Unlocking the Effects of Gender Faultlines on Team Creativity:
Is Activation the Key?

Matthew J. Pearsall, Aleksander P. J. Ellis, and Joel M. Evans
The University of Arizona

The purpose of this study was to use faultline theory to examine the effects of gender diversity on team creativity. Results from 80 teams working on an idea generation task indicated that the activation of gender faultlines negatively affected the number and overall creativity of ideas. However, gender faultlines that were not activated had no effect. Results also indicated that the relationship between activated gender faultlines and team creativity was partially mediated by the level of conflict within the team. Specifically, emotional conflict partially mediated the effects of activated gender faultlines on the number of ideas generated. Implications are discussed, as well as possible limitations and directions for future research.

Keywords: faultlines, teams, creativity, conflict

In an increasingly competitive global marketplace, organizations must exploit their ability to nurture innovation in order to survive and prosper (e.g., Danneels, 2002; Pi & Cohen, 2006; Subramaniam & Youndt, 2005). As noted by Porter (1990), “prosperity is created, not inherited . . . competitiveness depends on the capacity of industry to innovate and upgrade” (p. 73). Organizational researchers have consistently linked a firm's innovative output to higher profitability, product quality, and market value (e.g., Cho & Pucik, 2005). Innovation, in turn, depends on the generation of creative ideas by employees (Amabile, 1983; Thompson & Brajkovich, 2003; Van de Ven, 1986).

Given its practical importance, researchers have attempted to identify methods to maximize the creative potential of the workforce. Since the early work of Osborn (1957), researchers have identified team-based work structures as a potential key to unlocking employee creativity. Teams allow access to a heterogeneous social circle that can capitalize on the varied skills and expertise of its members (e.g., Mohrman, Cohen, & Mohrman, 1995; Taggar, 2002; Tesluk, Farr, & Klein, 1997). Exposure to different backgrounds, approaches, and perspectives is thought to stimulate crucial processes such as divergent and flexible thinking (Cosier, 1975; Granovetter, 1982). Researchers have suggested that communicating and interacting with diverse others may enhance creativity by enabling new pathways of thought and preventing groupthink (e.g., Amabile, 1994; De Dreu & West, 2001; Watson, Kumar, & Michaelson, 1993).

However, a number of empirical studies have been much less positive, as researchers have found that increased demographic diversity can actually hamper creativity and innovation (e.g., Ancona & Caldwell, 1992; O’Reilly & Platt, 1989) by inducing internal friction (e.g., Argote & McGrath, 1993; Maznevski, 1994; Pelled, Eisenhardt, & Xin, 1999) and negatively influencing the exchange of creative ideas (e.g., Bhappu, Griffith, & Northcraft, 1997; van Knippenberg, De Dreu, & Homan, 2004).

This inconsistency has led researchers to suggest that simply examining the amount of heterogeneity does not adequately capture this phenomenon and have called for the development of more complex theoretical conceptualizations of diversity (Harrison & Klein, 2007; van Knippenberg & Schippers, 2007). One theory that may help to clarify the nature of the relationship between diversity and team creativity is faultline theory (Lau & Murnighan, 1998, 2005; Li & Hambrick, 2005; Thatcher, Jehn, & Zanutto, 2003). Faultlines are defined as “hypothetical dividing lines that may split a group into subgroups based on one or more attributes” (Lau & Murnighan, 1998, p. 328). They arise from variation in a number of team member attributes, including age, ethnicity, and gender. Lau and Murnighan (1998) argued that when team members share a characteristic with one or more of their teammates, they will tend to align themselves on the basis of that attribute. This alignment can create fractures within the team that have the potential to inhibit essential processes such as communication, coordination, cohesion, and trust. However, in order for such fractures to exert an influence on team outcomes, faultlines must be activated by features of the team task or context that stimulate the recognition of differences and similarities within the team.

Although faultlines can arise from differences across a number of dimensions, this study focuses on demographic diversity—more specifically, gender diversity. Given the increasing representation of women in management positions and work groups (see Duehr & Bono, 2006; Powell & Graves, 2003), gender embodies one of the more salient components of diversity in teams. Therefore, it is important to determine the role, if any, that gender diversity plays in the creative process of teams.

The purpose of this study is twofold. First, we use faultline theory to examine the effects of gender diversity on team creativity and expect that activated gender faultlines will negatively affect team creativity. Second, we attempt to uncover the mechanism...
through which activated gender faultlines exert their influence. More specifically, we argue that the relationship between activated gender faultlines and creativity is determined, in part, by the level of conflict within the team.

This study contributes to the literature in a number of ways. Specifically, we hope to further our understanding of the effects of gender diversity on team creativity, looking beyond simple dispersion to examine the influence of subgroup formation. We also explore the critical role of activation in faultline theory and uncover the mechanisms through which activated gender faultlines affect creativity. Finally, by theoretically bounding our hypotheses, we hope to offer advice regarding when and how managers can arrange teams to avoid the potential pitfalls associated with combining men and women in the same group when creativity is the goal.

Diversity in Teams

Research on demographic diversity in groups has traditionally taken a social categorization perspective, focusing on how dissimilarities among members lead to their conception of others as different from themselves (e.g., Jackson, 1992; Williams & O’Reilly, 1998). The social categorization perspective suggests that the classification of group members by their demographic characteristics will disrupt group processes and negatively affect attitudes and performance outcomes (e.g., Brewer, 1979; Jehn, 1995; Murnighan & Conlon, 1991; Tajfel & Turner, 1986). However, support for the hypothesized effects of demographic diversity on group outcomes has been inconsistent (e.g., Bowers, Pharmer, & Salas, 2000; Jehn, Northcraft, & Neale, 1999; Pelled et al., 1999).

Meta-analytic reviews have suggested that the traditional reliance on simple measures of demographic dispersion may be preventing researchers from adequately accounting for the effects of group heterogeneity (e.g., Bowers et al., 2000; Webber & Donahue, 2001). Instead, researchers have called for the choice of more sophisticated conceptualizations of diversity that go beyond dispersion (Harrison & Klein, 2007) and that “pay greater attention to the processes mediating the effects of diversity” (van Knippenberg & Schippers, 2007, p. 517). Recently, Lau and Murnighan (1998) proposed that researchers might better account for the effects of diversity on team processes and performance by examining the relational demography of subgroups.

Faultline Theory

Lau and Murnighan (1998) introduced their theory of faultlines to explain the effects of demographic subgroup formation within a larger work group. Lau and Murnighan suggested that group faultlines depend on the “compositional dynamics of the multiple demographic attributes that can potentially subdivide a group” (Lau & Murnighan, 1998, p. 325). Demographic faultlines can form along a number of perceivable dimensions, such as gender, ethnicity, age, managerial status, education, tenure, or functional area. A faultline can be said to exist within a group or team if two or more subgroups can form along one or more demographic lines. For example, in a team with two men and two women, a single gender faultline would exist between those two subgroups. This faultline may lead to social categorizations between team members, resulting in impaired communication, coordination, and cohesion.

On the basis of theories of cross-categorization (e.g., Crisp, Hewstone, & Rubin, 2001; Deschamps & Doise, 1978; Tajfel & Turner, 1979), Lau and Murnighan (1998) argued that when multiple categories align, the faultline is considered stronger as the differences between subgroups become more readily available for comparison. In contrast, if categories are crossed such that members can be classified in multiple, overlapping demographic subgroups, the faultline is said to be weaker.

A number of studies have provided empirical support for this theory. Li and Hambrick (2005) found that within teams with strong faultlines members evaluated their own subgroups more positively, they also found that such teams as a whole had lower levels of conflict, greater psychological safety, and higher satisfaction than did teams with weak faultlines. Similarly, Thacher, Jehn, and Zanutto (2003) found that the magnitude of demographic faultlines was negatively related to group conflict, and Gibson and Vermeulen (2003) found that the formation of subgroups resulted in greater learning behavior than in teams without subgroups.

One explanation for the problems between linking faultlines and outcomes is the way the construct has been operationalized (van Knippenberg & Schippers, 2007). Although many faultline researchers have relied on indices of group differences to measure faultline potential (e.g., Gibson & Vermeulen, 2003; Thacher et al., 2003), such measures presume that some or all of the possible demographic differences will be relevant to team members and lead to disruptive subgroup formation. Other researchers have echoed this concern, suggesting that social categorizations resulting from faultlines may be contingent on the salience of the group task or context to the faultline in question (Gibson & Vermeulen, 2003; van Knippenberg et al., 2004).

Faultline Activation

Lau and Murnighan (1998) referred to the process of triggering social categorization based on the salience of demographic differences within a group as activation. The salience of social categories is dependent on the comparative fit, normative fit, and cognitive accessibility of the categorization (Oakes, Haslam, & Turner, 1994; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Comparative fit refers to the degree to which the categorization reflects the actual differences between members, normative fit refers to the extent to which the categorization is meaningful to members within the group’s context, and cognitive accessibility relates to how easily members perceive the differences between them and how quickly they come to mind. Social categorization theory suggests that all three components must be present for subgroups to form (Turner et al., 1987). For example, the social categorization of a work team along gender lines meets the requirement of
cognitive accessibility, as gender differences are readily apparent, and will have strong comparative fit if there are no other cross-cutting differences between team members. To achieve normative fit, however, a gender categorization also requires that the team task or context makes gender differences meaningful to team members.

Building on this theory, Lau and Murnighan (1998) argued that for demographic subgroups to form, the group task or context must contain faultline-relevant elements that activate the salience of similarities and differences. Lau and Murnighan (1998) offered examples of this process, suggesting that the activation of ethnic faultlines might result from discussing affirmative action or that those based on occupational roles might be triggered by a team resource-allocation decision. If, however, faultlines are not activated, they will remain dormant, and the group will continue functioning as usual.

Though the role of faultline activation was critical in Lau and Murnighan’s (1998) theory, researchers have yet to explicitly examine the effects of faultline activation in teams. Although Lau and Murnighan (2005) recently studied faultlines in teams that performed tasks without obvious faultline-relevant issues, they suggested that such groups had received only “minimal faultline stimulation” (p. 655) and that the activation of group faultlines would have resulted in significantly different effects. Because the presence of multiple crosscutting faultlines makes it unclear along which differences subgroups will form, the activation of faultlines is most effectively studied within a single-issue heterogeneity characteristic, such as gender (Lau & Murnighan, 2005).

Therefore, in an attempt to examine the role of demographic diversity in team creativity, we focus on activating gender faultlines in teams. Research examining the effects of gender diversity on team interactions has been mixed. Some researchers have suggested that gender-heterogeneous teams are able to take advantage of their range of mental capacities and perspectives to develop new ideas and solve complex problems (e.g., Frink, Robinson, & Reithel 2003; Hirschfeld, Jordan, Feild, Giles, & Armenakis, 2005), whereas others have identified gender differences as a source of friction (e.g., Jehn, 1995; Pelled et al., 1999; Randel, 2002), negatively influencing the exchange of creative ideas. Applying faultline theory to gender provides one avenue toward understanding the conditions under which the gender demography of a team influences its ability to be creative.

Team creativity requires building on, combining, and critically improving each member’s ideas through open interaction (Amabile, Conti, Coon, Lazenby, & Herron, 1996). Creativity is further stimulated through challenging majority opinions, forcing members to rethink old beliefs and solve problems more imaginatively (Nemeth, 1986). Such divergent thinking emerges from the participation of all team members and the free flow of ideas between them (De Dreu & West, 2001; Nemeth & Staw, 1989). When activated, we expect gender faultlines to cause inter-subgroup communication to break down and member participation to wane, preventing the sharing of unique experiences and viewpoints. The potential benefits of diversity will be stifled, leading to fewer, less creative ideas.

Therefore we hypothesize the following:

**Hypothesis 1:** Activated gender faultlines will negatively affect team creativity.

**The Role of Emotional Conflict**

We expect activated gender faultlines to negatively affect team creativity by influencing the level of conflict within the team. Researchers studying the effects of demographic diversity have suggested that differences in impermeable demographic characteristics such as gender are more likely to lead to emotional conflict between members (Jehn et al., 1999; Randel, 2002). Emotional conflict, sometimes called relationship conflict (Jehn, 1995), is affective in nature and refers to the interpersonal disputes between team members that result in anger, resentment, and mistrust (Pelled, 1996; Jehn et al., 1999). Emotional conflict distracts members from the pursuit of team goals, making them reluctant to share or listen to opposing viewpoints (e.g., Evan, 1965; Jehn, 1997) and to waste time in contentious, non-task-related activities (Jehn & Mannix, 2001). Emotional conflict is particularly detrimental to team creativity, as it encourages cynical attributions regarding the ideas and contributions of team members, thwarting productive discussion (Janssen, Van de Vliert, & Veenstra, 1999; Torrance, 1957). Research has supported the negative influence of emotional conflict on team performance (e.g., Amason, 1996; Jehn & Mannix, 2001; Jehn et al., 1999), cognitive interaction, and innovation (e.g., De Dreu, 2006; De Dreu & Weingart, 2003).

Emotional conflict is likely to occur in teams with faultlines, as they contain implicit social categories along which subgroups may form. Once social categorization occurs, members become biased toward their subgroup and hold negative opinions about members of other subgroups, leading to disparagement of their ideas and interpersonal discord (Jehn, 1995; Tajfel, 1982; Tajfel & Turner, 1986). Although the presence of nonsalient gender differences do not stimulate emotional conflict (e.g., Lau & Murnighan, 2005; Li & Hambrick, 2005; Pelled et al., 1999), faultline theory suggests that gender faultlines will begin to cause antipathy between members when a faultline-related issue arises (Lau & Murnighan, 1998). Providing evidence, Randel (2002) found that the salience of gender differences within workgroups moderated the degree to which gender diversity became a source of emotional conflict.

Therefore, we expect activated gender faultlines to positively affect the team’s level of emotional conflict and emotional conflict to reduce team creativity. Given that we also expect that activated gender faultlines will negatively affect team creativity, we hypothesize the following:

**Hypothesis 2:** The effects of activated gender faultlines on team creativity will be partially mediated by emotional conflict.

**Method**

**Participants**

Participants included 320 students from introductory management courses at a large southwestern university who were arrayed into 80 four-person teams. Out of the 320 students, 160 (50%) were female, and 212 (66.3%) were White, with an average age of 21 years. In exchange for their participation, each participant earned class credit and became eligible for cash prizes ($200 per team) based upon the team’s performance. The experiment consisted of a 2 (gender faultline vs. no gender faultline) × 2 (activation vs. nonactivation) factorial design. Individuals were ran-
domly assigned to teams and randomly assigned to one of the four experimental conditions (faultline/activation, faultline/nonactivation, no-faultline/activation, and no-faultline/nonactivation), resulting in 20 teams per condition.

Procedure

Before they arrived, participants were randomly assigned online to either a gender-faultline or a no-gender-faultline team. Upon arrival, each team was placed in a private conference room for the duration of the experiment (approximately 1 hr). Teams initially encountered the activation manipulation, where they completed a 15-min task that was either gender biased (faultline/activation, no-faultline/activation) or gender neutral (faultline/nonactivation, no-faultline/nonactivation). Teams in all four conditions then completed the 15-min, gender-neutral experimental task. Upon conclusion of the experimental task, participants completed the manipulation check and conflict measures.

Manipulations

Gender faultlines. We manipulated gender faultlines by placing individuals in either gender-heterogeneous or gender-homogenous teams. The 40 gender-heterogeneous teams consisted of 2 men and 2 women, whereas the gender-homogenous teams included 20 teams composed of 4 men and 20 teams composed of 4 women. Because we included both all-male and all-female teams in the same condition, we wanted to ensure that the creativity performance of both groups was the same. Results indicated that there were no significant differences between the all-male and all-female teams on any of the outcomes of interest. Therefore, we collapsed both sets together into a single gender-homogenous group within this condition.

Activation. Lau and Murnighan (1998) argued that faultlines will lie dormant until activated by salient aspects of a task or context that act as a “trigger.” Therefore, faultline activation was manipulated by having teams complete a task that was either gender biased or gender neutral. We employed idea generation tasks, commonly used by researchers to stimulate and examine group creativity (e.g., Friedman & Forster, 2001; Goncalo & Staw, 2006). Team members were told to envision themselves as innovative product development teams and instructed to work together to generate new designs for specific products, focusing on distinguishing their ideas from those currently available on the market. For the faultline activation conditions, gender salience was triggered by instructing teams to generate new designs for a men’s electric razor. Teams were instructed that this razor was to be purchased and used by men and that their design ideas should appeal to the male market. In the nonactivation conditions, teams were instructed to generate new design ideas for a digital alarm clock: a gender-neutral product. Teams were instructed that this alarm clock was to be marketed to both male and female consumers.

Measures

Team creativity. After completing the activation session, team members were given instructions for completing the experimental task. For this task, teams were again told to envision themselves as innovative product development teams and were instructed to work together to generate original, practical ideas for a gender-neutral product: a student backpack. Teams were instructed that this backpack was to be marketed to all consumers. Teams were also told that their productivity would be rated according to two criteria: (a) the total number of creative ideas generated and (b) the originality or novelty of those ideas. Finally, they were informed that, based on those criteria, top-performing teams would be awarded a $200 cash prize to be shared evenly among team members ($50 per person).

We utilized two distinct but complementary evaluations of team creativity performance. First, we recorded the amount of idea generation, or the total number of nonrepetitious responses created by each team during a specific session (e.g., Diehl & Stroebe, 1987; Goncalo & Staw, 2006). Creativity researchers have noted the importance of examining the volume of ideas in determining a group’s ability to develop creative solutions (e.g., Campbell, 1969; Simonton, 1999).

Second, we employed a variation of Amabile’s (1983) consensual assessment technique, which emphasizes subjective creativity ratings provided by independent judges (e.g., Shalley, 1991, 1995; Shalley & Perry-Smith, 2001; Zhou, 2003). Because many teams generated a large number of ideas, we undertook a more rigorous assessment of creativity by subjectively rating each item in a team’s list of ideas rather than making a broader overall creativity judgment (e.g., Friedman & Forster, 2001; Goncalo & Staw, 2006). These ratings were performed by two independent coders who were blind to the conditions. Creativity for each item was assessed on a 6-point scale from 1 (very uncreative: ideas are already standard product designs or else fail to provide any meaningful creative improvement to the product) to 6 (extremely creative: ideas are extremely unique and original). Two experimenters were in charge of coding. To ensure that the coding was accurate and consistent, both experimenters participated in a 2-hr training session, which included a review of the creativity literature as well as the definition of each creativity outcome utilized in this study. The experimenters then coded 10 practice teams, where any discrepancies were discussed and resolved between the coders. Following the training session, to establish interrater reliability, the two raters independently coded the same 30 teams (674 ideas). Their item-by-item subjective ratings demonstrated adequate reliability between raters (r = .79), so the scores for those items were averaged together, and the remaining 50 teams were split evenly between the raters.

In order to make overall creativity independent from the number of ideas generated, the creativity scores of all of a team’s ideas were averaged to form each team’s index of overall creativity (e.g., Goncalo & Staw, 2006; Shalley, 1991). The two measures (idea generation and overall creativity) therefore provided unique and independent assessments of team creativity (r = -.09, ns).

Emotional conflict. The emotional conflict scale included four items from Pellet et al. (1999) designed to assess the level of emotional tension within the team. Example items read “How much emotional conflict is there among members in your team?” and “How much tension is there among members in your team?” Responses were recorded on a 5-point Likert scale, ranging from 1 (none) to 5 (a lot). Coefficient alpha for this study was .82. To examine the extent to which the measure was sufficiently reliable to be aggregated at the team level, we computed intraclass correlation coefficients (ICC; see Klein et al., 2000). ICC(1) represents
the reliability of a single rating of the team mean, or the statistical agreement among team members regarding a rated variable. ICC(2) represents the reliability of the average of team member responses. An ICC(1) of .47 and an ICC(2) of .78 suggested that individual scores could be aggregated to the team level (see Bliese, 2000).

Results

Manipulation Check

To examine the potential of our faultline activation manipulation to trigger gender salience within the team, participants completed the manipulation check after both the activation session and experimental task. Each was asked to rate the degree of gender salience in the task using a four-item manipulation check (e.g., “This task seems to be biased toward men” and “The subject matter of this task seems more appropriate for men”). Participants responded to each item on a scale from 1 (not at all true) to 5 (very true). Coefficient alpha for the scale reached .90 for both the activation session and experimental task. Results indicated that the electric razor task was perceived as more gender salient (M = 3.38, SD = 0.80) than the alarm clock task (M = 1.64, SD = 0.74), t(319) = 20.23, p < .01, supporting the effectiveness of the activation manipulation.

The student backpack (M = 1.67, SD = 0.80) was also perceived as significantly less gender salient than the electric razor, t(158) = 19.58, p < .01, but was not significantly different than the alarm clock, t(158) = 0.73, ns, supporting the gender neutrality of the experimental task.

Tests of Hypotheses

Means, standard deviations, and intercorrelations among all the variables included in the hypotheses tests are included in Table 1.1 Hypothesis 1 predicted that activated gender faultlines would negatively affect team creativity. Given the design of the study, we tested Hypothesis 1 using hierarchical regression. In Step 1, we entered both the gender faultline and faultline activation manipulations as independent variables predicting our two creativity outcomes. In Step 2, we entered the interaction term. As shown in Table 2, neither the gender faultlines nor the faultline activation had any direct effects on either idea generation (β = .00, ns; β = .04, ns) or overall creativity (β = .09, ns; β = .03, ns). However, as shown in Table 2 and Figure 1, the interaction significantly affected idea generation (β = −.41, p < .05), explaining an additional 6% of its variance, and overall creativity (β = −.35, p < .05), explaining an additional 4% of its variance.

To test whether the interactions supported our hypothesis (i.e., that faultlines only exert their effects when activated), we first tested whether teams in the faultline/activation condition were significantly less creative than teams in the faultline/nonactivation condition by conducting a planned contrast. Results indicated that creativity outcomes for the faultline/activation condition were significantly lower than those for the faultline/nonactivation condition for idea generation, t(76) = −2.92, p < .01, and overall creativity, t(76) = −2.23, p < .05. Second, we tested whether teams in the other three conditions had similar levels of creativity through planned contrasts between the faultline/nonactivation condition and the two no-faultline conditions. Results indicated that creativity outcomes for the faultline/nonactivation condition were not significantly different than the no-faultline/activation [idea generation, t(76) = 0.31, ns; overall creativity, t(76) = −0.52, ns] or no-faultline/nonactivation conditions [idea generation, t(76) = −0.09, ns; overall creativity, t(76) = 0.41, ns]. In sum, these results support Hypothesis 1.

Hypothesis 2 proposed that effects of the activated gender faultlines on team creativity would be partially mediated by emotional conflict. Since we hypothesized that the effects of the interaction between activation and faultlines would be partially mediated by emotional conflict, we followed Baron and Kenny’s

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive statistics totals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.45</td>
<td>18.86</td>
<td>2.59</td>
</tr>
<tr>
<td>SD</td>
<td>0.35</td>
<td>8.71</td>
<td>0.37</td>
</tr>
<tr>
<td>Faultline/activation condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.60</td>
<td>13.05</td>
<td>2.43</td>
</tr>
<tr>
<td>SD</td>
<td>0.45</td>
<td>8.24</td>
<td>0.35</td>
</tr>
<tr>
<td>Faultline/nonactivation condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.38</td>
<td>20.60</td>
<td>2.65</td>
</tr>
<tr>
<td>SD</td>
<td>0.35</td>
<td>8.83</td>
<td>0.44</td>
</tr>
<tr>
<td>No-faultline/activation condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.40</td>
<td>20.60</td>
<td>2.62</td>
</tr>
<tr>
<td>SD</td>
<td>0.29</td>
<td>8.92</td>
<td>0.27</td>
</tr>
<tr>
<td>No-faultline/nonactivation condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.42</td>
<td>21.20</td>
<td>2.64</td>
</tr>
<tr>
<td>SD</td>
<td>0.29</td>
<td>6.52</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Note. N = 80 (n = 20 for each of the four conditions).

To help rule out the possibility that perceptions of gender bias in the activation task put women at a performance disadvantage, we compared gender-homogenous teams within the activation condition. In the experimental task we found no difference in the number of ideas generated by teams of all women (M = 21.7, SD = 10.73) and teams of all men (M = 19.5, SD = 7.11), F(1, 19) = 0.29, ns, or in the overall creativity of their ideas (women, M = 2.57, SD = 0.33; men, M = 2.67, SD = 0.22), F(1, 19) = 0.72, ns. These results suggest that completing the activation task did not result in reduced efficacy or engagement for women in the subsequent gender-neutral experimental task. To address the possibility that faultline effects were due to status differences stemming from the activation task, we analyzed individual team member contributions within the faultline/activation condition to verify there were no status-related differences between women and men. We asked participants to rate their own personal contribution to the team’s total creative output with the following item: “About ___% of my team’s ideas came from me,” with responses ranging from 1 (0%–20%) to 5 (81%–100%). Within the faultline/activation condition, we found no differences between women’s (M = 2.50, SD = 1.01) and men’s (M = 2.43, SD = 1.06), F(1, 79) = 0.05, ns, perceptions of their own contributions. These results suggest that women were as equally engaged as men during the experimental task and viewed their contributions as equally significant.
procedure for testing mediated moderation, which requires that regression equations involving the interaction must also include the main effects of the interaction term. As noted above, the interaction term was significantly correlated with the two creativity indicators, meeting the first requirement for mediated moderation. As shown in Table 2, the interaction was also significantly related to emotional conflict ($\beta = .30$, $p < .05$), meeting the second requirement. As shown in Table 3, emotional conflict significantly predicted idea generation ($\beta = -.24$, $p < .05$) but not overall creativity ($\beta = -.04$, ns) when controlling for the main effects and interaction, meeting the third requirement for mediation for idea generation. After controlling for emotional conflict, the variance in idea generation accounted for by the interaction was reduced from 6% to 3%. This reduction was significant by Sobel’s (1982) test, $Z = 1.82$, $p < .05$ (one-tailed). These results provide partial support for Hypothesis 2.

Discussion

At Zig, a leading Canadian ad agency, promotional campaigns are developed for products ranging from beer to home appliances to breast cancer awareness. Recognizing the potential divisiveness of these products, Zig staffs creative teams composed of all men, all women, and mixed genders and then assigns them according to the gender focus of the project (Yaccato, 2003). However, researchers have identified a number of other organizations that are doing just the opposite, striving to increase the diversity of their teams in an attempt to stimulate creativity and innovation (Taylor & Greve, 2006; Thompson & Brajkovich, 2003). Unfortunately the literature has been unable to offer any consistent results regarding the superiority of either approach (e.g., Gibson & Gibbs, 2006; Sethi, Smith, & Park, 2002).

Therefore, the primary purpose of this study was to use faultline theory to examine the effects of gender diversity on team creativity and to explore the critical role of faultline activation. According to faultline theory, whether gender diversity will benefit or inhibit creativity is contingent on both the possible social categorizations within the team and the salient features of the team’s task. As shown in Figure 1, results supported our main hypothesis, as both the level and creativity of idea generation were significantly lower in teams with activated gender faultlines than in teams with no faultlines. It is more notable that teams with activated gender faultlines were also significantly less creative than teams with faultlines that were not activated.

These results suggest that mere existence of gender faultlines is not sufficient for social comparison to occur; subgroup formation requires a trigger in the form of a gender-salient team task. This extends our understanding of the contingencies influencing subgroup formation within teams and provides support for the role of activation in faultline theory. These results also extend faultline theory to encompass the effects of gender diversity on team creativity, as, consistent with theory, the activation of gender faultlines by an exogenous task led to the formation of subgroups that impaired team participation and divergent thinking.

Our results also support researchers’ arguments that the effects of diversity are complex and cannot adequately be accounted for by examining the simple main effects of gender dispersion (e.g.,

![Figure 1](image.png)
Harrison & Klein, 2007; Jehn et al., 1999; van Knippenberg & Schippers, 2007). In this study, traditional measures of gender dispersion would have failed to recognize the multifaceted relationship between gender and team creativity. Gender differences appear to exhibit an effect on team creativity only under specific conditions: The gender composition of the team must support the presence of potential subgroups, and social categorizations that activate those subgroups must be triggered by a gender salient contingency, such as the task. Previous research has, for the most part, taken a more one-dimensional approach, believing that it is enough for team members to just be different from one another.

The second purpose of this study was to identify the mediating processes through which gender faultline activation affects team creativity. We expected that the activated gender faultlines would impair team creativity by influencing the level of conflict within the team. Our results partly supported our hypothesis. We found that emotional conflict partially mediated the negative effects of activated gender faultlines on the number of ideas a team generated. Our results suggest that when gender faultlines are triggered, friction develops between male and female subgroups, reducing communication about ideas and opinions that benefits team creativity. Due to their antipathy, teams with activated faultlines failed to take advantage of opposing viewpoints, to delve into issues more deeply, or to develop more innovative alternatives (Tjosvold, 1986). When dormant, however, gender faultlines in the team did not affect emotional conflict within the team. Social categorizations leading to subgroup formation began to affect interpersonal friction only when gender was made salient to team members through activation.

Managerial Implications

Researchers have suggested that managers should attempt to inoculate creative teams against subgroup formation in the initial stages of interaction by increasing the identification of team members with their team rather than their subgroup (e.g., Pratt, Fuller, & Northcraft, 2000). One approach is to select team members with strong commonalities in backgrounds and values. For example, Hargadon and Sutton (1997) described the way creative design firm IDEO addresses this issue by carefully recruiting engineers with common educational backgrounds, selecting 70% of design team members from the mechanical engineering program at Stanford University.

Limitations and Directions for Future Research

One limitation of our study is the nature of the activation task. In that task we increased the salience of gender differences by having mixed-gender teams brainstorm about a gender-relevant product. Activating other faultlines (e.g., racial), however, may be difficult without relying on inflammatory or stereotypical differences. Such activations may capitalize more on stereotype threat or heightened intergroup bias than on social categorizations, resulting in greater conflict and even lower performance. Similarly, it is possible that the negative effects of activating faultlines based on more status-laden demographic differences such as tenure or managerial level may be due in part to awareness of status differences within the team. Although status differences between demographic subgroups are consistent with faultline theory (Lau & Murnighan, 1998), activation tasks that highlight these status differences may increase intrateam divisiveness. For more deep-level nondemographic characteristics such as geographic location and conscientiousness (e.g., Polzer et al., 2006; Rico, Molleman, Sanchez-Manzanares, & Van der Vegt, 2007), it is unclear how they become activated within the team and whether there are stereotypes or status differences at work. Future research is needed to explore the activation of different faultlines to parse out the effects of social categorization from stereotypes or status-induced bias.

A second limitation is our focus on the activation of a single faultline. Most team demographic profiles are likely to contain multiple faultlines from which subgroup formations can arise. Based on faultline theory, we would expect the results of this study to hold true and perhaps be even stronger for teams with multiple, similarly aligned subgroups. Nevertheless, a consideration for future research is how the activation of one faultline is likely to influence the potential activation of others. Once subgroups are formed along one demographic difference, the creation of new subgroups based on different characteristics may be repressed (Dyck & Starke, 1999). Faultline theory would benefit from research that explores the activation of multiple faultlines once initial subgroup formation has occurred.

Third, because this study was conducted in a laboratory context, future research needs to examine the external validity of these results. Although the nature of the research participants’ experiences did not exactly mirror those of a real organizational situation, there were certain features of this task and our participants that achieved what Berkowitz and Donnerstein (1982) referred to as “mundane realism.” Generating ideas for commercial products

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Idea generation</th>
<th>Overall creativity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Step 1</td>
<td>Step 2</td>
</tr>
<tr>
<td>Emotional conflict</td>
<td>-.32**</td>
<td>-.27**</td>
</tr>
<tr>
<td>Faultline</td>
<td>-.20*</td>
<td>-.01</td>
</tr>
<tr>
<td>Activation</td>
<td>-.16</td>
<td>-.03</td>
</tr>
<tr>
<td>Faultline × Activation interaction</td>
<td>.10</td>
<td>.17</td>
</tr>
<tr>
<td>Total $R^2$</td>
<td>.07</td>
<td>.03</td>
</tr>
</tbody>
</table>

$p < .05$. **$p < .01$. One-tailed tests.
is not that far removed from the actual work of advertising and engineering design teams. We are simply asking the “can it happen” question, which is exactly the type of question that bears investigation in a laboratory setting (Ilgen, 1986). As Driskell and Salas (1992) noted, “experimental research is generalized on the basis of the theoretical relationships that are tested, not through the concrete results of a single study” (p. 113).

Finally, we examined teams in the early stages of interaction where faultline activation and creativity occurred during a short period of time. It may be that the effects of subgroup formation triggered by a single task will fade as team members continue to interact in non-gender-salient contexts. As well, the influence of surface-level characteristics such as gender may decline over time as team members build relationships based on nondemographic commonalities, such as values or personality (e.g., Harrison, Price, Gavin, & Florey, 2002; Jehn et al., 1999). However, if faultlines are activated early in a team’s development, repeated faultline-salient interactions may act instead to legitimize and polarize the subgroup (Lau & Murnighan, 1998). Clearly, additional research is needed to examine the effects of single and repeated faultline activation over time.

References
constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, 32, 1199–1228.


Received February 5, 2007
Revision received June 18, 2007
Accepted August 31, 2007