Journal of Organizational Behavior
J. Organiz. Behav. **31**, 1032–1054 (2010)
Published online 15 September 2009 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/job.661

Group-level organizational citizenship behavior: Effects of demographic faultlines and conflict in small work groups

JIN NAM CHOI[1]* AND THOMAS SY[2]
[1]*Seoul National University, Seoul, South Korea*
[2]*University of California, Riverside, California, U.S.A.*

Summary Organizational scholars have recently shifted their attention to examining organizational citizenship behavior (OCB) as a group-level phenomenon. Adopting the input-mediator-outcome model of group performance, we examined antecedents and intermediate processes that predict group-level OCB (GOCB) in small work groups. The results, based on data from 62 work groups representing a variety of industries, revealed that demographic faultlines based on relation-oriented attributes (gender, age, and race) and a task-related attribute (tenure) had differentiated relationships with task and relationship conflict, which mediated the relationships between faultlines and group outcomes (GOCB and group performance). Both task and relationship conflict were negative predictors of group performance. However, task conflict increased GOCB, whereas relationship conflict decreased it. The present study offers evidence of the relationship between demographic faultlines and various group processes and outcome variables in natural work groups. Copyright © 2009 John Wiley & Sons, Ltd.

Introduction

Researchers have made repeated calls for research on organizational citizenship behavior (OCB) at levels other than the individual level (e.g., Organ & Ryan, 1995; Podsakoff, MacKenzie, Paine, & Bachrach, 2000). Nevertheless, most empirical studies have persisted in framing OCB as an individual-level phenomenon (Bommer, Dierdorff, & Rubin, 2007; Organ, Podsakoff, & MacKenzie, 2006). Somech and Drach-Zahavy (2004) maintained that studies of OCB at the individual level of analysis ''fall short of fully capturing the OCB phenomenon'' (p. 292). Choi (in press) also pointed out that ''empirical findings related to antecedents of individual-level OCB cannot be automatically generalized to the group level'' (p. 3). Prior investigations have shown that there is significant group-level variation in OCB (Schnake & Dumler, 2003), indicating that OCB may be a group-level phenomenon. Empirical studies have also revealed that group-level OCB is a significant predictor of work unit performance (Podsakoff, Whiting, Podsakoff, & Blume, 2009).

* Correspondence to: Jin Nam Choi, College of Business Administration, Seoul National University, Shinlim-dong, San 56-1, Gwanak-gu, Seoul, 151-742, South Korea. E-mail: jnchoi@snu.kr

Received 2 May 2008
Revised 8 July 2009
Accepted 30 July 2009

Recently, scholars have begun to investigate antecedents of group-level OCB, although the focus has been limited to various leadership issues (e.g., Lau & Lam, 2008; Mayer, Kuenzi, Greenbaum, Bardes, & Salvador, 2009). The present study contributes to this emerging literature on group-level OCB (hereafter referred to as *GOCB* to distinguish it from OCB at the individual level) by identifying and empirically investigating a set of group-level antecedents of GOCB in small work groups. To this end, we adopt the widely accepted framework of the input-mediator-outcome model that conceptualizes a group as a system that produces outcomes using various inputs to the group through intermediate interactive processes (Mathieu, Maynard, Rapp, & Gilson, 2008). Thus, we examine group composition, particularly demographic faultlines, as an ''input'' to the group that shapes interpersonal processes in terms of task and relationship conflict among members. We further predict that these input and mediating processes affect group outcomes including GOCB (a behavioral outcome) and actual group performance (see Figure 1 for a summary of this framework).

To examine the effects of group composition on GOCB, we adopted the construct of ''demographic faultlines'' (Lau & Murnighan, 1998), defined as *hypothetical dividing lines* that may split members of a group into subgroups based on the combined effects of various attributes of the group members (e.g., gender, age, etc.). Owing to dissatisfaction with traditionally simple conceptualizations of diversity (e.g., measures of dispersion), faultlines emerged as a more promising and sophisticated conceptualization to explain the conflicting results in the diversity literature (Gibson & Vermeulen, 2003; Thatcher, Jehn, & Zanutto, 2003; van Knippenberg & Schippers, 2007). Faultlines and the resulting formation of subgroups may shape interpersonal perceptions and behaviors among members. Despite their intuitive appeal, investigations of faultlines have been limited to either laboratory settings and/or work groups with clearly preexisting subgroups such as gender-divided groups and two groups with different cultural backgrounds (Homan, van Knippenberg, van Kleef, & De Dreu, 2007b; Pearsall, Ellis, & Evans, 2008; Sawyer, Houlette, & Yeagley, 2006). Empirical research on faultlines is still in its early stage, and related findings are not consistent and straightforward perhaps due to the lack of sufficient empirical evidence in various settings (van Knippenberg & Schippers, 2007).

Despite repeated calls to examine the effects of faultlines in natural work group settings (Homan, Hollenbeck, Humphrey, van Knippenberg, Ilgen, & van Kleef, 2008), the implications of demographic faultlines in natural organizational groups have received limited attention perhaps due to the complexity involved in computing faultlines with no predetermined, salient dividing lines among members. Moreover, existing studies have typically employed a single, collapsed measure of faultlines often because a predetermined dividing line among members already exists (Lau & Murnighan, 2005; Thatcher, Jehn, & Zanutto, 2003). The present study provides an empirical examination of the faultline construct in a natural work group setting considering the simultaneous effects of multiple attributes and their combined effects (e.g., gender-driven faultlines, tenure-driven faultlines). To this end, we adopt Shaw's (2004) measure of faultlines that accurately reflects Lau and Murnighan's original conceptualization of group faultlines.

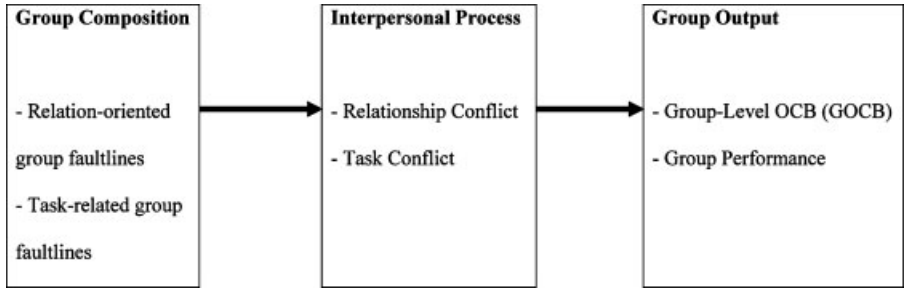


Figure 1. Input-mediator-output model of group-level OCB

DOI: 10.1002/job

In summary, the present theoretical framework and empirical findings meaningfully expand the OCB and diversity literature in several ways. First, it provides greater understanding of the group characteristics that lead to a greater level of GOCB by identifying theoretically plausible input and mediating processes. Second, it addresses the limitation of the diversity literature that has principally relied on simple measures of diversity based on proportion or dispersion of a single attribute instead of measures that consider configurations of multiple member characteristics. Third, our empirical findings provide an elaborate understanding of task and relationship conflict as meaningful mediators between group input and outcome variables, as well as their distinct dynamics in predicting GOCB and group performance. Finally, examining the study variables within the input-mediator-outcome framework addresses the historical weakness in group research of not considering these variables simultaneously (Hackman, 1992).

Group-Level OCB (GOCB)

We define GOCB as ''the overall level of group members' behavior that contributes to the maintenance and enhancement of the social and psychological context of the group that facilitates its task performance'' (adapted from Borman & Motowidlo, 1997). Individual-level OCB and GOCB represent conceptually and empirically distinct phenomena (Bommer et al., 2007). For this reason, individual OCB and GOCB are likely to have distinct construct functions in terms of both antecedents and outcomes (Chen, Mathieu, & Bliese, 2004). For example, whereas individual OCB can be predicted by individual dispositions (e.g., personality, values) and attitudes (organizational commitment), GOCB is more likely to be predicted by collective properties (e.g., group processes, task interdependence) (Choi, in press).

The majority of existing studies of OCB at the group level have paid attention to the GOCB-performance link (Podsakoff et al., 2000). Empirical studies have shown that GOCB is significantly related to the quantity and quality of group performance, including efficiency of operation, profitability, and customer satisfaction (Podsakoff et al., 2009). Scholars have recently begun to identify antecedents of GOCB, mostly focusing on various leadership constructs such as servant leadership (Ehrhart, 2004), transformational leadership (Richardson and Vandenberg, 2005), ethical leadership (Mayer et al., 2009), and leader trustworthiness (Lau & Lam, 2008).

However, collective behaviors such as GOCB may be better understood by looking into the interpersonal dynamics among members. Inter-member relationships, instead of or in addition to leader–member relationships, should be particularly meaningful for GOCB, which represents members' voluntary collaborations and contributions beyond those required and possibly enforced by the leader and formal system (Organ et al., 2006). For this reason, we focus on inter-member dynamics that may affect the level of GOCB. We specifically examine these dynamics within the context of small work groups because it is at this meso level (House, Rousseau, & Thomas-Hunt, 1995) that members have the most contact and interaction with each other; thus, inter-member dynamics could have the most robust effect on GOCB in small work group settings.

Input-Mediator-Outcome Model of Group-Level OCB

As depicted in Figure 1, with regard to input to the group, we focus on group composition, particularly the demography of group members. Group composition has been regarded as the most

DOI: 10.1002/job

fundamental ''input'' to any social situation that shapes various social processes and interpersonal attitudes (Williams & O'Reilly, 1998). Regarding the mediating process, we attend to task and relationship conflict because the effect of group composition on group outcomes, such as GOCB and performance, are most likely indirect in that group composition acts on intervening group dynamics (i.e., conflict) that have a more proximal effect on group outcomes (van Knippenberg, De Dreu, & Homan, 2004). The linkage between group composition and conflict is well established (Jehn, Northcraft, & Neale, 1999; Mohammed & Angell, 2004; Pelled, Eisenhardt, & Xin, 1999). Lau and Murnighan (1998) also proposed that faultlines brew conflict across subgroups. Furthermore, conflict has been widely recognized as critical intermediate process that have significant implications for both behavioral and task outcomes of the group (De Dreu & Weingart, 2003). As such, we focus on task and relationship conflict as key intervening processes in our model. Thus, our selection of variables, specified relationships, and hypothesized direction of causality are informed by, and consistent with, existing theory and empirical findings (e.g., De Dreu & van Vianen, 2001; Ng & Van Dyne, 2005). Below we develop the hypotheses for each relationship in this model.

Group input: Demographic faultlines

Given that members compose a group's main source of input and resources and that the most salient social cues often stem from demographic characteristics (Harrison & Klein, 2007; Williams & O'Reilly, 1998), group composition carries significant implications for both the internal functioning and performance of the group. Based on Lau and Murnighan's (1998) proposition that group behavior is affected by *configurations* of multiple demographic characteristics (*cf*. diversity structure, Sawyer et al., 2006), scholars have recently argued that the lack of consistent findings in the diversity literature may be due to the focus on disconnected effects of single demographic attributes—a focus that ignores the combined effects of multiple demographic attributes (Thatcher et al., 2003).

Lau and Murnighan (1998) argued that the dimensions of demographic diversity may impede group process and performance when they combine and develop strong dividing lines (i.e., *faultlines*) among members. Strong faultlines are likely to bring about destructive perceptions and interactions between members of different subgroups (e.g., stereotypes, biases, disliking, conflict, scorn, competition, etc.). Strong faultlines are characterized by high member similarity *within* subgroups and low member similarity *between* subgroups (Shaw, 2004). The lack of meaningful demographic overlaps *between* subgroups under the strong faultlines scenario is likely to result in the development of barriers between subgroups and diminish the quality of inter-subgroup communication and empathy (Lau & Murnighan, 2005). The detrimental effects of faultlines on interpersonal processes were demonstrated in a recent study by Homan, van Knippenberg, van Kleef, and De Dreu (2007a), in which groups exhibited more conflict, less satisfaction, and more negative climate particularly when informational diversity converged with the faultline based on the alignment of gender and bogus personality feedback.

Expanding prior findings based on lab-manipulated faultline strengths, we propose that faultlines based on different demographic attributes have distinct implications for task and relationship conflict in organizational teams. Jehn (1995) defined relationship conflict as interpersonal incompatibilities among group members that are characterized by frustration, friction, and interpersonal clashes within the group. Task conflict refers to group members' disagreements over ideas, content, and issues related to the task. The diversity literature developed a similar distinction between relation-oriented and task-

DOI: 10.1002/job

related diversity (Jackson, May, & Whitney, 1995), although researchers have used various labels such as more and less job-related diversity (Webber & Donahue, 2001) and informational and social-category diversity (Jehn et al., 1999).

A central aspect of faultlines is that groups may have multiple faultlines that remain dormant until they are activated (Lau & Murnighan, 1998; Pearsall et al., 2008). The activation of a faultline may be triggered by (a) the salient phenotypic attributes of group members (e.g., when a six-member group consists of three Caucasian males and three African-American females) or (b) a group's task context. Gender, age, and race are salient phenotypic attributes from which social categorizations and identities are naturally formed (Tajfel, 1978; Turner, 1987). For this reason, when faultlines are activated based on attributes such as gender, age, and race, they create social categories that subsequently lead to stereotyping and other interpersonal biases that promote relationship conflict (Byrne, 1971; Lincoln & Miller, 1979). For example, groups with gender–age faultlines, which become salient when members' gender and age are well aligned in a group (e.g., old males vs. young females), may suffer from interpersonal, emotional clashes due to the strong subgroups resulting in stereotypes, and biased perceptions. Thus, we expect that faultlines generated by phenotypic attributes (e.g., gender–age, gender–race, and age–race faultlines) will increase relationship conflict among members.

Unlike gender, age, and race, tenure is not a phenotypic attribute (e.g., there is no direct physical correlate that indicates how much tenure one has with a particular company). Instead, tenure is work-based and represents organization and task seniority. As such, task contexts are most likely to activate tenure-based faultlines. Likewise, tenure-based faultlines are most likely to trigger task-related disagreements because how work gets done and who should be responsible for certain aspects of the work is often determined by one's previous work and task experience (i.e., tenure) (Ancona & Caldwell, 1992; Zenger & Lawrence, 1989). Thus, we suggest that tenure-based subgroupings (e.g., tenure–gender, tenure–age, etc.) will increase task conflict among group members.

In summary, subgroups with strong faultlines based on relation-oriented social categories such as age, gender, and race are likely to stimulate social cognitive processes (e.g., stereotyping) and affective reactions (e.g., anxiety, hostility) (Jackson et al., 1995), increasing relationship conflict. Subgroups with strong faultlines involving task-related characteristics such as tenure or work-related experiences and expertise, on the other hand, may promote information seeking/sharing and learning among members (Jackson et al., 1995), which may promote task conflict. Thus, faultlines triggered by different demographic attributes are expected to increase different types of conflict.

Hypothesis 1a. Strengths of relationship-oriented faultlines (gender–age, gender–race, and age–race faultlines) will be positively related to relationship conflict.

Hypothesis 1b. Strengths of task-related faultlines (tenure–gender, tenure–age, and tenure–race faultlines) will be positively related to task conflict.

Task and relationship conflict and group outcomes

Conflict has been one of the most widely investigated process variables in the group literature (Hempel, Zhang, & Tjosvold, 2009; De Dreu & Weingart, 2003; Tekleab, Quigley, & Tesluk, 2009). However, to the best of our knowledge, the only empirical study on the relationship between group conflict and group-level OCB was conducted by De Dreu and Van Vianen (2001): Relationship conflict is negatively related to group-level helping and group member compliance. Relationship conflict

DOI: 10.1002/job

represents ''conflict over work group members' personal preferences or disagreements about interpersonal interactions'' (Jehn et al., 1999, p. 745) and thus tends to generate negative emotions and dysfunctional attitudes (such as antagonism) that may reduce cooperation as well as group performance (Sy, Cote, & Saavedra, 2005). Relationship conflict due to differences in interpersonal styles, preferences, and personality is more likely to generate (mostly negative) affect-laden interactions (van Knippenberg et al., 2004) because values and personalities are central to members' personal identity and sense of self-worth (Pelled et al., 1999). Relationship conflict may also generate a vicious cycle of detrimental social exchange in which the lack of mutual appreciation among members leads to a reduced level of GOCB within the group (e.g., ''You're not nice to me, so I won't help you'').

Hypothesis 2a. Relationship conflict will be negatively related to GOCB.

Hypothesis 2b. Relationship conflict will be negatively related to group performance.

Unlike relationship conflict, task conflict has been proposed to be a source of creativity and sound decision making, thereby increasing group performance (Jehn, 1995). Despite many intuitively appealing reasons for why task conflict enhances group effectiveness, a recent meta-analysis by De Dreu and Weingart (2003) showed that task and relationship conflict have negative effects on member satisfaction and group performance, although the negative effects are weaker for task conflict. Perhaps task conflict bears a relational consequence (Homan et al., 2007a) because conflict is inherently affect-laden (Sy et al., 2005) and reflects back on group members' relationships (Mooney, Holahan, & Amason, 2007). Thus, task and relationship conflict may be concomitantly linked, and group members may find it difficult to distinguish between them (e.g., providing critical feedback for a colleague's ideas may be interpreted as a personal attack on that colleague's intelligence).

The decision making literature has demonstrated that people feel uncomfortable when considering ideas and information different from their own perspectives and that the group decision process is driven by many biases and motivational factors that may not be rational (Brodbeck, Kerschreiter, Mojzisch, & Schulz-Hardt, 2007). Thus, encountering different ideas (that may negate one's position) may be interpreted as a threat to one's self-esteem (Turner, Pratkanis, Probasco, & Leve, 1992), a situation in which individuals are unlikely to support others' ideas. Moreover, task conflict tends to impose high cognitive load and arousal on members that may distract and waste group resources (De Dreu & Weingart, 2003). In addition, task conflict may slow down the decision making process owing to the increased need for coordination and negotiation, which deprives the group of valuable time for implementation and hinders group performance (Choi & Kim, 1999).

Hypothesis 3a. Task conflict will be negatively related to GOCB.

Hypothesis 3b. Task conflict will be negatively related to group performance.

Mediating role of group conflict

Although demographic faultlines have meaningful implications for group performance (Lau & Murnighan, 2005; Thatcher et al., 2003), they are more likely to influence group performance indirectly by shaping interpersonal dynamics among members (van Knippenberg & Schippers, 2007). van Knippenberg et al. (2004) pointed out that differences among members in social categories, values, and ideas tend to be *inherently* neutral in their implications for performance: Any negative or positive consequences result from the way the group cognitively and emotionally responds to them. Similarly, we argue that demographic faultlines, as a form of group input, will exert significant effects on

DOI: 10.1002/job

behavioral and performance outcomes only when they affect interpersonal processes, such as task and relationship conflict, that are pertinent to the outcomes.

Hypothesis 4. The relationship between demographic faultlines and group outcomes (GOCB and group performance) will be mediated by task and relationship conflict.

Method

Sample and data collection procedure

The current study is part of a larger data collection effort. Accordingly, we report only the details pertaining to the current study. Participants were recruited by a group of trained undergraduate students who received extra credit in their grades in management courses. All students received approximately an hour of training on study protocol and ethical guidelines. In addition to the one-hour training session, all students received detailed written protocol instructions (including a standardized participant recruitment script). The research group distributed surveys to non-student working adults from their existing network of contacts. Data were collected from groups composed of a leader and a minimum of three group members. Students approached work units that fit the definition of groups as semi-autonomous work groups with members who are mutually responsible for accomplishing a set of tasks (Guzzo & Shea, 1991). Participants independently returned their completed surveys using a pre-addressed, stamped envelope provided by the researchers.

Consistent with past research employing similar methods of recruiting participants from existing contact networks (e.g., De Dreu & van Vianen, 2001), the response rate was quite high, reaching 94.4 per cent. The initial sample consisted of 1103 individuals in 292 work groups. At the group level, missing data invite both random and systematic biases that often cannot be properly controlled (Timmerman, 2005), and the problem seems to persist with diversity measures (Allen, Stanley, Williams, & Ross, 2007). The missing data problem can be particularly serious for group-level measures, such as social network properties or faultline strengths, because these measures are most likely to be distorted even by a small number of missing group members. For this reason, we selected groups with 100 per cent membership representation for our analysis. With this screening criterion applied, the final analysis sample included 248 individuals in 62 groups (all four-person groups).

Participants were recruited from a variety of companies located in a large metropolitan area in the western region of the United States. The average company employed 4458 employees. The present sample included a wide range of industries including retail (15 per cent), food services (14 per cent), professional services (10 per cent), healthcare (9 per cent), transportation (9 per cent), telecommunication (7 per cent), and financial services (7 per cent). Participants worked a minimum of 20 hours per week, with an average work week of 35.4 hours ($SD = 10.35$). Forty-four per cent were men. Average company tenure of the participants was 2.8 years ($SD = 4.23$). Age ranged from 18 to 65, with an average age of 30.6 ($SD = 9.76$). Participants' education levels were high school (36 per cent), vocational school (17 per cent), bachelors degree (33 per cent), and graduate degree (8 per cent). The sample consisted of Asian Americans (33 per cent), Hispanics (25 per cent), White (22 per cent), African Americans (8 per cent), and other (12 per cent). These demographic characteristics were not different between the members of the final analysis sample of 62 groups and those of the initial sample of 292 works groups (all $p > .10$), except that our analysis sample included slightly more Asians (33 per cent vs. 25 per cent) and shorter organizational tenure (2.8 years vs. 3.7 years) than the initial sample.

DOI: 10.1002/job

At the group level, the current sample included 12.9 per cent all male, 17.7 per cent all female, and 69.4 per cent mixed-gender groups. Whites, Blacks, Asians, and Hispanics were represented in 50.0, 22.6, 53.2, and 53.2 per cent of the groups, respectively. The average age for each group was 30.6 ($SD = 6.54$), with a range of 19.5–43.3. The average organizational tenure for each group (ranging between 1.0 and 13.8) was 3.7.

Measures

Group members and leaders completed different survey instruments. Group leaders rated the level of task and relationship conflict within their groups, as well as their groups' overall performance. Group members reported the level of GOCB within their groups. Given the small size of the groups and the centrality of the leaders within their groups, leaders were likely to have daily contact with group members and intimate knowledge about the group, its members, and its functioning. Thus, leaders were likely very well informed and reliable judges for task and relationship conflict, and overall group performance.

Demographic faultlines

In this study, we examined the effects of faultlines based on three relation-oriented attributes (gender–age, gender–race, and age–race faultlines) and those involving task-related attributes (tenure–gender, tenure–age, and tenure–race faultlines). These attributes have been investigated the most in the diversity literature (Williams & O'Reilley, 1998). We employed the faultline strength formula developed by Shaw (2004) that offers a way to assess appropriately the extent to which a group is divided into subgroups based on multiple demographic attributes. Shaw's (2004) measure assesses faultline strength in several steps: (a) Assessing *subgroup internal alignment* (IA) (ranging between 0 and 1), which refers to ''the extent to which members within a particular subgroup are similar to one another on all other relevant attributes'' (p. 72); (b) computing *cross-subgroup alignment* (CGAI) (ranging between 0 and 1), which refers to the extent to which members of different subgroups are similar to each other on all other relevant attributes; and (c) calculating the faultline strength score: IA $\times$ (1-CGAI). Thus, the faultline strength score is high when subgroup internal alignment is high, *and* cross-subgroup alignment is low (we present a detailed numerical example of the computation procedure involved in Shaw's measure of faultline strength in the Appendix).

For example, in a group consisting of two 20-year-old females and two 50-year-old males, gender–age faultlines become strong because the gender subgroups also exhibit similar levels of age (i.e., high subgroup internal alignment). In contrast, in a group composed of a 20-year-old female, a 20-year-old male, a 50-year-old female, and a 50-year-old male, gender–age faultlines are weak because the gender subgroups do not share the age level internally. Instead, the age level is similar across gender subgroups (i.e., high cross-subgroup alignment). This simultaneous consideration of within-subgroup similarity and cross-subgroup similarity accurately reflects the idea of group faultlines (Lau & Murnighan, 1998).

To calculate the strengths of faultlines based on the four demographic characteristics, we used Chung, Shaw, and Jackson's (2006) program that was designed to calculate Shaw's faultlines measure. We recoded age and tenure because Chung et al.'s program allowed only categorical data. Specifically, we created four categories for both age (1 = 20 or younger; 2 = between 21 and 30; 3 = between 31 and 40; 4 = 41 or older) and tenure (1 = one year or less; 2 = between one and two years; 3 = between two and five years; 4 = five years or more). These categories represented typical age or tenure blocks, and provided a relatively equal allocation of participants into age and tenure subgroups.

DOI: 10.1002/job

Relationship conflict

We measured relationship conflict with a three-item scale ($\alpha = .94$) developed by Jehn (1995). This scale included (a) ''How much tension is there among members in your work unit?,'' (b) ''How much friction is there among members in your work unit?,'' and (c) ''How much are personality conflicts evident in your work unit?'' Group leaders rated these items using a seven-point Likert-type scale (1 = *a little*, 7 = *a lot*).

Task conflict

To measure task conflict, we adopted three items ($\alpha = 0.85$) developed by Jehn (1995): (a) ''How often do people in your work unit disagree about opinions regarding the work being done?,'' (b) ''How frequently are there conflicts about ideas in your work unit?,'' and (c) ''How much conflict about the work is there in your work unit?'' Group leaders rated these items using the same Likert-type scale for relationship conflict.

Group-level OCB

We measured GOCB with five items (individual-level $\alpha = 0.91$, group-level $\alpha = 0.94$) adopted from Wayne, Shore, and Liden (1997). Sample items included ''Members of my group help others with their work when they have been absent even when they are not required to do so,'' ''Members of my group make innovative suggestions to improve the overall quality of the department,'' and ''Members of my group willingly attend functions not required by management but which help our group's overall image.'' The scale items reflected various content domains of OCB including helping, conscientiousness, initiative, and civic virtue (Organ et al., 2006). Group members rated these items using a seven-point Likert scale (1 = *never*, 7 = *always*). We used r_{wg} and ICC statistics to determine whether there was sufficient within-group agreement and between-group variation to justify group-level aggregation of the GOCB scale (Bliese, 2000; Chen et al., 2004; James, Demaree, & Wolf, 1984). The GOCB scale exhibited an r_{wg} value of 0.83 and statistically significant ICC(1) and ICC(2) values of 0.24 and 0.50, respectively ($F = 1.95$, $p < 0.001$). All these indices were in the acceptable range, and thus we aggregated the responses of group members to create a group-level measure of GOCB for each group.

Group performance

Leaders of each group evaluated their group's performance with two items ($\alpha = .89$) adopted from Wayne et al. (1997): (a) ''In my estimation, members of my group get their work done very effectively'' (1 = *strongly disagree*, 7 = *strongly agree*) and (b) ''Rate the overall level of performance of your group members'' (1 = *low performance*, 7 = *high performance*).

Results

Table 1 presents the descriptive statistics of the study variables. Our final analysis sample included only 62 groups, which invites concern regarding the reliability of estimates owing to the small sample size. To address this issue, we employed partial least square (PLS) modeling (Chin, 1998), which tends to produce robust results facing various inadequacies such as missing values, model misspecification, multicollinearity among observed and latent variables, and violation of the usual statistical assumptions such as multivariate normality (Cassel, Hackl, & Westlund, 1999; Chin & Newsted, 1999). Studies based on Monte Carlo simulations have demonstrated that compared with the structural equation modeling (SEM) analysis (Bollen, 1989), the PLS procedure generates more robust estimates particularly when the sample size is relatively small (Qureshi & Compeau, 2009).

DOI: 10.1002/job

Table 1. Means, standard deviations, and correlations among study variables (group level, $N = 62$)

Variables	*M*	*SD*	ICR	AVE	1	2	3	4	5	6	7	8	9	10
1. Gender–Age faultlines	0.14	0.23			—									
2. Gender–Race faultlines	0.03	0.06			0.23	—								
3. Age–Race faultlines	0.04	0.07			0.19	0.32*	—							
4. Tenure–Gender faultlines	0.08	0.11			0.067	−0.11	−0.05	—						
5. Tenure–Age faultlines	0.09	0.10			−0.04	0.12	−0.02	0.34**	—					
6. Tenure–Race faultlines	0.05	0.09			−0.15	−0.09	0.25*	−0.06	0.06	—				
7. Relationship conflict	2.97	1.42	0.96	0.89	0.22	0.14	0.14	0.23	0.28*	−0.19	**0.94**			
8. Task conflict	2.83	1.16	0.91	0.76	0.17	0.07	0.00	0.29*	0.23	0.12	0.58***	**0.87**		
9. Group-Level OCB	4.65	1.05	0.95	0.80	−0.15	−0.03	−0.06	0.06	−0.01	−0.03	−0.27*	−0.01	**0.89**	
10. Group performance	5.64	0.97	0.94	0.90	−0.15	−0.01	−0.18	−0.02	0.03	−0.03	−0.40**	−0.44***	0.39**	**0.95**

Note: ICR, Internal Composite Reliability; AVE, Average Variance Extracted; bold figures on the diagonal represent the square root of AVE.
*$p < .05$; **$p < .01$; ***$p < .001$.

DOI: 10.1002/job

PLS implementation guidelines

Although the PLS analysis offers a reasonable alternative to other approaches such as ordinary least-squares (OLS) regressions and covariance-based SEM, it should not be regarded as a panacea for deficient research design. To avoid the potential pitfall of blindly applying PLS to inadequate data, we implemented Marcoulides and Saunders' (2006) guidelines. First, we ensured that our theoretical model reflects the most recent developments and theoretical knowledge in the relevant literature. Second, we implemented proper data screening procedures and removed groups with non-responses from members and/or the leader. We also checked the normality of the observed variables and found that all estimates of skewness and kurtosis were within the acceptable range (although the leaders' group performance ratings were significantly and positively skewed).

Third, we examined the psychometric properties of the variables in our model. All measures were based on validated scales used in prior studies and exhibited acceptable reliability coefficients. The statistical procedure of PLS provides the internal composite reliability (ICR) of the measures in the model (Chin, 1998). As reported in Table 1, the four variables in this study exhibited high ICR coefficients. In PLS, convergent validity is established by examining the average variance extracted (AVE) or average communality for the latent factor. In the present data, AVE ranged between 0.76 and 0.90, which was greater than the minimum of 0.50 (Fornell & Larcker, 1981). The current measures also exhibited discriminant validity in that the square root of AVE for each construct was greater than its correlations with other constructs (i.e., within-scale communality is greater than shared variance across measures) (Fornell & Larcker, 1981). In the present data, the square roots of AVE ranged between 0.87 and 0.95, whereas inter-scale correlations were between -0.44 and 0.58.

Finally and perhaps most importantly, Marcoulides and Saunders (2006) underscored that PLS should not be used as a cure for a study with an insufficient sample size. A typical rule of thumb that has been cited often was Chin's (1998) recommendation that the sample size should be greater than ten times the largest number of indicators for a latent factor or the largest number of predictors for a latent outcome. In our model, the largest number of indicators was five (for GOCB) and that of the predictors was three (for relationship and task conflict). Thus, following Chin's recommendation, the minimum sample size required is 50, and the present sample size of 62 is satisfactory. Marcoulides and Saunders challenged the wide use of this rule of thumb and suggested the need to assess actually the statistical power of the study. According to their Monte Carlo analysis results (shown in Table 1, Marcoulides & Saunders, 2006, p. vii), under the situation of high factor loadings (in our study, approaching 0.90) and moderate inter-factor correlations (in our study between 0.40 and 0.50), the sample size of 62 groups is adequate to achieve the power of 0.80 (an arbitrary, but typical guideline for power).

Conducting PLS path modeling

Because all the latent constructs in the model exhibited decent psychometric properties as described above, we tested the structural relations among the constructs as proposed in our theoretical model summarized in Figure 1. To this end, we performed an omnibus test of the entire model using SmartPLS 2.0, a PLS-based path modeling program (Ringle, Wende, & Will, 2005).

Following the convention of testing alternative theoretical models in conducting SEM (MacCallum, Wegener, Uchino, & Fabrigar, 1993), we tested theoretically plausible alternative structural relations among the present constructs. For example, it is possible that demographic faultlines directly affect GOCB and group performance, in addition to their indirect effects via task and relationship conflict. Therefore, we tested the possibility that either relation-oriented (gender–age, gender–race, and age–race) faultlines or task-related (tenure–gender, tenure–age, and tenure–race) faultlines exert direct

DOI: 10.1002/job

effects on GOCB and group performance. The significance of the increase in explained variance with the introduction of new structural relations was tested by Cohen's f^2, which can be computed by the formula: $(R^2_{\text{full}} - R^2_{\text{reduced}})/(1 - R^2_{\text{full}})$. f^2 values of 0.02, 0.15, and 0.35 are considered to represent small, medium, and large effect sizes, respectively. When we added three paths from faultline measures (either relation-oriented or task-related) to GOCB or group performance, the resulting f^2 values for the four alternative models ranged between 0.01 and 0.05. Given the relatively small effect sizes associated with these direct paths linking faultlines to GOCB and group performance, it seemed appropriate to retain the original conceptual model because parsimonious models are desirable.

Another theoretical possibility is that various faultlines affect both types of conflict. Although Hypotheses 1a and 1b propose distinct relationships between different types of faultlines and two types of conflict, it may be unrealistic to expect that the two mechanisms based on relationship and task are completely blocked from each other. For example, De Dreu and Weingart's (2003) meta-analysis showed that the two conflict measures are highly correlated ($r = 0.58$). Mooney, Holahan, and Amason (2007) showed that cognitive conflict tends to generate emotional conflict. For this reason, we empirically tested the possibility of crossover effects (e.g., from relationship-oriented faultlines to task conflict). The f^2 value for the three paths from the relation-oriented faultlines to task conflict was 0.05, which does not justify their inclusion in the model. In contrast, the three paths from the task-related faultlines to relationship conflict resulted in a meaningful change in explained variance from 0.07 to 0.20, which represents a medium effect size ($f^2 = 0.17$). This pattern indicates that the effects of task-related faultlines on relationship conflict are not negligible, and omitting them from the analysis may substantially reduce the explanatory power and validity of the model.

Hypothesis testing

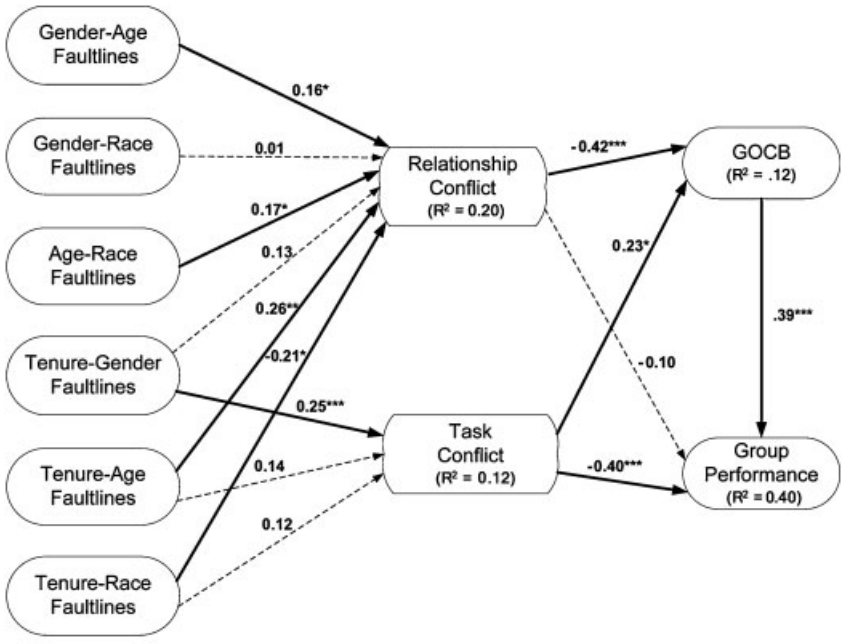
Figure 2 displays the structural model that includes the paths between task-related faultlines and relationship conflict.[1] For a simple presentation, this figure does not show the indicators of each latent construct (e.g., five items indicating GOCB), but as explained above, all factor loadings were high (all > 0.85). We used bootstrapping to obtain the standard errors of each path coefficients because the distribution of estimated parameters is unknown in PLS. Figure 2 reports the significance of each path based on *t*-values obtained by resampling 300 random samples (Preacher, Rucker, & Hayes, 2007).

Direct effects

Supporting Hypothesis 1a, two of the three relationship-oriented faultlines were significant, positive predictors of relationship conflict. Specifically, gender–age and age–race faultlines were positively related to relationship conflict ($\beta = 0.16$ and 0.17, respectively, both $p < 0.05$). Although not hypothesized, tenure–age faultlines also increased relationship conflict ($\beta = 0.26$, $p < 0.01$). Surprisingly, tenure–race faultlines were negatively related to relationship conflict ($\beta = -0.21$,

[1]Although we utilized a robust analytic strategy (i.e., PLS analysis) to deal with the relatively small sample size, it would be ideal to replicate the present findings with larger samples. To check if our results were substantially affected by the small sample size and the screening procedure, we conducted the same PLS path modeling using different sets of data: (a) 62 groups with full membership, (b) 112 groups with 70% of membership, and (c) 246 groups with approximately 30% of membership represented in this sample. In all three samples, the overall pattern of results remained the same. However, with the increasing number of groups included in the analysis, the results became less significant perhaps due to the increasing levels of non-responses and the ensuing problem of weakening the representativeness of the data against the entire group. The pattern of decreasing significance of research findings was particularly apparent for the relationships between faultlines and conflict measures than for the relationships between conflict and group outcomes. It was clear that compared with the measures of collective perceptions or behavior (e.g., conflict, GOCB), the adequacy or validity of faultlines measures meaningfully depend on the representativeness of the sample analyzed.

DOI: 10.1002/job



[a] Dotted lines represent non-significant paths. Solid lines represent statistically significant paths.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Figure 2. PLS-based path modeling of faultlines and conflict predicting GOCB and group performance[a]. [a]Dotted lines represent non-significant paths. Solid lines represent statistically significant paths. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

$p < 0.05$), indicating the possibility that alignment of tenure and race (e.g., more-experienced Whites and less-experienced Asians) decreases relationship conflict instead of increasing it. Of the three task-related faultlines, tenure–gender faultlines were a significant positive predictor of task conflict ($\beta = 0.25$, $p < 0.001$), providing partial support for Hypothesis 1b.

When both types of conflict were simultaneously entered to predict GOCB, task and relationship conflict exhibited *opposite* effects on GOCB, each having positive and negative effects ($\beta = 0.23$, $p < 0.05$ and $\beta = -0.42$, $p < 0.001$, respectively). Thus, the results supported Hypotheses 2a but disconfirmed Hypothesis 3a. This counterintuitive pattern will be discussed later. With regard to team performance, task conflict was a significant negative predictor ($\beta = -0.40$, $p < 0.001$), supporting Hypothesis 3b. Relationship conflict, however, was not a significant predictor of team performance after controlling for its indirect effect via GOCB. Reinforcing prior findings (Podsakoff et al., 2009), GOCB showed a significant, positive association with a team's task performance ($\beta = 0.39$, $p < 0.01$).

Indirect effects

Hypothesis 4 predicted that the effects of faultlines on GOCB and group performance would be mediated by task and relationship conflict. As described above, we tested the possibility that faultlines had direct effects on GOCB and group performance in addition to the indirect, mediated effects through the two types of conflict. Given the low effect sizes (f^2) for those direct effects, the present data provide evidence for full mediation, supporting Hypothesis 4.

DOI: 10.1002/job

Table 2. Indirect effects of faultlines and conflict on group outcome

Dependent variable	Independent variable	Indirect effect	Standard error	*t*-value
Group-Level OCB	Gender–Age faultlines	−0.07	0.04	1.79
	Gender–Race faultlines	−0.01	0.04	0.11
	Age–Race faultlines	−0.07	0.04	1.71
	Tenure–Gender faultlines	0.01	0.04	0.10
	Tenure–Age faultlines	−0.08	0.05	1.59
	Tenure–Race faultlines	0.12**	0.04	2.62
Group performance	Gender–Age faultlines	−0.04	0.02	1.81
	Gender–Race faultlines	−0.01	0.03	0.11
	Age–Race faultlines	−0.05	0.03	1.45
	Tenure–Gender faultlines	−0.11*	0.05	2.25
	Tenure–Age faultlines	−0.11*	0.06	1.96
	Tenure–Race faultlines	0.02	0.06	.30
	Relationship conflict	−0.16***	0.05	3.35
	Task conflict	0.09*	0.04	1.97

*$p < .05$; **$p < .01$; ***$p < .001$.

Various approaches for testing the significance of mediation and indirect effects have been proposed (for review, see MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). Among them, MacKinnon et al.'s simulation study (2002) showed that approaches using the product of coefficients perform better than other methods, resulting in the greatest power and the most accurate Type 1 error. Recently, Preacher et al. (2007) advocated bootstrapping based on resampling as a robust strategy for assessing indirect effects. Table 2 reports the indirect effects computed using the product of coefficients generated through the bootstrapping method (300 iterations). The results indicated that only task-related faultlines possessed significant indirect effects on group outcomes via task and relationship conflict: (a) GOCB was positively related to tenure–race faultlines ($\beta = 0.12$, $p < 0.01$), and (b) group performance was negatively related to tenure–gender and tenure–age faultlines (both $\beta = -0.11$, $p < 0.05$). In addition, task and relationship conflict exhibited significant indirect effects on group performance via GOCB, although their effect directions were opposite ($\beta = 0.09$, $p < 0.05$ and $\beta = -0.16$, $p < 0.001$, respectively).

Discussion

Reflecting the increasing need for group-level citizenship behavior in the current team-based organizational context (Ng & Van Dyne, 2005), we conducted a systematic investigation of GOCB by theoretically identifying its antecedents. Our analysis indicated that demographic faultlines predict intermediate processes such as task and relationship conflict, which in turn affect GOCB and group performance. Unlike previous studies that examined faultlines in a laboratory setting or in groups with clearly preexisting subgroups (Homan et al., 2007a; Pearsall et al., 2008; Sawyer et al., 2006), the present study provides a field validation of the faultlines construct in a natural, organizational group setting, taking into account the effects of multiple faultlines simultaneously. Interestingly, task and relationship conflict showed *opposite* effects on GOCB. In this section, we will discuss the theoretical implications of the study, along with its limitations.

DOI: 10.1002/job

Faultlines based on relation-oriented and task-related attributes

This study presents empirical evidence regarding the differential roles of faultlines based on relation-oriented and task-related attributes in natural work groups. Diversity researchers have argued that less-job related attributes are prone to creating impermeable social categories and inter-category clashes, thus exasperating emotional and relational conflict among members (Jackson et al., 1995). Jehn et al.'s (1999) empirical study showed that social category diversity in gender and age is positively related to relationship conflict but not related to task conflict. The present results are consistent with prior theories and findings. Specifically, faultlines originating from social categories (gender–age and age–race faultlines) increased relationship conflict but their contribution to task conflict was not significant.

Compared with other visible social categories, members' tenure is a more task-related attribute that reflects distinct work experiences and histories with the organization, and thus likely to be a source of different task-related perspectives and information (Jackson et al., 1995). Confirming this expectation, tenure-based faultlines showed positive associations with task conflict (see Figure 2). Particularly, tenure–gender faultlines exhibited a highly significant positive effect on task conflict, indicating that task conflict is exacerbated when tenure-based subgroups are meaningfully differentiated by gender (e.g., female senior members vs. male junior members).

Unexpectedly, tenure–age faultlines, which is based on a task-related attribute, also increased relationship conflict. This finding is consistent with Homan et al.'s (2007a) study on informational diversity (i.e., differences in knowledge and ideas), which revealed that the convergence of informational diversity and faultlines based on gender and personality induces both task and relationship conflict. In the present study, the convergence of differences in knowledge and perspectives based on tenure and the age levels of members within each tenure subgroup seems to have increased relationship conflict. These findings suggest that task-related faultlines characterized by informational diversity is apt to activate subgroup categorization processes that are further promoted by social-category diversity dimensions such as gender, age, and race (i.e., relationship-oriented faultlines), resulting in increased interpersonal conflict (Homan et al., 2007b). Prior studies have indicated that these counterproductive social categorization processes can be thwarted by cross-cutting informational diversity and faultlines (Homan et al., 2007a) or promoting beliefs, norms, or climate that value diversity (Homan et al., 2007b). Further studies should reveal dynamic interactions between task-related faultlines and relationship-oriented faultlines.

Interestingly, all three predictors of relationship conflict (gender–age, age–race, and tenure–age faultlines) were driven by age. Strong age-based faultlines may divide group members into age-based subgroups and hinder the development of cohesion and identification with the entire group (Jehn et al., 1999). This detrimental effect seems particularly strong when age-based subgroup members also share other characteristics (high within-subgroup similarity) and do not have much overlap with members of other age-subgroup (low cross-subgroup similarity) (Shaw, 2004). The disparity in preferences, values, and styles resulting from age differences could become highly salient when the dividing lines are amplified by other member attributes.

As we hypothesized, task and relationship conflict mediated the effects of faultlines on group outcomes. Our bootstrap analysis indicated that all three faultlines involving tenure had significant indirect effects on group outcomes. Tenure–gender and tenure–age faultlines exerted significant negative indirect effects on group performance. Interestingly, although it was not hypothesized, tenure–race faultlines had a significant, positive indirect effect on GOCB. This is consistent with prior studies indicating the significance of tenure in organizational settings (Ancona & Caldwell, 1992; Zenger & Lawrence, 1989), as well as a recent meta-analysis showing that task-related diversity has greater implications for group performance than do relation-oriented social categories (Horwitz & Horwitz, 2007).

DOI: 10.1002/job

All in all, our results indicate that the presence of strong faultlines may harm the group by begetting negative interpersonal dynamics and increasing both task and relationship conflict (Homan et al., 2007a; Sawyer et al., 2006). The present analysis using faultlines provides a more sophisticated understanding of the dynamics generated by membership diversity. If we had used the traditional measures of diversity based on dispersion or Euclidean distance (e.g., Choi, in press), the conclusion could be quite simple: Gender diversity increases both task and relationship conflict, whereas race and tenure diversity decreases relationship conflict. The present findings, as shown in Figure 2, offers a more refined interpretation of the roles associated with different aspects of diversity when they are considered in combination, although the overall pattern of results may be more complex than we have hypothesized. In this regard, prior studies have indicated that to fully understand the effects of diversity on group processes and outcomes, it is important to consider moderating or contingency variables. For example, Mohammed and Angell (2004) showed that a strong group orientation among members neutralizes the negative effect of gender diversity on relationship conflict. Recently, Van der Vegt, Van de Vliert, and Huang (2005) also reported that the relationship between tenure diversity and innovative climate is positive in a low power-distance culture, but it is negative in a high power-distance culture. Future studies should examine moderators and boundary conditions that shape the effects of various faultline constructs on group processes and outcomes.

Distinct roles of task and relationship conflict

Our analysis revealed that task and relationship conflict were positive and negative predictors of GOCB, respectively. In line with Jehn's (1995) proposition, scholars and practitioners alike have endorsed the belief that task and relationship conflict respectively influence performance positively and negatively. Nevertheless, De Dreu and Weingart's (2003) meta-analysis showed that *both* types of conflict are negative predictors of group performance and member satisfaction. The results of our study provide potential explanations for this discrepancy. That is, while the direct effects of task and relationship conflict on performance are negative in accordance with the recent meta-analytic findings (De Dreu & Weingart, 2003) and the pattern observed in the present study, they may exert disparate indirect effects on group performance by shaping members' behaviors, which supports Jehn's (1995) framework. As shown in Figure 2, although both types of conflict had negative effects on group performance, they contributed differently to GOCB. Also, as reported in Table 2, task and relationship conflict had positive and negative indirect effects on group performance.

Task and relationship conflict may initiate different group-level affective and cognitive processes that lead to distinct behavioral outcomes for the group. In the case of relationship conflict, avoidance is the only way to achieve effective group functioning (De Dreu & Van Vianen, 2001) perhaps because it would be quite difficult to resolve interpersonal differences caused by personal identity and values. In contrast, ample evidence suggest that task-related conflict can be effectively dealt with via information sharing and collaborative strategies adopted by group members (De Dreu, 2006). By allowing different perspectives to emerge and to be shared among members, task conflict increases members' understanding and satisfaction with group decisions (Jehn, 1995; Tekleab et al., 2009). In addition, the presence of different ideas, information, and viewpoints expressed by members may indicate group members' level of energy, involvement, and enthusiasm for the task. In fact, studies have shown that when conflict remains task-focused, group members tend to experience more commitment to the group and trust towards each other (Tekleab et al., 2009; Tjosvold, Dann, & Wong, 1992). For example, Hempel et al. (2009) showed that when a team adopts a cooperative conflict management strategy, it tends to exhibit strong affect-based trust, which has positive implications for performance. Strong task-orientation and high group commitment (perhaps task-based) among members may induce voluntary

DOI: 10.1002/job

engagement with group task activities, which should increase the overall level of voluntary contribution and helping within the group.

Given the unexpected positive link between task conflict and GOCB and the strong correlation between task and relationship conflict, it is plausible to attribute the unexpected finding to the suppression caused by multicollinearity among predictors. In the present data, however, when both task and relationship conflict were entered into the equation predicting GOCB, the variance inflation factor value was only 1.52 (much lower than the typical cutoff of 10), indicating that multicollinearity was not a serious threat to the validity of the present finding. In addition, the partial correlation between task conflict and GOCB after controlling for relationship conflict was 0.19, supporting the pattern reported in Figure 2.

Perhaps, after controlling for the negative effect of relationship conflict and its dysfunctional consequences (Ng & Van Dyne, 2005), group members may interpret the presence of task conflict as an indication of others' involvement and enthusiasm for the task at hand. This positive assessment regarding other members' task-related attitudes and motivation may reassure members that their discretionary behaviors and voluntary contributions are likely to be acknowledged and reciprocated (*cf.* social exchange theory, Blau, 1964). Task conflict may also increase GOCB perhaps owing to an increased understanding of other members' views and appreciation of differences (De Dreu & Van Vianen, 2001; Jehn, 1995). All these speculations, however, should be further validated by additional empirical research that examines emotional, motivational, and cognitive reactions to task and relationship conflict. As suggested by recent studies (Hempel et al., 2009; Tekleab et al., 2009), it is critical to understand how conflict is managed by the group rather than how much conflict is present in it.

Study limitations

This study has several limitations. First, some of the measures used in this study may have limited validity. Specifically, tenure was the only attribute used to assess the task-related faultlines, which could have been expanded using other characteristics such as function, expertise, and education (Jackson et al., 1995; Horwitz & Horwitz, 2007). In addition, group performance was based on the leader's assessment instead of more objective forms of assessment (e.g., output quantity, error rate, customer satisfaction). Future research should expand the current framework by employing faultlines based on additional attributes and objective group performance measures.

Second, while the selection of variables and specified relationships were informed by theory and empirical findings, the causality of relationships was not definitive, given the cross-sectional nature of the present research design. However, our hypothesized direction of causality is consistent with theory (e.g., De Dreu & van Vianen, 2001; Ng & Van Dyne, 2005). We join other researchers (Organ et al., 2006; Podsakoff et al., 2009) in the call for research to obtain better evidence on the direction of causality between GOCB and its antecedents and consequences.

A third limitation of this study relates to our data collection strategy. In this study, we collected data from a variety of companies in diverse industries. This design limited our ability to control for certain extraneous variables (e.g., organizational culture, group structure, task characteristics, etc.), and thus does not take into account potential confounding variables that might have influenced the study findings. However, the sampling of a variety of companies and industries could also be considered a strength in that it increases the generalizability of the study findings across various settings. In addition, given that studies on faultlines and GOCB are still in the early stage of development, it is important to broaden the conceptual and empirical scope and investigate additional mediators (e.g., trust, integration, and team commitment) and moderators or boundary conditions for the proposed relationship (e.g., leader behavior, average group tenure, and task interdependence).

DOI: 10.1002/job

Conclusion

Given that OCB's contribution to group and organizational performance is based on employees' *collective* engagement in such behaviors (Organ et al., 2006), it is critical that we understand the mechanisms through which (or the conditions under which) a group of individuals collectively exhibits high levels of OCB in organizational settings (Choi, in press). The present study answers the call to examine the relationships among GOCB and its antecedents and consequences within a coherent framework (Mathieu et al., 2008; Podsakoff et al., 2000). Moreover, as one of the few empirical studies using field data to assess the effect of faultlines in natural work groups, this study demonstrates that the nature and effects of demographic faultlines largely depend on the specific attributes that are activated to create subgroups in organizational groups. Given that faultlines based on different attributes have different implications for group processes and outcomes, it is critical that researchers separately examine faultlines based on different attributes, instead of developing a single, collapsed measure (*cf*. Guideline 7, Harrison & Klein, 2007). Finally, this study highlights the need to elaborate further on the underlying mechanisms that shape the effects of task and relationship conflict on behavioral and performance outcomes of work groups. Our study demonstrates that diversity should not be viewed as a monolithic concept with homogenous implications; rather, diversity should be examined from diverse perspectives with potentially differential underlying mechanisms and consequences.

Author biographies

Jin Nam Choi is Associate Professor of Organizational Behavior at the Graduate School of Business, Seoul National University, Korea. He earned his Ph.D. in Organizational Psychology from the University of Michigan. He is currently serving as an Associate Editor for *Journal of Occupational and Organizational Psychology.* His research interests include team processes and effectiveness, multilevel dynamics involving innovation and creativity, and organizational citizenship behavior.

Thomas Sy is Assistant Professor at the Department of Psychology, University of California, Riverside. He earned his Ph.D. in Organizational Psychology from the University of Michigan. His research interests include the role of affect in organizations, leadership within different cultural contexts, and group processes and performance.

References

Allen, N. J., Stanley, D. J., Williams, H. M., & Ross, S. J. (2007). Assessing the impact of nonresponse on group diversity effects. *Organizational Research Methods*, *10*, 262–286.

Ancona, D. G., & Caldwell, D. F. (1992). Demography and design: Predictors of new product team performance. *Organization Science*, *3*, 321–341.

Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein, & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

DOI: 10.1002/job

Bommer, W. H., Dierdorff, E. C., & Rubin, R. S. (2007). Does prevalence mitigate relevance? The moderating effect of group-level OCB on employee performance. *Academy of Management Journal*, *50*, 1481–1494.

Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance*, *10*, 99–109.

Brodbeck, F. C., Kerschreiter, R., Mojzisch, A., & Schulz-Hardt, S. (2007). Group decision making under conditions of distributed knowledge: The information asymmetries model. *Academy of Management Review*, *32*, 459–479.

Byrne, D. E. (1971). *The attraction paradigm*. New York: Academic Press.

Cassel, C., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, *26*, 435–446.

Chen, G., Mathieu, J. E., & Bliese, P. D. (2004). A framework for conducting multi-level construct validation. In F. J. Yammarino, & F. Danereau (Eds.), *Multi-level issues in organizational behavior and processes* (pp. 273–303). Oxford, UK: Elsevier.

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Mahwah, NJ: Lawrence Erlbaum Associates.

Chin, W. W., & Newsted, P. R. (1999). Structural equation modeling analysis with small samples using partial least squares. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 307–341). Thousand Oaks, CA: Sage.

Choi, J. N. (in press) Group-level dynamics of organizational citizenship behavior: What group characteristics promote citizenship in groups? *Journal of Management Studies.*

Choi, J. N., & Kim, M. U. (1999). The organizational application of groupthink and its limitations in organizations. *Journal of Applied Psychology*, *84*, 297–306.

Chung, Y., Shaw, J. B., & Jackson, S. E. (2006). *Faultline Index for groups, version 1.0: A guide for SAS users*. Piscataway, NJ: Rutgers University.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

De Dreu, C. K. W., & van Vianen, A. E. M. (2001). Managing relationship conflict and the effectiveness of organizational teams. *Journal of Organizational Behavior*, *22*, 309–328.

De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, *88*, 741–749.

Ehrhart, M. G. (2004). Leadership and procedural justice climate as antecedents of unit-level organizational citizenship behavior. *Personnel Psychology*, *57*, 61–94.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*, 39–50.

Gibson, C., & Vermeulen, F. (2003). A healthy divide: Subgroups as a stimulus for team learning behavior. *Administrative Science Quarterly*, *48*, 202–239.

Guzzo, R. A., & Shea, G. P. (1991). Group performance and intergroup relations in organizations. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 269–313). Palo Alto, CA: Consulting Psychologists Press.

Hackman, J. R. (1992). Group influences on individuals in organizations. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 199–267). Palo Alto, CA: Consulting Psychologists Press.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*, 1199–1228.

Hempel, P. S., Zhang, Z.-X., & Tjosvold, D. (2009). Conflict management between and within teams for trusting relationships and performance in China. *Journal of Organizational Behavior*, *30*, 41–65.

Homan, A. C., Hollenbeck, J. R., Humphrey, S. E., van Knippenberg, D., Ilgen, D. R., & van Kleef, G. A. (2008). Facing differences with an open mind: Openness to experience, salience of intragroup differences, and performance of diverse work groups. *Academy of Management Journal*, *51*, 1204–1222.

Homan, A. C., van Knippenberg, D., van Kleef, G. A., & De Dreu, C. K. W. (2007a). Interacting dimensions of diversity: Cross-categorization and the functioning of diverse work groups. *Group Dynamics*, *11*, 79–94.

Homan, A. C., van Knippenberg, D., van Kleef, G. A., & De Dreu, C. K. W. (2007b). Bridging faultlines by valuing diversity: Diversity beliefs, information elaboration, and performance in diverse work groups. *Journal of Applied Psychology*, *92*, 1189–1199.

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management*, *33*, 987–1015.

DOI: 10.1002/job

House, R., Rousseau, D. M., & Thomas-Hunt, M. (1995). The meso paradigm: A framework for integration of micro and macro organizational theories. In L. L. Cummings, & B. Staw (Eds.), *Research in organizational behavior* (Vol. 17, pp. 71–114). Greenwich, CT: JAI.

Jackson, S. E., May, K. E., & Whitney, K. (1995). Understanding the dynamics of diversity in decision-making teams. In R. A. Guzzo, & E. Salas Associates (Eds.), *Team effectiveness and decision making in organizations* (pp. 204–261). San Francisco: Jossey-Bass.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*, 85–98.

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, *40*, 256–282.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly*, *44*, 741–763.

Lau, D. C., & Lam, L. W. (2008). Effects of trusting and being trusted on team citizenship behaviours in chain stores. *Asian Journal of Social Psychology*, *11*, 141–149.

Lau, D. C., & Murnighan, J. K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of Management Review*, *23*, 325–340.

Lau, D. C., & Murnighan, J. K. (2005). Interactions within groups and subgroups: The effects of demographic faultlines. *Academy of Management Journal*, *48*, 645–659.

Lincoln, J. R., & Miller, J. (1979). Work and friendship ties in organizations: A comparative analysis of relational networks. *Administrative Science Quarterly*, *24*, 181–199.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, *114*, 184–199.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83–104.

Maddux, W. W., & Brewer, M. B. (2005). Group gender differences in the relational and collective bases for trust. *Processes & Intergroup Relations*, *8*, 159–171.

Marcoulides, G. A., & Saunders, C. (2006). Editor's comments: PLS: A silver bullet? *MIS Quarterly*, *30*, iii–ix.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*, 410–476.

Mayer, D. M., Kuenzi, M., Greenbaum, R., Bardes, M., & Salvador, R. (2009). How low does ethical leadership flow? Test of a trickle-down model. *Organizational Behavior and Human Decision Processes*, *108*, 1–13.

Mohammed, S., & Angell, L. C. (2004). Surface- and deep-level diversity in workgroups: Examining the moderating effects of team orientation and team process on relationship conflict. *Journal of Organizational Behavior*, *25*, 1015–1039.

Mooney, A. C., Holahan, P. J., & Amason, A. C. (2007). Don't take it personally: Exploring cognitive conflict as a mediator of affective conflict. *Journal of Management Studies*, *44*, 733–758.

Ng, K. Y., & Van Dyne, L. (2005). Antecedents and performance consequences of helping behavior in work groups: A multilevel analysis. *Group & Organization Management*, *30*, 514–540.

Niehoff, B. P., & Moorman, R. H. (1993). Justice as a mediator of the relationship between methods of monitoring and organizational citizenship behavior. *Academy of Management Journal*, *36*, 527–556.

Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenships behavior. *Personnel Psychology*, *48*, 775–802.

Organ, D. W., Podsakoff, P. M., & MacKenzie, S. B. (2006). *Organizational citizenship behavior: Its nature, antecedents, and consequences*. Thousand Oaks, CA: Sage.

Pearsall, M. J., Ellis, A. P. J., & Evans, J. M. (2008). Unlocking the gender faultlines on team creativity: Is activation the key? *Journal of Applied Psychology*, *93*, 225–234.

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, *44*, 1–28.

Podsakoff, N. P., Whiting, S. W., Podsakoff, P. M., & Blume, B. D. (2009). Individual- and organizational-level consequences of organizational citizenship behavior: A meta-analysis. *Journal of Applied Psychology*, *94*, 122–141.

Podsakoff, P. M., Ahearne, M., & MacKenzie, S. B. (1997). Organizational citizenship behavior and the quantity and quality of work group performance. *Journal of Applied Psychology*, *82*, 262–270.

Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management*, *26*, 513–563.

DOI: 10.1002/job

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185–227.

Qureshi, I., & Compeau, D. (2009). Assessing between-group differences in information systems research: A comparison of covariance- and component-based SEM. *MIS Quarterly*, *33*, 197–214.

Richardson, H. A., & Vandenberg, R. J. (2005). 'Integrating managerial perceptions and transformational leadership into a work-unit level model of employee involvement'. *Journal of Organizational Behavior*, *26*, 561–589.

Ringle, C. M., Wende, S., & Will, A. (2005). *SmartPLS 2.0.* Hamburg: University of Hamburg.

Sawyer, J. E., Houlette, M. A., & Yeagley, E. L. (2006). Decision performance and diversity structure: Comparing faultlines in convergent, crosscut, and racially homogeneous groups. *Organizational Behavior and Human Decision Processes*, *99*, 1–15.

Schnake, M. E., & Dumler, M. P. (2003). Levels of measurement and analysis issues in organizational citizenship behaviour research. *Journal of Occupational and Organizational Psychology*, *76*, 283–301.

Shaw, J. B. (2004). The development and analysis of a measure of group faultlines. *Organizational Research Methods*, *7*, 66–100.

Somech, A., & Drach-Zahavy, A. (2004). Exploring organizational citizenship behavior from an organizational perspective: The relationship between organizational learning and organizational citizenship behavior. *Journal of Occupational and Organizational Psychology*, *77*, 281–298.

Sy, T., Cote, S., & Saavedra, R. (2005). The contagious leader: Impact of leader's affect on group member affect and group processes. *Journal of Applied Psychology*, *90*, 295–305.

Tajfel, H. (1978). The achievement of group differentiation. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 483–507). London: Academic Press.

Tekleab, A. G., Quigley, N. R., & Tesluk, P. E. (2009). A longitudinal study of team conflict, conflict management, cohesion, and team effectiveness. *Group & Organizational Management*, *34*, 170–205.

Thatcher, S. M. B., Jehn, K. A., & Zanutto, E. (2003). Cracks in diversity research: The effects of diversity faultlines on conflict and performance. *Group Decision and Negotiation*, *12*, 217–241.

Timmerman, T. A. (2005). Missing persons in the study of groups. *Journal of Organizational Behavior*, *26*, 21–36.

Tjosvold, D., Dann, V., & Wong, C. (1992). Managing conflict between departments to serve customers. *Human Relations*, *45*, 1035–1054.

Turner, J. C. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford, England: Basil Blackwell.

Turner, M. E., Pratkanis, A. R., Probasco, P., & Leve, C. (1992). Threat, cohesion, and group effectiveness: Testing a social identity maintenance perspective on groupthink. *Journal of Personality and Social Psychology*, *63*, 781–796.

Van der Vegt, G. S., Van de Vliert, E., & Huang, X. (2005). Location-level links between diversity and innovative climate depend on national power distance. *Academy of Management Journal*, *48*, 1171–1182.

van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, *58*, 515–541.

van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, *89*, 1008–1022.

Wayne, S. J., Shore, L. M., & Liden, R. C. (1997). Perceived organizational support and leader-member exchange: A social exchange perspective. *Academy of Management Journal*, *40*, 82–111.

Webber, S. S., & Donahue, L. M. (2001). Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of Management*, *27*, 141–162.

Williams, K. Y., & O'Reilly, C. A. III. (1998). Demography and diversity in organizations: A review of 40 years of research. In B. M. Staw, & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 20, pp. 77–140). Greenwich, CT: JAI Press.

Zenger, T., & Lawrence, B. (1989). Organizational demography: The differential effects of age and tenure distributions on technical communications. *Academy of Management Journal*, *32*, 353–376.

DOI: 10.1002/job

Appendix

The faultline strength scores used in the present study were computed using four demographic characteristics. For the sake of simplicity, below we provide an illustration of the computation procedure based on two factors (gender and age) using a fictitious group as an example. This illustration is based on an example from Shaw (2004).

Observed frequencies					
	Age category				
	20s	30s	40s	50 and older	Subgroup n
Males	$N_{m/20s} = 2$	$N_{m/30s} = 2$	$N_{m/40s} = 0$	$N_{m/50+} = 0$	$N_m = 4$
Females	$N_{f/20s} = 1$	$N_{f/30s} = 1$	$N_{f/40s} = 0$	$N_{f/50+} = 0$	$N_f = 2$

Note: Expected frequencies for males: $E_m = 1$; for females: $E_f = 0.5$.

1. Subgroup internal alignment (IA) of the male subgroup in terms of age is calculated using the formula shown below (see Shaw, 2004, p. 72):

$$IA_{m/age/obs} = \sum \frac{(O_{mi} - E_{mi})^2}{E_{mi}}$$

where $IA_{m/age/obs}$ is the observed male alignment index across age categories, O_{mi} is the observed number of males in the *i*th age category, and E_{mi} is the expected number of males in the *i*th age category assuming random distribution. In the above example,

$$IA_{m/age/obs} = [(2-1)^2]/1 + [(2-1)^2]/1 + [(0-1)^2]/1 + [(0-1)^2]/1 = 4.0$$

If there was perfect alignment of males across age categories, then,

$$IA_{m/age/perfect} = [(4-1)^2]/1 + [(0-1)^2]/1 + [(0-1)^2]/1 + [(0-1)^2]/1 = 12.0$$

If there was total nonalignment of males across age categories, then,

$$IA_{m/age/nonaligh} = [(1-1)^2]/1 + [(1-1)^2]/1 + [(1-1)^2]/1 + [(1-1)^2]/1 = 0.0$$

To adjust for the variation in $IA_{m/age/perfect}$ and $IA_{m/age/nonalign}$ due to differences in number of categories and subgroup sample sizes, the final IA index for males across age categories is calculated as,

$$IA_{males/age} = \frac{(IA_{m/age/obs} - IA_{m/age/nonalign})}{IA_{m/age/perfect}} = \frac{(4-0)}{12} = 0.333$$

2. In the above example, the cross-subgroup alignment index (CGAI) can be calculated by identifying the frequency of ''match-ups'' between subgroup members in each age category (Shaw, 2004, pp. 74–76):

$$CGAI_{gender/age} = [(N_{m/20s} \times N_{f/20s}) + (N_{m/30s} \times N_{f/30s}) + (N_{m/40s} \times N_{f/40s}) + (N_{m/50} + \times N_{f/50+})] / (N_m \times N_f)$$
$$= [(2 \times 1) + (2 \times 1) + (0 \times 0) + (0 \times 0)]/(4 \times 2) = [2 + 2 + 0 + 0]/8 = 4/8 = 0.5$$

DOI: 10.1002/job

3. Subgroup internal alignment (IA) and cross-subgroup alignment (CGAI) indices can be combined in the following formula to calculate the faultline strength score (FLS):

$$\text{FLS} = \text{IA} \times (1 - \text{CGA}) = 0.33 \times (1 - 0.5) = 0.165$$

Thus, in the above example, the strength of gender-based faultlines, taking into account the age distribution within the group, is 0.165.

DOI: 10.1002/job