were given formal feedback on their performance. Students had the option of working with the same group of students the following semester.

3.2. Measures

Evenness of Group Size. Within our sample we had groups mainly consisting of five or six members. Two of the groups had only four members. It is conceivable that group size makes a considerable difference when looking at the effect of faultlines. For the groups with four or six members faultlines may create subgroups of equal sizes. Faultlines are unable to

Table 2. Demographic Descriptions of Team Members

Age
Average age: 27
Range: 21-39

Gender
Male: 315 (74%)
Female: 109 (26%)

Race

Caucasian:	267 (63%)
African-American:	20 (5%)
Asian:	95 (22%)
American Indian:	1 (0%)
Hispanic:	23 (5%)
Other:	18 (4%)

Country of origin
Members were asked to write out their country of origin and the responses were coded into the following 7 categories:

North America	308 (72%)
South America	16 (4%)
Asia	65 (15%)
Europe	31 (7%)
Australia	1 (0%)
Africa	2 (0%)
Middle East	3 (1%)

Graduate major
Individuals were asked to write down their current major. Their responses were coded into 10 categories:

Finance:	237 (56%)	Insurance:	28 (7%)
Management:	44 (10%)	Real Estate:	17 (4%)
Marketing:	51 (12%)	Public Policy:	4 (.1%)
Health Care:	13 (3%)	Accounting:	4 (.1%)
MIS:	15 (3.5%)	Other:	9 (2%)

Years of work experience
Individuals were asked to write down their years of work experience.
Average years of work experience: 4.7 years
Range: 4 months-16.8 years

Functional Area (Work Experience)
Individuals were provided with a list of functional areas and asked to mark the one that best represented his/her work experience. The eight options were:

Finance and Accounting:	158 (37%)
Manufacturing:	12 (2.8%)
Sales and Marketing:	40 (9.4%)
Computers and Engineering:	34 (8.0%)
Human Resources and Training:	17 (4.0%)
Consulting:	91 (21%)
General Management:	24 (5.7%)
Other:	48 (11%)

form equally– sized subgroups for teams with five members. The effects of this distribution are unknown and thus, we felt it necessary to control for the evenness of group size.

Faultlines. Faultlines were measured by combining demographic characteristics (years of work experience, type of functional background, degree major, sex, age, race, and country of origin). Faultline strengths were measured by calculating the percent of total variation in overall group characteristics accounted for by the strongest group split (see Appendix 1 for specific examples). We consider only faultlines that split groups into two subgroups for two reasons. First, because group sizes in this dataset were small (4–6 members) few groups were likely to have faultlines that split the group into more than two subgroups. Second, calculating *Fau* for fautlines that split groups into more than two groups would be much more computationally complex than the algorithm used here (see the Discussion section for more details).

To define this faultline measure, consider a group containing a total of n members who are measured on p characteristics. A faultline can split this group into two subgroups in a total of $S = 2^{n-1} - 1$ ways. For each of these possible subgroupings we measure the percent of total variation in overall group characteristics accounted for by the variation between subgroups by calculating the ratio of the between group sum of squares to the total sum of squares:

$$Fau_g = \left(\frac{\sum_{j=1}^{p} \sum_{k=1}^{2} n_k^g (\bar{x}_{\cdot jk} - \bar{x}_{\cdot j \cdot})^2}{\sum_{j=1}^{p} \sum_{k=1}^{2} \sum_{i=1}^{n_k^g} (\bar{x}_{ijk} - \bar{x}_{\cdot j \cdot})^2} \right) \quad g = 1, 2, \ldots S.$$

where x_{ijk} denotes the value of the j^{th} characteristic of the i^{th} member of subgroup k, $\bar{x}_{\cdot j\cdot}$ denotes the overall group mean of characteristic j, $\bar{x}_{\cdot jk}$ denotes the mean of characteristic j in subgroup k, and n_k^g denotes the number of members of the k^{th} subgroup (k = 1, 2) under split g (more details on how to incorporate categorical variables by using dummy variables in this calculation are described later in this section). The faultline strength, *Fau*, is then calculated as the maximum value of Fau_g over all possible splits g = 1,2, ... S. For small group sizes it is possible to calculate *Fau* by enumerating all possible splits and calculating the maximum Fau_g, g = 1, 2, ... S. This was the analysis approach taken in this paper. To apply this measure to larger groups, for which it is not practical to enumerate all possible splits, it may be necessary in future research to employ a clustering algorithm (e.g., Jobson 1992; Sharma 1996) to find the strongest split. *Fau* was calculated considering only group splits in which each subgroup has at least two members (e.g., not allowing subgroups of size one). The faultline scores used in this paper do not consider subgroups of size one because subgroups with only one person do not really constitute a group.

Fau can take on values between 0 and 1, with larger values indicating stronger faultlines. Groups that split into two relatively homogenous subgroups will have large values of *Fau*. Examples of two groups from our data are shown in Table 3.

Table 3. Example of *Fau* scores for two groups

Group A: *Fau* = 0.634

Member	Age	Gender	Race	Major	Work Experience (# years)	Work Function	Region Of Origin
1	27	Female	Caucasian	Marketing	4	Finance & Accounting	North America
2	29	Male	Caucasian	Finance	6	Finance & Accounting	North America
3	25	Male	Asian	Finance	4	Other	Asia
4	26	Female	Caucasian	Finance	4	Finance & Accounting	North America
5	29	Male	Asian	Management	6	Computers & Engineering	Asia

Group B: *Fau* = 0.282

Member	Age	Gender	Race	Major	Work Experience (# years)	Work Function	Region Of Origin
1	25	Male	Hispanic	Management	3	Sales & Marketing	North America
2	26	Male	Caucasian	Public Policy	2	Other	North America
3	27	Male	Asian	Management	3	HR & Training	Asia
4	27	Male	Caucasian	Finance	5	Sales & Marketing	North America
5	25	Male	African-American	Finance	3	Finance	North America
6	34	Male	Caucasian	Finance	11.5	Manufacturing	Europe

Group A has a relatively large *Fau* (*Fau* = 0.634) with the strongest split occurring between members 1, 2, 4 versus 3, 5. This splits the group into one subgroup of Asians and another group of North American Caucasians with work experience in finance and accounting. Group B has a relatively small *Fau* (*Fau* = 0.282) with the strongest split occurring between members 1, 2, 3, 5 versus 4, 6 (recall that subgroups of size one are not considered here). This splits the group into a subgroup of Caucasian Finance majors and another subgroup of heterogeneous people. Use of this *Fau* measure is motivated by its use in multivariate statistical cluster analysis (e.g., Jobson 1992; Sharma 1996) to solve a related problem of determining the number of clusters present in a population.

This *Fau* measure can be calculated using both continuous and categorical variables. However, since *Fau* is based on the Euclidean distance between people, care must be taken when using both continuous and categorical variables in this measure. For example, how do we combine a distance of 10 years in age with a difference in race on one common scale? Our approach is to use rescaling factors (Morrison 1967). The first step is to recode each categorical variable into a series of dummy variables. A categorical variable with c categories, such as race or degree major, should be re-expressed as c dummy variables (Jobson 1992). To understand why it should not be c-1 dummy variables, consider the following example. Suppose we have three categories for degree major: arts (A), social science (S), physical science (P), and we want to measure the distance between an arts major and a social science major. Using three dummy variables, the arts major is coded as A = 1, S = 0, P = 0, the social science major is coded as A = 0, S = 1, P = 0 and the Euclidean distance between them (ignoring all other characteristics) is $\sqrt{(1-0)^2 + (0-1)^2 + (0-0)^2} = \sqrt{2}$. The same distance would be obtained for any two people with different degree majors (e.g., people with different degree measures are a distance of $\sqrt{2}$ apart, and people with the same degree major are a distance of zero apart). Alternatively, consider using only two dummy variables, A and S. Then the arts major is coded as A = 1, S = 0, the social science major is coded as A = 0, S = 1 and the distance between them is $\sqrt{(1-0)^2 + (0-1)^2} = \sqrt{2}$. However, the distance between an arts major, A = 1, S = 0, and a physical science major, A = 0, S = 0, is $\sqrt{(1-0)^2 + (0-0)^2} = \sqrt{1} = 1$. Therefore, using only two dummy variables (or, more generally c-1 dummy variables) is unacceptable because the resulting distance does not measure all differences in degree major equally. Binary categorical variables ($c = 2$) are a special case in which it is possible to obtain consistent distance measures using only one dummy variable. However, for generality and to avoid complicating the discussion of rescaling factors below, we will assume that all categorical variables with c categories have been recoded into c dummy variables.

The second step is to rescale the continuous variables and the dummy variables so that they can be reasonably combined into one distance measure. The discussion above illustrates why it is convenient to rescale the dummy variables by dividing by $\sqrt{2}$ (or equivalently, by multiplying by $1/\sqrt{2}$) so that the dummy variables equal $1/\sqrt{2}$ (rather than 1) when a person falls into that category and equal zero otherwise. This will ensure that a difference in this categorical variable (e.g., a difference in degree major) will contribute a distance of 1 to the overall distance between these two people (rather than contributing a distance of $\sqrt{2}$ as in the original dummy variable coding above). This makes the rescaling factors for the continuous variables straightforward. For example, in the distance measure used for this analysis, a difference in race or sex or country of origin or functional background or degree major contributes as much to the distance between two people as a difference of 10 years in age or a difference of 8 years of experience. To accomplish this in the *Fau* measure described above, age needs to be rescaled by dividing all ages by 10 and experience needs to be rescaled by dividing all years of experience by 8.

This combination of recoding and rescaling allows categorical and continuous variables to be combined into a single distance measure that can then be used to measure faultline strength through *Fau*. In the absence of further information about the relative importance of differences in specific covariates to the strength of faultlines, these rescaling factors are unavoidably subjective (Morrison 1967). The rescaling factors used in this analysis seem reasonable, however, further research is needed into determining optimal rescaling factors.

Group size is not explicitly incorporated into the *Fau* measure. *Fau* measures the degree of aligment of covariates (e.g., the proportion of variability explained by a group split) which is independent of group size. It is possible that this aligment might have a bigger impact on group performance in some groups (e.g., in smaller groups or groups with an even number of people). We control for these effects by including "evenness of group size" as a separate predictor in our regression models.

Conflict. Relationship, task, and process conflict were measured by the intragroup conflict scale (Jehn 1995, 1997). This measure includes ten 7-point Likert-type questions anchored by 1 = "None" and 7 = "A lot." Relationship conflict was measured by questions such as, "How much emotional tension was there in your team?" and "How much were personality clashes evident in your team?" Questions such as "How much conflict of ideas was there in your team?" and "How much did you talk through disagreements about your team projects?" measured task conflict. Process conflict was measured by questions such as "How much disagreement was there about procedures in your team?" and "To what extent did you disagree about the way to do things in your team?" The Cronbach alpha scores for the scales of relationship, task, and process conflict were 0.92, 0.70, and 0.83, respectively. Compared to the Cronbach alphas for scales measuring task conflict reported in previous research, our score of .70 was quite low but acceptable (Amason 1999; Jehn 1995; Shah and Jehn 1997).

Group Performance. Objective performance was measured by an average of the final scores given to groups on their group projects in different classes over the observation period. The raters of these projects were the professors of the classes that the students were in. Since a different rater was employed for each class and each class had substantively different tasks, interrater reliability is not computed; however, all grades were standardized by class using z-scores.

Group Morale. Morale was measured by taking the group average of an individual's commitment to the team, satisfaction with the team, and intent to remain with the team. Commitment was measured with four 7-point Likert questions, anchored from 1 = "strongly disagree" to 7 = "strongly agree" from O'Reilly and Chatman's commitment questionnaire (1986). Some of the questions measuring commitment were, "I talk about my learning team to my friends as a great team to work with" and "I feel strongly committed to my learning team." Satisfaction was measured by the Kunin faces scale (1955). This 7-point Likert scale asks members to circle the face that most adequately portrays how they feel about working in the group. The faces range from a very sad face to a very happy face. Intent to remain with the team was a 1-item question asking the respondent the degree to which they intended to remain with their team in the next semester (7-point Likert scale). The Cronbach alpha for the morale scale was 0.93.

4. Results

Table 4 provides the means, standard deviations, and correlations of the variables in the model. As found in previous research, relationship, task, and process conflict were highly correlated and all three types of conflict were negatively related to group morale and performance. Contrary to our expectations, diversity faultlines were negatively correlated with conflict but positively correlated with morale and performance. We examine these relationships further using hierarchical regression analyses.

Evenness of Group Size. Evenness of group size as a control variable was not significant in any of the regression analyses that were run thus, this did not have any effect on the relationship between diversity faultlines, conflict experiences and group outcomes.

Diversity Faultlines and Conflict. See Table 5 for results of our tests of the hypotheses. The results of the regressions testing hypotheses 1 and 3 were significant but in the direction opposite that hypothesized. Thus, groups that had strong faultlines experienced less relationship and process conflict than groups with weak faultlines. Hypothesis 2 was not supported. Groups with stronger faultlines did not experience more task conflict.

Diversity Faultlines and Outcomes. Both hypotheses 4 and 5 were not supported. Strong faultlines did NOT significantly affect performance or morale.

5. Discussion of Initial Analyses

This study tests a model of the effects of diversity faultlines on conflict experiences and group outcomes. This is the first study we are aware of that empirically measures the theoretical construct of faultlines proposed by Lau and Murnighan (1998). Therefore, a main

Table 4. Means, Standard Deviations, and Intercorrelations (n = 79)

Variable	1	2	3	4	5	6	7
1. Diversity faultlines	–						
2. Relationship conflict	–.31**	–					
3.Task conflict	–.21	.56**	–				
4. Process conflict	–.25*	.81***	.66***	–			
5. Group morale	.16	–.58***	–.31*	–.57***	–		
6. Group performance	.17	–.64***	–.29**	–.66***	.79***	–	
7. Even group size	–.079	.069	.063	.099	–.077	–.095	–
Mean	.409	2.54	2.90	2.54	4.39	4.81	.392
Standard deviation	.076	1.23	.741	.726	.810	.903	.491

*p < .05, **p < .01, ***p < .001.

Table 5. Regression Analyses for Faultlines, Conflict, and Group Outcomes[a]

	Types of conflict			Group outcomes	
	Relationship (n = 79)	Task (n = 79)	Process (n = 79)	Performance (n = 79)	Morale (n = 79)
Step 1: Controls					
Even group size	.069	.063	.099	–.095	–.007
Adjusted R^2	–.008	–.009	–.003	–.004	–.013
F	.365	.304	.755	.694	.004
Step 2: Main effects					
Diversity faultlines	–.314**	–.208	–.248*	.166	.161
Adjusted R^2	.087	.031	.050	.015	.013
F change	8.430**	3.465	5.066*	2.170	2.036

[a]Entries represent standardized coefficients
*p < .05, **p < .01, ***p < .001.

contribution of this research, we believe, is the theoretical-based algorithm developed to assess the concept of faultlines, which up to this point has been entirely theoretical without operationalization capabilities.

Unfortunately, our results did not support our hypotheses or the predictions put forth by Lau and Murnighan (1998). We will discuss possible reasons for our findings and describe the supplemental analyses we conducted based on further examination of our data. There are two possible reasons why our findings did not support our hypotheses. First, the theoretical foundations of faultlines could be incorrect or incomplete. Second, there may be limitations in the data we used to test the hypotheses.

The faultline theory as developed by Lau and Murnighan (1998) suggests that more conflict will emerge as strong subgroupings form but we found in the case of relationship and process conflict that stronger subgroupings led to reduced perceptions of conflict. It is possible that immediate subgroups provide a comfortable and supportive atmosphere that lessens perceptions of conflict and increases performance and morale. In addition, when strong subgroups are formed there may be limited interaction across subgroups creating less opportunity for communication and conflict. The lack of interdependence across subgroups decreases actual conflict and perceptions thereof. This is similar to the results in studies of intragroup conflict which find lower levels of conflict in groups with low levels of interdependence than in highly interdependent groups (Jehn 1995). In addition to the (lack of) interdependence argument, Lau and Murnighan (1998) also suggested that groups that split into subgroups of comparable power may originally experience intense conflict but if resolved quickly would become less susceptible to future conflict. Therefore, theoretical considerations that were not tested may be responsible for the results of our analyses.

A second possibility, and one that we are able to further investigate with our data, is that there were limitations in our data that did not allow us to fully test the hypotheses. We found a restriction of range problem in our data. There were no groups in this study that were homogeneous and no groups that had complete homogenous subgroup splits, or very strong faultlines. The strength of the faultlines in our sample ranged from very weak to strong; they did not fit the full range of “none” to “very strong” as Table 1 based on Lau and Murnighan suggests are theoretically possible, and upon which our hypotheses were based.

6. Supplemental Analysis

Given the restriction of range problem in our data, we conducted post hoc supplemental analyses to look at the following, more appropriate, propositions based on our unique data. In the discussion above we suggested that lack of interdependence or the ability to resolve conflict quickly may reduce conflict, increase performance, and increase morale in groups with strong faultlines. However, both the work on group composition and boundary-spanning (Ancona and Caldwell 1992) suggest other alternatives. Ancona and Caldwell (1992) found that communication and group performance improve when there is overlapping membership across subgroups. Thus, groups in which there are more than one potential subgroup (medium faultlines) may be *less* likely to have conflict and *more* likely to have high levels of group performance and morale than in groups where there are no potential faultlines or in groups where there are strong faultlines. Belonging to multiple potential subgroups provides an opportunity for individuals to act as boundary spanners and thus facilitate communication among the multiple subgroups. Therefore, in the supplemental analyses, we examine whether groups with very weak (very diverse) or strong (two fairly distinct subgroups) of faultlines will experience high levels of relationship, task, and process conflict and groups with medium faultlines will experience low levels of relationship, task, and process conflict due to the boundary spanning possibilities. As regards morale and performance, we examine a similar curvilinear prediction that groups with very weak or strong (two fairly distinct subgroups) faultlines will have low levels of performance and morale whereas groups with medium faultlines will have high levels of performance and morale.

The results of the supplemental analyses can be seen in Table 6. The curvilinear regressions conducted in the supplemental analyses are a better fit of the data compared to the linear regressions conducted in the initial analyses as shown by the higher R^2’s and the significance of the coefficients.

Results: Diversity Faultlines and Conflict. We find a curvilinear effect of diversity faultlines on relationship and process conflict such that groups with very weak or strong faultline scores experience high levels of relationship and process conflict but groups with medium faultline scores experience low levels of relationship and process conflict. In other words, the curvilinear effects of diversity faultlines on process and relationship conflict is U-shaped. A curvilinear effect of diversity faultlines on task conflict was not supported.

Results: Diversity Faultlines and Outcomes**.** Groups with very weak or strong fautline scores had low levels of group performance and morale and groups with medium faultline scores had high levels of performance and morale. Thus, a curvilinear relationship in the

Table 6. Supplemental Regression Analyses for Faultlines, Conflict, and Group Outcomes[a]

	Types of conflict			Group outcomes	
	Relationship (n = 79)	Task (n = 79)	Process (n = 79)	Performance (n = 79)	Morale (n = 79)
Step 1: Controls					
Even group size	.069	.063	.099	–.095	–.007
Adjusted R^2	–.008	–.009	–.003	–.004	–.013
F	.365	.304	.755	.694	.004
Step 2: Main effects					
Diversity faultlines	–2.065*	–1.699	–2.062*	2.597**	2.434*
Diversity faultlines^2	1.763$^{\Psi}$	1.502	1.826$^{\Psi}$	–2.448*	2.289*
Adjusted R^2	.118	.049	.083	.084	.072
F change	6.206**	3.008$^{\Psi}$	4.530*	4.595*	4.040*

[a]Entries represent standardized coefficients
$^{\Psi}p < .06$, *$p < .05$, **$p < .01$, ***$p < .001$.

direction of an inverted U-shape exists between diversity faultlines and group outcomes.

6. Discussion of Results from Supplemental Analysis

The results of this study suggest that the distribution of demographic characteristics that team members possess can greatly influence conflict experiences and outcomes. Consistent with previous diversity research groups that were very diverse (very weak faultline scores) experienced high levels of process and relationship conflict. However, different from past diversity research, we consistently found that groups with strong faultlines (moderate diversity scores) also experienced high levels of relationship conflict and process conflict. These findings support the idea that the relationship between group composition and conflict may be nonlinear (Lau and Murnighan 1998). Having a moderate potential for developing faultlines and coalitions may decrease the potential for relationship and process conflict because individual team members who belong to more than one subgroup may move easily between subgroups smoothing over potential conflicts. As previous research on boundary spanning shows (Ancona and Caldwell 1992), the ability to move between subgroups improves communication between the subgroups.

We also found that diversity faultlines predict group morale and group performance in our supplemental analyses when we consider the limited range in our data (no truly homogenous groups and no groups with total alignment on all characteristics across subgroups which would be Lau and Murnighan's "very strong" faultline condition). The curvilinear relationship found between diversity faultlines and group outcomes was consistent with the relationship found between diversity faultlines and process and relationship conflict. Groups that were very heterogeneous across demographic characteristics or had two relatively homogeneous subgroups reported low levels of group morale and had low levels of performance. The groups with medium faultlines not only had less process and relationship conflict but also had higher levels of morale and group performance. Once again, team members that belong to more than one subgroup (the case in medium faultline groups) may create a communication link as they move between the various subgroups. Consistent with Thatcher's (1999) findings, diversity in and of itself is not negative; it is the composition and arrangement of the diversity among members that ultimately influence group process and performance. Future studies should examine not only the degree to which diversity exists in a team but the composition of that diversity. This faultline algorithm (*Fau*) is the first step toward addressing a major limitation of past diversity research: the ability to take into account a number of demographic characteristics simultaneously and incorporating the group composition.

8. Limitations and Future Research

One of the limitations of this quasi-field study is that the student sample limited variability on demographic differences and thus limited the range of the faultline construct. However, the restricted range of the sample provides for a conservative test of the hypotheses. In addition, we also feel that the restricted range of the sample is a realistic reflection of the American workforce. It would be unusual to find groups that are completely homogeneous on a range of demographic characteristics or fall nicely into two or three distinct subgroups. Despite this conservative test, there was support for the curvilinear model. Given that, there is a dire need for more research, especially field studies, in this area utilizing the new diversity construct of faultlines. Individuals in a school setting, regardless of the tasks and assignments they are told to do, are not exposed to the pressures and consequences of work in organizations. Conducting field work in this area is a difficult task, since issues of diversity and conflict are sensitive topics which many organizations may not wish to study for fear of upsetting their employees and incurring legal costs (Jehn 1997). Methodological triangulation using interviews, observations, and other-report measures (Taylor and Bogdan 1984) may help researchers untangle some of the effects of diversity faultlines on types of conflict and group performance. Again, we hope our study of student groups will be a useful first step in the development and use of the faultline algorithm.

For simplicity and due to the small group sizes in this data, we restricted our consideration to faultlines that split the group into two subgroups. Future research should examine the possibility of faultlines that split groups into more than two subgroups. Our faultline measure, *Fau*, is appropriate for quantifying differences between subgroups when the number of possible subgroups is fixed (e.g., in our case the number of possible subgroups is fixed at 2). However, when the number of possible subgroups is variable (e.g., a faultline may split some groups into 2 subgroups and other groups into 3 or 4 subgroups depending on group size and coalition formation) further criteria are necessary to identify the strongest split in a group (i.e., to decide whether the group splits into 2 or 3 subgroups). This is because *Fau* increases as the number of subgroups increases, complicating the choice between 2-subgroup splits and 3-subgroup splits, for example. The multivariate statistical cluster analysis literature (e.g., Jobson 1992; Sharma 1996) may suggest some solutions to this problem.

Our faultline strength measure, *Fau*, reflects the idea that faultlines become stronger as more attributes align themselves to form homogenous subgroups. However, it may also be important to measure not only how homogenous subgroups are, but how far apart the subgroups are. For example, a group with 3 members who are 27 years old each with 2 years experience and 3 members who are 55 years old each with 15 years of experience will have *Fau* = 1. On the other hand, a group with 3 members who are 27 years old each with 2 years experience and 3 members who are 31 years old each with 3 years experience will also have *Fau* = 1. Because both groups split into two perfectly homogenous groups, they both have extremely strong faultlines. However, the two subgroups in the first group are farther apart in terms of both age and years of experience than the two subgroups in the second group. Thus, it may be important to measure not only how "deep" a faultline is (e.g., the more attributes that align, the deeper the faultline), but also how "wide" it is (e.g., the faultline in the first group above is wider than in the second). Again the multivariate statistical cluster analysis literature suggests some measures for between-subgroup distance that may be used or adapted for this purpose. A measure that combines faultline depth and width may also be useful.

9. Implications

The implications of these findings for managers and group leaders are quite important. Using a more comprehensive and sophisticated construct and measurement of diversity we find important links to past research. Managers and group leaders need to learn how to identify and monitor groups that are likely to split into detrimental subgroups along faultlines. Strong heterogeneity and weak alignment among members (low faultline strength) does indeed cause more conflict and lower morale and performance. On the other hand, if the heterogeneity is structured so that multiple subgroups with overlapping members are created, then there may be little to fear from the diversity. This is especially important given the limitations of past diversity research which often examines only one demographic characteristic at a time. Although group leaders often do not have control over group diversity levels, our findings suggest they may be able to predict the type of conflict that is most likely to occur by identifying the diversity-motivated faultlines within the group. This may enable managers to more adequately supervise the group process. This could take the form of proactive training, which would teach team members and leaders the benefits and detriments of diversity, diversity composition and conflict within organizational groups.

In the interests of promoting diversity, organizations form teams that attempt to include many people who are different from one another, without taking into account the multitude of problems that diversity brings. The results of this study demonstrate that differences on a variety of demographic characteristics may be able to be managed if the distribution of team members does not contribute to strong faultlines, and if alignment does not contribute to negative stereotyping of outgroup members. This study should caution organizations engaged in haphazard development of diverse teams without provisions for training in group processes, diversity management, and conflict resolution. On the other hand, the results of this study suggest that certain levels and forms of diversity can indeed have very positive effects for teams.

10. Conclusion

In one of the first empirical research attempts to measure faultlines, we found that diversity faultlines play an important role in predicting relationship conflict, process conflict, morale and performance. The most interesting result of this research is that the effects of faultlines on groups are not linear. Groups with strong faultlines may experience negative effects as a result of the strong forces pitting the two sides against each other. Groups that experience moderate faultlines may experience turbulence more frequently but as the team members move between subgroups and become boundary spanners, they are able to ensure smooth communication and functioning. As the American workforce becomes more diverse (Johnston and Packer 1987), the findings in this study provide managers and organizations with a new tool for understanding group composition and diversity's effects on the workplace.

Appendix 1

Calculating *Fau*

As an example of how to calculate *Fau* we look at one hypothetical group with 4 members shown in the table below. In this example, the number of races is limited to 3, but in the actual data for this paper there were a total of 6 races.

Table A1: Raw Data

Member	Sex	Age	Race
1	Male	26	2
2	Male	31	1
3	Male	24	2
4	Female	29	3

This raw data is recoded to include two dummy variables for sec and 3 for race, so that the raw data matrix for this group is as shown in Table A2 (next page).

Table A2: Recoded Data

Member	Sex1	Sex2	Age	Race1	Race2	Race3
1	1	0	26	0	1	0
2	1	0	31	1	0	0
3	1	0	24	0	1	0
4	0	1	29	0	0	1

The raw data matrix is then rescaled. Here we rescale the categorical variables, sex and race, by dividing by $\sqrt{2}$ or (equivalently, multiplying by $1/\sqrt{2}$ which equals 0.707 approximately) so that a difference in sex or race contributes one to the total distance between two people, instead of contributing two as it would in the data in Table A2. We also rescale age by dividing by 10 so that a difference of 10 years in age contributes the same amount to the distance between two people as a difference in race or sex. The rescaled data is shown below in Table A3.

Table A3: Rescaled Data

Member	Sex1	Sex2	Age	Race1	Race2	Race3
1	0.707	0	2.600	0	0.707	0
2	0.707	0	3.100	0.707	0	0
3	0.707	0	2.400	0	0.707	0
4	0	0.707	2.900	0	0	0.707

Group Means (for rescaled data)					
Sex1 = $\bar{x}_{\bullet 1 \bullet}$	Sex2 = $\bar{x}_{\bullet 2 \bullet}$	Age = $\bar{x}_{\bullet 3 \bullet}$	Race1 = $\bar{x}_{\bullet 4 \bullet}$	Race2 = $\bar{x}_{\bullet 5 \bullet}$	Race3 = $\bar{x}_{\bullet 6 \bullet}$
0.530	0.177	2.750	0.177	0.354	0.177

Fau calculations for this group are shown in Table A4. Descriptions of the cells in Table A4 using the notation of this paper are given below.

Table A4

Raw Data

Member	Sex1	Sex2	Age	Race1	Race2	Race3
1	1	0	26	0	1	0
2	1	0	31	1	0	0
3	1	0	24	0	1	0
4	0	1	29	0	0	1

Rescaling Factors	Sex1	Sex2	Age	Race1	Race2	Race3
	0.707	0.707	0.100	0.707	0.707	0.707

Rescaled Data

Member	Sex1	Sex2	Age	Race1	Race2	Race3
1	0.707	0.000	2.600	0.000	0.707	0.000
2	0.707	0.000	3.100	0.707	0.000	0.000
3	0.707	0.000	2.400	0.000	0.707	0.000
4	0.000	0.707	2.900	0.000	0.000	0.707
Rescaled Means	0.530	0.177	2.750	0.177	0.354	0.177

Total Sum of Squares
(denominator of Fau-g)
2.289

			Subgroup Characteristic Averages						Between Group Sum of Squares for Characteristics						Subgroup	Total	
Split (g)	Members	ng	Sex1	Sex2	Age	Race1	Race2	Race3	Sex1	Sex2	Age	Race1	Race2	Race3	Between SS	Between SS	Fau-g
Split #1 (g=1)																	
Subgroup 1 (k=1)	1	1.000	0.707	0.000	2.600	0.000	0.707	0.000	0.031	0.031	0.023	0.031	0.125	0.031	0.272		
Subgroup 2 (k=2)	2,3,4	3.000	0.471	0.236	2.800	0.236	0.236	0.236	0.010	0.010	0.007	0.010	0.042	0.010	0.091	0.363	0.159
Split #2 (g=2)																	
Subgroup 1 (k=1)	2	1.000	0.707	0.000	3.100	0.707	0.000	0.000	0.031	0.031	0.123	0.281	0.125	0.031	0.622		
Subgroup 2 (k=2)	1,3,4	3.000	0.471	0.236	2.633	0.000	0.471	0.236	0.010	0.010	0.041	0.094	0.042	0.010	0.207	0.830	0.362
Split #3 (g=3)																	
Subgroup 1 (k=1)	3	1.000	0.707	0.000	2.400	0.000	0.707	0.000	0.031	0.031	0.123	0.031	0.125	0.031	0.372		
Subgroup 2 (k=2)	1,2,4	3.000	0.471	0.236	2.867	0.236	0.236	0.236	0.010	0.010	0.041	0.010	0.042	0.010	0.124	0.497	0.217
Split #4 (g=4)																	
Subgroup 1 (k=1)	4	1.000	0.000	0.707	2.900	0.000	0.000	0.707	0.281	0.281	0.023	0.031	0.125	0.281	1.022		
Subgroup 2 (k=2)	1,2,3	3.000	0.707	0.000	2.700	0.236	0.471	0.000	0.094	0.094	0.007	0.010	0.042	0.094	0.341	1.363	0.595
Split #5 (g=5)																	
Subgroup 1 (k=1)	1,2	2.000	0.707	0.000	2.850	0.354	0.354	0.000	0.062	0.062	0.020	0.062	0.000	0.062	0.270		
Subgroup 2 (k=2)	3,4	2.000	0.354	0.354	2.650	0.000	0.354	0.354	0.062	0.062	0.020	0.062	0.000	0.062	0.270	0.540	0.236
Split #6 (g=6)																	
Subgroup 1 (k=1)	1,3	2.000	0.707	0.000	2.500	0.000	0.707	0.000	0.062	0.062	0.125	0.062	0.250	0.062	0.625		
Subgroup 2 (k=2)	2,4	2.000	0.354	0.354	3.000	0.354	0.000	0.354	0.062	0.062	0.125	0.062	0.250	0.062	0.625	1.250	0.546
Split #7 (g=7)																	
Subgroup 1 (k=1)	1,4	2.000	0.354	0.354	2.750	0.000	0.354	0.354	0.062	0.062	0.000	0.062	0.000	0.062	0.250		
Subgroup 2 (k=2)	2,3	2.000	0.707	0.000	2.750	0.354	0.354	0.000	0.062	0.062	0.000	0.062	0.000	0.062	0.250	0.500	0.218

Overall Fau= 0.595

Fau (not allowing subgruops of size one)= 0.546

Subgroup Characteristic Averages					
Sex1 = $\bar{x}_{\bullet 1k}$	Sex2 = $\bar{x}_{\bullet 2k}$	Age = $\bar{x}_{\bullet 3k}$	Race1 = $\bar{x}_{\bullet 4k}$	Race2 = $\bar{x}_{\bullet 5k}$	Race3 = $\bar{x}_{\bullet 6k}$

Between Group Sum of Squares for Characteristics					
Sex1 = $n_k^g(\bar{x}_{\bullet 1k}-\bar{x}_{\bullet 1\bullet})^2$	Sex2 = $n_k^g(\bar{x}_{\bullet 2k}-\bar{x}_{\bullet 2\bullet})^2$	Age = $n_k^g(\bar{x}_{\bullet 3k}-\bar{x}_{\bullet 3\bullet})^2$	Race1 = $n_k^g(\bar{x}_{\bullet 4k}-\bar{x}_{\bullet 4\bullet})^2$	Race2 = $n_k^g(\bar{x}_{\bullet 5k}-\bar{x}_{\bullet 5\bullet})^2$	Race3 = $n_k^g(\bar{x}_{\bullet 6k}-\bar{x}_{\bullet 6\bullet})^2$

$$\text{Subgroup Between SS} = \sum_{j=1}^{p=6} n_k^g \left(\bar{x}_{\bullet jk} - \bar{x}_{\bullet j\bullet}\right)^2$$

$$\text{Total Between SS} = \sum_{k=1}^{2}\sum_{j=1}^{p=6} n_k^g \left(\bar{x}_{\bullet jk} - \bar{x}_{\bullet j\bullet}\right)^2$$

$$\text{Total Sum of Squares} = \sum_{k=1}^{2}\sum_{j=1}^{p=6}\sum_{i=1}^{n_k^g} \left(x_{ijk} - \bar{x}_{\bullet j\bullet}\right)^2$$

$$Fau_g = \frac{\sum_{k=1}^{2}\sum_{j=1}^{p=6} n_k^g \left(\bar{x}_{\bullet jk} - \bar{x}_{\bullet j\bullet}\right)^2}{\sum_{k=1}^{2}\sum_{j=1}^{p=6}\sum_{i=1}^{n_k^g} \left(x_{ijk} - \bar{x}_{\bullet j\bullet}\right)^2}$$

$$Fau = \max_{g=1,2,\ldots 7}(Fau_g)$$

To calculate Fau based only on group splits in which each subgroup has more than one member, calculations are the same as described above except that the maximum is taken over all splits where each subgroup contains at least two members.

References

Abrams, D., M. Wetherell, S. Cochrane, M. A. Hogg, and J. C. Turner. (1990). "Knowing What to Think by Knowing Who You Are: Self-Categorization and the Nature of Norm Formation, Conformity and Group Polarization," *British Journal of Social Psychology* 29, 97–119.

Allen, N. J., and J. P. Meyer. (1990). "The Measurement and Antecedents of Affective, Continuance, and Normative Commitment to the Organization," *Journal of Occupational Psychology* 63, 1–18.

Amason, A. C. (1999). "The Effects of Past Performance on Top Management Team Conflict in Strategic Decision Making," *International Journal of Conflict Management* 10 (4), 340–360.

Amason, A. 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, D. G., and D. F. Caldwell. (1992). "Demography and Design: Predictors of New Product Team Performance," *Organization Science* 3, 321–341.

Bacharach, S. B., P. Bamberger, and B. Mundell. (1993). "Status Inconsistency in Organizations: From Social Hierarchy to Stress," *Journal of Organizational Behavior* 14, 21–36.

Bantel, K. A., and S. 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.

Bezrukova, K., S. M. B. Thatcher, and K. A. Jehn. (2001). "Comparing the Effects of Group Heterogeneity and Faultlines on Performance: An Empirical Assessment of Contrasting Models," Paper Presented at the Annual Meetings of the Academy of Management, Washington, D.C.

Blau, P. M. (1977). *Inequality and Composition: A Primitive Theory of Social Structure*. New York: Free Press.

Byrne, D. E. (1971). *The Attraction Paradigm*. New York: Academic Press.

Clark, M. A., V. Anand, and L. Roberson. (2000). "Resolving Meaning: Interpretation in Diverse Decision-Making Groups," *Group Dynamics: Theory, Research, and Practice* 4, 211–221.

Cummings, A., J. Zhou, and G. R. Oldham. (1993). "Demographic Differences and Employee Work Outcomes: Effects on Multiple Comparison Groups," Paper Presented at the Annual Meetings of the Academy of Management, Atlanta, Georgia.

Eisenhardt, K., J. Kahwajy, and L. Bourgeois. (1997). "Conflict and Strategic Choice: How Top Management Teams Disagree," *California Management Review* 39, 42–62.

Ely, R. J. (1995). "The Power in Demography: Women's Social Constructions of Gender Identity at Work," *Academy of Management Journal* 38 (3), 589–634.

Frey, L. R. (2000). "Diversifying Our Understanding of Diversity and Communication in Groups: Dialoguing with Clark, Anand, and Roberson," *Group Dynamics: Theory, Research, and Practice* 4, 222–229.

Hambrick, D., T. Cho, and M. Chen. (1996). "The Influence of Top Management Team Heterogeneity on Firms' Competitive Moves," *Administrative Science Quarterly* 41, 659–684.

Hogg, M. A., J. C. Turner, and B. Davidson. (1990). "Polarized Norms and Social Frames of Reference: A Test of the Self-Categorization Theory of Group Polarization," *Basic and Applied Social Psychology* 11: 77–100.

Jackson, S. E., J. F. Brett, V. I. Sessa, D. M. Cooper, J. A. Julian, and K. Peyronnin. (1991). "Some Differences Make a Difference: Individual Dissimilarity and Group Heterogeneity as Correlates of Recruitment, Promotion, and Turnover," *Journal of Applied Psychology* 76, 675–689.

Jehn, K. (1995). "A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict," *Administrative Science Quarterly* 40, 256–282.

Jehn, K. (1997). "A Qualitative Analysis of Conflict Types and Dimensions in Organizational Groups," *Administrative Science Quarterly* 42, 530–557.

Jehn, K. A., C. Chadwick, and S. M. B. Thatcher. (1997). "To Agree or Not to Agree: The Effects of Value Congruence, Individual Demographic Dissimilarity, and Conflict on Workgroup Outcomes," *International Journal of Conflict Management* 8 (4), 287–306.

Jehn, K. A., G. B. Northcraft, and M. A. Neale. (1999). "Why Differences Make a Difference: A Field Study of Diversity, Conflict and Performance in Workgroups," *Administrative Science Quarterly* 44 (4), 741–763.

Jobson, J. D. (1992). *Applied Multivariate Data Analysis. Volume II: Categorical and Multivariate Methods*. New York: Springer-Verlag.

Johnston, W. B., and A. H. Packer. (1987). *Work Force 2000*. Indianapolis, IN: Hudson Institute.

Kabanoff, B. (1991). "Equity, Equality, Power, and Conflict," *Academy of Management Review* 16, 416–441.

Kirchmeyer, C. (1995). "Demographic Similarity to the Workgroup: A Longitudinal Study of Managers at the Early Career Stage," *Journal of Organizational Behavior* 16, 67–83.

Kunin, T. (1955). "The Construction of a New Type of Attitude Measure," *Personnel Psychology* 8, 65–78.

Lau, D. C., and J. K. Murnighan. (1998). "Demographic Diversity and Faultlines : The Compositional Dynamics of Organizational Groups," *Academy of Management Review* 23 (2), 325–340.

Lincoln, J. R., and J. Miller. (1979). "Work and Friendship Ties in Organizations: A Comparative Analysis of Relational Networks," *Administrative Science Quarterly* 24, 181–199.

Michel, J., and D. Hambrick. (1992). "Diversification Posture and the Characteristics of the Top Management Team," *Academy of Management Journal* 35, 9–37.

Morrison, D. G. (1967). "Measurement Problems in Cluster Analysis," *Management Science*, 13 775–780.

Murnighan, J. K. (1978). "Models of Coalition Behavior: Game Theoretic, Social Psychological, and Political Perspectives," *Psychological Bulletin* 85, 1130–1153.

Murnighan, J. K., and D. J. Brass. (1991). "Intraorganizational Coalitions," *Research on Negotiation in Organizations* 3, 283–306.

O'Reilly, C. A., K. Williams, and S. Barsade. (1997). "Group Demography and Innovation: Does Diversity Help?" Paper Presented at the Annual Meetings of the Academy of Management, Boston, MA.

O'Reilly, C. A., R. Snyder, and J. Boothe. (1993). "Effects of Executive Team Demography on Organizational Change," *in* G. Huber, and W. Glick (eds.), *Organizational Change and Redesign*. New York: Oxford University Press, pp. 147–175.

O'Reilly, C. A., D. F. Caldwell, and W. P. Barnett. (1989). "Work Group Demography, Social Integration, and Turnover," *Administrative Science Quarterly* 34, 21–37.

O'Reilly, C. A., and J. Chatman. (1986). "Organizational Commitment and Psychological Attachment: The Effects of Compliance, Identification, and Internalization of Prosocial Behavior," *Journal of Applied Psychology* 71, 492–499.

Pelled, L. H. (1996a). "Relational Demography and Perceptions of Group Conflict and Performance: A Field Investigation," *International Journal of Conflict Resolution* 7 (3), 230–247.

Pelled, L. H. (1996b). "Demographic Diversity, Conflict, and Work Group Outcomes: An Intervening Process Theory," *Organization Science* 7 (6), 615–631.

Pfeffer, J., and C. A. O'Reilly. (1987). "Hospital Demography and Turnover Among Nurses," *Industrial Relations* 26, 158–173.

Polzer, J. T., E. A. Mannix, and M. A. Neale. (1998). "Interest Alignment and Coalitions in Multiparty Negotiation," *Academy of Management Journal* 41, 42–54.

Pool, J. (1976). "Coalition Formation in Small Groups With Incomplete Communication Networks," *Journal of Personality and Social Psychology* 34 (1): 82–91.

Riordan, C. M., and L. M. Shore. (1997). "Demographic Diversity and Employee Attitudes: An Empirical Examination of Relational Demography Within Work Units," *Journal of Applied Psychology* 82 (3), 342–358.

Shah, P. P., and K. A. Jehn. (1993). "Do Friends Perform Better than Acquaintances: The Interaction of Friendship, Conflict, and Task," *Group Decision and Negotiation* 2: 149–166.

Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Wiley & Sons.

Stevenson, W. B., J. L. Pearce, and L. W. Porter. (1985). "The Concept of 'Coalition' in Organization Theory and Research," *Academy of Management Review* 10 (2): 256–268.

Tajfel, H. (1978). "The Achievement of Group Differentiation," *in* H. Tajfel (ed.), *Differentiation Between Social Groups.* London: Academic Press, pp. 483–507.

Taylor, S., and R. Bogdan. (1984). *Introduction to Qualitative Research Methods: The Search for Meanings.* New York, NY: John Wiley & Sons.

Thatcher, S. M. B. (1999). "The Contextual Importance of Diversity: The Impact of Relational Demography and Group Diversity on Individual Performance and Satisfaction," *Performance Improvement Quarterly* 12 (1), 97–112.

Tsui, A. S., T. Egan, and C. A. O'Reilly. (1992). "Being Different: Relational Demography and Organizational Attachment," *Administrative Science Quarterly* 37, 549–579.

Turner, J. C. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, England: Basil Blackwell.

Turner, J. C. (1982). "Towards a Cognitive Redefinition of the Social Group," *in* H. Tajfel (ed.), *Social Identity and Intergroup Relations*. Cambridge, England: Cambridge University Press, pp. 15–40.

Wagner, G. W., J. Pfeffer, and C. A. O'Reilly. (1984). "Organizational Demography and Turnover in Top Management Groups," *Administrative Science Quarterly* 29, 74–92.

Weingart, L. R. (1992). "Impact of Group Goals, Task Component Complexity, Effort and Planning On Group Performance," *Journal of Applied Psychology* 77, 682–693.

Wharton, A., and J. Baron. (1987). "So Happy Together? The Impact of Gender Segregation on Men at Work," *American Sociological Review* 52, 574–587.

Wiersema, M., and A. Bird. (1993). "Organizational Demography in Japanese Firms: Group Heterogeneity, Individual Dissimilarity and Top Management Team Turnover," *Academy of Management Journal* 36, 996–1025.

Wiersema, M., and K. Bantel. (1992). "Top Management Team Demography and Corporate Strategic Change," *Academy of Management Journal* 35, 91–121.

Williams, K. Y., and C. A. O'Reilly. (1998). "Demography and Diversity in Organizations: A Review of 40 Years of Research," *Research in Organizational Behavior* 20, 77–140.

Zajac, E., B. Golden, and S. Shortell. (1991). "New Organizational Forms for Enhancing Innovation: The Case of Internal Corporate Joint Ventures," *Management Science* 37: 170–84.

Zenger, T., and B. Lawrence. (1989). "Organizational Demography: The Differential Effects of Age and Tenure Distributions on Technical Communications," *Academy of Management Journal* 32, 353–376.