Title
Rating performance or contesting status: Evidence against the homophily explanation for supervisor demographic skew in performance ratings

Permalink
https://escholarship.org/uc/item/1tc9t7wx

Journal
Organization Science, 23(2)

ISSN
1047-7039

Authors
Pearce, JL
Xu, QJ

Publication Date
2012-03-01

DOI
10.1287/orsc.1100.0585

License
https://creativecommons.org/licenses/by/4.0/ 4.0

Peer reviewed
Rating Performance or Contesting Status: Evidence Against the Homophily Explanation for Supervisor Demographic Skew in Performance Ratings

Jone L. Pearce
The Paul Merage School of Business, University of California, Irvine, Irvine, California 92697, jlpearce@uci.edu

Qiumei Jane Xu
College of Business and Management, Northeastern Illinois University, Chicago, Illinois 60625, q-xu@neiu.edu

We propose and test an argument in which the well-documented skew in supervisory performance appraisal ratings toward those with the same demography as themselves is better explained by the status contests than the reigning theory of homophily. We conduct the test in a field study of 358 supervisor–subordinate dyads in 10 organizations, using hierarchical linear modeling with various controls. We find that supervisors’ ratings of subordinates’ contextual and task performance only skew toward similar subordinates when supervisors’ status is contested by a higher demographic status subordinate, as predicted by social dominance and status characteristics theories. None of the general homophily preference hypotheses is supported. This study provides a richer theory more consistent with the accumulating evidence about demography effects in organizations and demonstrates the value of head-to-head strong inference tests and status explanations for the field of organizational behavior.

Key words: status; diversity; demography; performance appraisal

Introduction
Supervisors’ ratings of their subordinates’ job performance are as pervasive as are concerns that they imperfectly reflect job performance, with a longstanding interest in the particular problems this poses for diverse workplaces (Roberson et al. 2007). As workplaces evolve, biases in employees’ performance assessments based on demographic differences become an increasingly serious matter. Performance ratings that are seen to reflect employees’ gender, race, age, or disability rather than their job performance undermine employees’ sense of fairness (Folger et al. 1992, Kanfer et al. 1987) and their trust in their organizations (Mayer and Davis 1999). They can also lead to rewards and promotions going to less able people (Folger et al. 1992). The best-documented and most problematic ratings inaccuracy is supervisory demographic similarity skew: the more similar subordinates are to their supervisors, the higher supervisors rate those subordinates’ job performance (Pulakos and Wexley 1983, Bauer and Baltes 2002). The most prominent explanation for this well-documented ratings skew is homophily bias. That is, supervisors exhibit favoritism to those similar to themselves. However, evidence inconsistent with homophily explanations of this demographic ratings skew continues to mount. We draw on theories of status defense to propose an alternative theory of this skew in supervisory ratings of subordinate performance and then test it against the reigning homophily account. In a test in 10 different organizations, we find that contests for relative status, not homophily, best explain these persistent demographic effects in supervisors’ ratings of their subordinates’ job performance.

Supervisors’ Ratings of Performance
Supervisors’ ratings of their subordinates’ job performance are the primary way employees’ job performance is measured for the vast majority of jobs in most organizations (Landy and Farr 1980). This reliance on supervisors’ ratings continues despite debate about their inaccuracy (Roberson et al. 2007). Supervisory ratings of performance persist because, except under rare circumstances, the alternatives are worse. There are few jobs with comprehensive objective measures of performance because the employment contract itself is based on a need for flexibility in the face of uncertainty about future demands (Murphy and Cleveland 1995). For most jobs, most of the time, employees’ performance needs to be assessed via retrospective assessments of job effort, citizenship, judgment, and performance after the fact (Pearce 1987, 2009).

There is strong evidence that supervisory ratings inaccuracy is not random but is instead the result of systematic distortions. The best-documented and most problematic of these distortions is demographic homophily skew: the more similar subordinates are to their supervisors,
the higher the subordinates’ relative performance ratings (Pulakos and Wexley 1983, Bauer and Baltes 2002). In the literature, this is most often called similarity bias; however, the word “bias” carries complex connotations, so we prefer to use the term “skew”—making it clear that our focus is on the empirical observation that supervisors tend to rate those more demographically similar to themselves more positively than those who are dissimilar (Roberson et al. 2007, Tsui and O’Reilly 1989).

Such demographic skew in supervisors’ evaluations of their subordinates’ performance is not merely a symbolic matter but has widespread effects on organizational practices and employee opportunities. For example, Wakabayashi et al. (1988) reported that subordinates more similar to their supervisors received relatively more frequent promotions; Somech’s (2003) research established that demographic differences between supervisors and subordinates were negatively associated with supervisor’s inclusion of subordinates in participative decision making; McCain et al. (1983) and Wagner et al. (1984) found that subordinates with demographically dissimilar supervisors were more likely to leave their jobs; and Castilla (2008) reported that women and minorities received lower merit pay increases from their white male supervisors. This demographic skew is widely documented and has pernicious effects on employees and organizational performance, yet, as DiTomaso et al. (2007a) decry, too many scholars have simply assumed that it is explained by supervisory homophily.

**Homophily and Ratings Skew**

The dominant theoretical explanation for demographic skew in supervisory ratings has been the similarity—attraction, or homophily, explanation (Byrne 1971, Glaman et al. 1996). There is a large body of research demonstrating that people tend to be more attracted to, and prefer to be with, those who are more similar to themselves. Homophily can invoke an attraction dynamic whereby demographically similar individuals accentuate the positive attributes of each other and derive positive social identities from identifying with similar others. This dynamic also is referenced in social identity, self-categorization, and related theories (e.g., Messick and Mackie 1989, Tajfel and Turner 1986, Turner 1987). Although all these theories are complex and have been used to explain a wide variety of social phenomena (see Hogg and Ridgeway 2003), all these explanations of demographic skew in organizational behavior assume that homophily drives the observed demographic skew in evaluations. Therefore, we shall use “homophily” to refer to all of these theoretical explanations.

Homophily has been used as a retrospective explanation for a wide range of observed demographic effects in organizational settings (e.g., Chatman and Sparato 2005, Ibarra 1995, Pelled et al. 1999, Tsui et al. 1992). Focusing on supervisors’ demographic skew in rating their subordinates, Tsui et al. (2002) argued that demographic differences between supervisors and subordinates can undermine supervisor–subordinate interpersonal attraction, supervisor perceptions of employee performance, and subordinate perceptions of managerial supportiveness.

**Status and Ratings Skew**

Here, we propose that it is not homophily but status contests that best account for demographic skew in supervisory ratings. The study of status dates back to the earliest days of sociology (Simmel 1950, Weber 1978) and social psychology (e.g., Blau 1977, 1994; Harvey and Consalvi 1960), but it has had a relatively minor role in organizational behavior, most often treated as a variable of secondary interest (Pearce 2010). There have been numerous definitional debates over the meaning of status and how it might differ from such ideas as estates (Weber 1978), class (Marx 1967), and power (Pearce et al. 2001), as well as the debate about status as a subjective evaluation or an objective structural reality (Wegener 1992). Building on the Pearce et al. (2001) review of the varied definitions of status, we adopt their definition of status as one’s relative respected social standing with reference to a particular social grouping or hierarchy.

There is no question that having comparatively high status is a source of many advantages in organizations. To cite just a few examples, those with higher status are listened to more, receive more deference from others, and are perceived as having more power (Berger et al. 1966, Sheets and Braver 1995). People with higher status have better health and longevity (Marmot 2004), and their decisions have more legitimacy (Keashly and Newberry 1995). Those randomly assigned a higher status outperform those randomly assigned a lower status (Lovaglia et al. 1998). Organizations with higher status were able to command higher prices and obtain other advantages in their markets (Podolny 2005), and organizations lose their business partners when they lose status (Jensen 2006). Because of its advantages, higher status usually is actively sought in social settings (Harvey and Consalvi 1960). Having more money (Nee 1996), displaying status objects (Sundstrom and Sundstrom 1986), having more education (Bidwell and Friedkin 1988), working in more prestigious occupations (Kanekar et al. 1989), and belonging to elite organizations (D’Aveni and Kesner 1993) all lead to greater status, and so status seeking is seen as a primary motive for pursuing money, status objects, education, certain occupations, and membership in elite voluntary associations.

Status did receive attention in early organizational behavior theory (e.g., Barnard 1968, Vroom 1964), yet only recently have scholars in organizational behavior begun to focus their attention on the role of status
and status striving in organizational behavior, focusing on discrimination (Umphress et al. 2007), within-group conflict (Phillips and Thomas-Hunt 2007), workplace harassment (Berdahl 2007), employee satisfaction (Chen et al. 2003, Elsbach and Kramer 1996), and organizational learning (Perretti and Negro 2006), among others. Here, we draw on two theories of status, social dominance theory (SDT) and status characteristics theory (SCT), to propose and test an argument that status contests, not homophily, better explain demographic similarity skew in supervisory ratings.

SDT focuses on how individual and institutional features combine to contribute to racial and gender oppression and is prominent in sociology and political science. This theory seeks to explain how processes at different levels interact, and it includes consideration of cultural, ideological, political, and societal structures as well as individual psychological and social-psychological processes (Sidanius et al. 2004). Although primarily a theory of power dynamics, it makes arguments that we can generalize to help develop status-based explanations of supervisory performance ratings skew.

According to SDT, those who are dominant, for example, in higher status groups, are motivated to subjugate less-dominant others to maintain their dominant or higher status positions, whereas members of other groups often provide willing support for their own subjugation. The theory proposes, and Sidanius et al. (1994) have found, that those in nondominant groups are consistently less homophilous than are those in dominant groups. Additional support found by Hinkle and Brown (1990) suggests that lower-status individuals did not favor similar others, and Levin (1992) reported that some low-status individuals favored higher-status others rather than similar others. In organizational research, Umphress et al. (2007) found that homophily preference was indeed more common among those in high-status groups, especially among those individuals with high social dominance orientation (see also Shao et al. 2009). Umphress et al. (2007) found that students from a nondominant group (women) were not more attracted to a prospective organization if it had more women, but men were more attracted to male-dominated groups.

Status characteristics theory also provides a nonhomophily explanation for why supervisors might skew their ratings of their subordinates’ job performance. Berger et al. (1972) proposed that observable characteristics such as gender, age, and race are invested with social value that generates a status hierarchy with those of higher status expected to be more competent. Furthermore, Cohen and Silver (1989) suggested that those whom working together are guided by two motives—contributing to collective performance and avoiding status loss. Building on earlier SCT ideas, Troyer and Younts (1997) argued that those who perceive themselves to have high status will become angry when confronted with a competing claim for high status. That is, those who perceive themselves to have high status are motivated to avoid a status loss, with its implications for their perceived competence. As Blader and Chen (2011) argue, those with higher relative status have a heightened sensitivity to whether those with lower status acknowledge, validate, and verify their high-status positions. This suggests that those supervisors not occupying high-status demographic groups might be highly sensitive to the relative status loss implied by having subordinates in high-status demographic groups and may be motivated to defend their relative standing.

Generalizing from these theories of status contest can help to account for the growing empirical evidence contradicting the homophily explanation. Tsui et al. (1992) found that men (higher status) responded more negatively than did women (lower status) to working in groups dominated by the other gender. However, they only note this empirical finding in passing and do not use it to question their homophily explanation. Although not citing status theories, Chatman and O’Reilly (2004) intuited that because men traditionally have occupied a higher status, men would be more sensitive to the status incongruence of working in female-dominated work groups. They found that men were more likely to want to transfer out of work groups with greater proportions of women. However, contrary to their predictions but consistent with SDT, lower-status women did not consistently favor either female-dominated or male-dominated work groups. Chattopadhyay and his colleagues (Chattopadhyay 1999, Chattopadhyay et al. 2008) proposed that those with higher status would react more negatively to being in low status-dominated work groups, whereas those with lower status would not react negatively; however, their empirical support was mixed. Likewise, Vecchio (1993) tested a status contest explanation for supervisor age skew in rating supervisors’ subordinates against homophily, social competition, and loyalty explanations. Unfortunately, his mixed results did not consistently support any of these explanations. Finally, Perry et al. (1999) conducted a direct test comparing the homophily and status incongruence explanations for demographic skew, but with mixed results: although the homophily explanations were not supported, neither were most of their status contest explanations. Unfortunately, both Vecchio (1993) and Perry et al. (1999) had small samples of low statistical power.

These empirical studies all demonstrate that homophily was not the best predictor of supervisory rating demographic skew in organizational settings, but they do not seem to have successfully unseated the attachment to homophily explanations for demographic skew in organizational behavior. For example, recent reviews by Roberson et al. (2007) and Williams and O’Reilly (1998) still rely on homophily explanations for demography-based biases in supervisory ratings. We believe this failure of contradictory data to shift
attachment to the preferred homophily explanation persists because there has been an absence of a strong competing theory that could better account for these conflicting empirical results (Kuhn 1970).

Task and Contextual Performance. Job performance has been split into two distinct categories: task performance and contextual performance (Motowidlo and Van Scotter 1994). Task performance is how well employees perform their assigned jobs; this is usually formally assessed and for many jobs can include objective indicators such as sales made, project deadlines met, and the like. Contextual performance comprises those non-required voluntary employee actions that facilitate team and organizational goal attainment. Examples include working hard to meet a deadline and taking the initiative to solve unexpected problems (Borman and Motowidlo 1993). Task and contextual performance are important for different reasons and so are analyzed separately here. Task performance is more important for organizational rewards but is likely to be constrained by objective measures of performance, providing limited opportunities for supervisory discretion; when rating contextual performance, supervisors have more scope to act to assert their status, and contextual performance has been found to be important for promotions.

Hierarchical vs. Demographic Status. There is substantial evidence that demography sends status signals—the status implications of differing ages, genders, races, and occupations all have been documented (Roberson et al. 2007). In assessing supervisors’ ratings of their subordinate employees’ job performance in field settings, we need to address the question of which status groups will be salient to supervisors. Because everyone at work occupies multiple status positions (Blau 1977), we need to propose which group, feature, or characteristic will form the basis for any individual’s status contestation. Social identity theory proposes that people will attend most to the parameter that best supports an attractive self-identity (Tajfel and Turner 1986), but this provides little concrete guidance because organizations are complex social settings with different parameters providing different potential identity-enhancing attractions. Therefore, we draw first on Blau (1977), who proposed that ascribed group parameters that do not allow for mobility (such as demography) will be more salient than more fluid achieved characteristics (such as formal organizational rank).

Organizations establish their own internal formal organizational status hierarchies and usually expend considerable effort directing attention to them. Organizational hierarchical rank is ascriptive but is assumed to reflect status rankings in most organizations. This higher status arises because supervisors have greater occupational prestige (Kanekar et al. 1989), hold power over others (Levine and Moreland 1990), and control more resources (Blau 1977). In fact, Driskell and Salas (1991), among others, have used hierarchical organizational position as the measure of status in their organizational research. There are exceptions, such as professional and artistic work, where the most knowledgeable or talented person’s superior task performance can trump hierarchical position. However, in most organizational settings, we would expect supervisors to be considered higher status compared with their subordinates.

We propose that any actions or characteristics, including demographic ones, serving to contest or undermine supervisors’ higher status claims relative to their subordinates would be contested by supervisors. Blau (1977) suggested that the more intersecting parameters (that is, the more those interacting occupy conflicting status positions, such as low-status age and high-status occupation), the greater the potential interpersonal conflict as individuals compete to assert the dominance of the parameter in which they have highest status. According to SDT supervisors, as members of the hierarchically high-status group, would be expected to be more subject to homophily preference, particularly because supervisory status is not ascriptive. Thus, according to Blau (1977), homophily preference is less salient and secure than ascribed statuses such as demography. Those with lower hierarchical status but higher demographic status need not act to assert their higher status claims based on a competing status parameter to initiate this dynamic. Research in SDT finds that those in dominant positions seek to protect their dominance regardless of any overt actions by subordinates (Sidanius et al. 2004). Furthermore, in Troyer and Younts’s (1997) extension of SCT, they found that high-status individuals responded to their perceptions of others’ expectations, whether any overt action to assert conflicting status claims took place. Thus, if supervisors (high organizational status) occupy lower demographic statuses than their subordinates, they face a potential challenge to their dominant position. In general, this means that higher demographic status subordinates challenge supervisors’ higher hierarchical status in ways that demographically similar or lower demographic status subordinates do not. Moreover, because competence claims are important to nonascriptive status claims (Blau 1977, DiTomaso et al. 2007a), denigrating the competence of higher demographic status subordinates is a way to remove or reduce the possibility of status loss that their higher demographic status poses. Brewer and Kramer (1985) found that the more insecure a person with dominant status, the more likely the person is to show similarity preference (Brewer and Kramer 1985). Contesting a subordinate’s claim to superior ascriptive status by giving them lower relative performance ratings would be a way for supervisors to defend their higher status.

This allows us to conduct strong inference tests (Platt 1964) pitting a status contest theory of supervisor ratings
skew against the homophily. We use two different demographic parameters to avoid the risk of inadvertently drawing general theoretical conclusions from relationships that could be particular to the meanings attached to any one demographic characteristic. Age and gender are two highly salient and visible demographic characteristics, conveying important meaning to organizational participants, and so both are used in these tests.

**Age-Based Status Incongruence.** Those who are older are expected to have more experience and expertise, and this enhances and supports the higher status of the supervisory role (Perry et al. 1999, Vecchio 1993). We propose that having a subordinate much older than themselves could be viewed as a claim to the greater status of experience and expertise, contesting the dominant hierarchical status of younger supervisors. Therefore, we expect to find that supervisors younger than their subordinates would tend to rate their subordinates’ task and contextual performance lower than they would rate their nonstatus contesting subordinates their same age or younger.

If the homophily explanation for demographic supervisory ratings skew is best, we would also expect to find that supervisors with younger subordinates would view them as negatively as supervisors with older subordinates, because both are equally dissimilar to themselves. However, the status contest argument predicts that only the younger supervisors of older subordinates would experience status contestation and seek to assert their claims by downgrading the competence of their older subordinates and demonstrate demographic skew in their ratings of their subordinates’ job performance. It is our contention that previous research findings of homophily effects were driven by status contest, and that once status contest supervisor–subordinate pairs are controlled, the homophily effects will disappear.

**Hypothesis 1A (H1A).** Supervisors with age-based high-status subordinates (supervisors younger than their subordinates) will give them poorer contextual performance ratings than they will their low-status and same-age status subordinates (supervisors older than their subordinates, or subordinates and supervisors of similar ages).

**Hypothesis 1B (H1B).** Supervisors with age-dissimilar subordinates (supervisors older or younger than their subordinates) will give them poorer contextual performance ratings than they will similar-age subordinates.

**Hypothesis 2A (H2A).** Supervisors with age-based high-status subordinates (supervisors younger than their subordinates) will give them poorer task performance ratings than they will to their low-status and same-age status subordinates (supervisors older than their subordinates, or subordinates and supervisor of similar ages).

**Hypothesis 2B (H2B).** Supervisors with age-dissimilar subordinates (supervisors older or younger than their subordinates) will give them poorer task performance ratings than they will to their similar-age subordinates.

**Gender-Based Status Incongruence.** Research on the effects of gender diversity on organizational behavior has burgeoned in recent decades (Shore et al. 2009), with a concomitant flowering of theoretical approaches. Although there is a large body of research finding that men hold higher status than women (Cohen and Zhou 1991, Thomas-Hunt and Phillips 2004), this insight has not been prominent in explaining gender-based supervisory performance rating skew. When jobs are dominated by women, they are perceived to have lower status (Kanekar et al. 1989). Women are less likely to engage in the higher status actions of interrupting, dominating, and confidently contributing their expertise than are men (Thomas-Hunt and Phillips 2004), and women are perceived as less competent than men (Driskell et al. 1993). This suggests that when (lower status) women occupy high-status supervisory roles over men (higher demographic status), these supervisors can feel that their status is contested.

If homophily best explained supervisory skew in performance ratings, we would also expect men supervising women to rate their women subordinates’ performances more negatively than that of their male subordinates, because both are equally dissimilar. We also do not expect the defense of status when women supervise women subordinates or men supervise men and women subordinates because the dominance of the supervisors’ formal hierarchical position would be uncontested by competing demographic status claims. This allows status contest and homophily preference theories to be tested against one another.

**Hypothesis 3A (H3A).** Supervisors with gender-based high-status subordinates (women supervising men) will give them poorer contextual performance ratings than they will their low-status and same-gender status subordinates (men supervising women, and same-gender supervisors and subordinates).

**Hypothesis 3B (H3B).** Supervisors with subordinates of dissimilar gender will give them poorer contextual performance ratings than those supervisors rating subordinates of the same gender as themselves.

**Hypothesis 4A (H4A).** Supervisors with gender-based high-status subordinates (women supervising men) will give them poorer task performance ratings than they will their low-status and same-gender status subordinates (men supervising women, and same-gender supervisors and subordinates).

**Hypothesis 4B (H4B).** Supervisors with subordinates of dissimilar gender will give them poorer task performance ratings than those with supervisors rating subordinates of the same gender as themselves.
Methods
The sample for this study consisted of 358 supervisor–subordinate dyads working in 10 U.S. companies, selected because they worked in highly competitive industries. Each company had more than 1,000 employees, and all sampled employees were nonsupervisory administrative or professionals. The sampled employees included engineers, scientists, technicians, marketing and sales professionals, administrative employees, and those working directly in food processing who packaged foods for restaurants and institutional food services. All of the supervisors rating employees’ performance were their immediate supervisors. Two organizations are in computer manufacturing, designing and producing small computers primarily for corporate and business customers. Two of the sampled organizations are in electronics and semiconductors, designing and producing components for computers and telephones. Two of the companies are nationwide telecommunications companies, providing equipment and services. One organization is a large national clothing retail chain, selling clothing and accessories to middle-market consumers. Finally, three nationwide food services companies are included. By sampling employees from more than one company, ranging from subordinates with doctorates in computer science to assembly-line food packers, we attempt to minimize the possibility of company- or industry-specific findings.

Eighty-five different jobs varying in ease with which employee performance could be assessed were selected. For each job, the names of the supervisors who directly supervised the employee performing the job were obtained. Subordinates and supervisors provided data independently via surveys that were mailed to the respondents’ work addresses, coded to enable the matching of supervisors and subordinates. Respondents were provided stamped, self-addressed envelopes to return completed surveys to the researchers’ university. Confidentiality of the survey was guaranteed. The supervisors had a response rate of 64% and provided data on their subordinates’ contextual and task performances, their own job tenure, and their own demographic characteristics. There were 96 different supervisors; 97% of the supervisors classified themselves as white and reported a mean organizational tenure of 11.1 years. For each supervisor three subordinates were randomly selected to receive questionnaires (60% response rate), with employees providing data on their personal demographics, job tenure, and trust in their supervisors; 87% of the subordinates classified themselves as white and reported a mean organizational tenure of 8.9 years. There were no identifiable differences between respondents and nonrespondents, and in none of these organizations did the low-status ascriptive group (women and older employees) dominate.

Measures
The independent variables are age and gender, with the supervisor–subordinate dyads grouped by whether or not they are status-incongruent or status-congruent (called Status incongruence) and by whether or not they are relationally dissimilar or similar (called Dissimilarity). Categorical age was measured in years. Categorical gender is measured by a dichotomous variable. A value of “1” was used to denote men and “2” women. The means, standard deviations, and intercorrelations for all study variables appear in Table 1. An alternative hypothesis is assessed using subordinates’ reporting Trust in supervisor, taken from Pearce et al. (1992) and consisting of seven items, with an $\alpha = 0.90$ in this study. Items included “I can rely on my supervisor” and “My supervisor seems willing to listen to my problems.”

Incongruence and Dissimilarity. For Status incongruence–Age, we calculated the age difference by subtracting the supervisor’s age from subordinate’s age. If the difference is larger than zero, it indicates that the subordinate is older than the supervisor and that this supervisor–subordinate dyad is Status incongruence–Age ($n = 58$). The number of age-incongruent supervisor–subordinate dyads is 58. We also tested for the alternative hypothesis that all younger supervisors may have poorer-performing subordinates or have (status undermining) less tenure in their jobs as supervisors by controlling for supervisor age and their job tenure in the tests.

The degree of age dissimilarity was indexed by the squared difference between the supervisor’s and the subordinate’s scores using a standard formula (Tsui et al. 1992). For example, a difference score of “0” on the age variable meant that a subordinate and a superior were identical in age. A difference score of “1” meant that they differed by one year, and a difference of “4” meant that they differed by two years, in either direction.

Status incongruence–Gender consists of supervisor–subordinate dyads with female supervisors and male subordinates (coded 1, $n = 10$), with all other dyads categorized as gender status-congruent (coded 0, $n = 348$). Supervisor–subordinate dyads are gender-relationally dissimilar if a man supervises a woman or a woman supervises a man ($n = 79$); all other dyads are considered gender-relationally similar ($n = 279$). We also controlled for the possibility that women may have had poorer-performing subordinates or may have had (status undermining) less tenure in their jobs as supervisors by controlling supervisor gender and job tenure in the tests. Dissimilarity–Gender was coded with a value of 1, and similarity was coded as 0. None of the organizations is female-dominated, and the dissimilar gender dyads were distributed widely among the 10 organizations.

Because race has been found to be a salient basis for demographic ratings skew, we control it in these tests:
for Dissimilarity–Race, supervisor–subordinate dyads are race similar if their races are the same (coded 0); all other dyads are considered race dissimilar (coded 1).

**Contextual Performance.** Supervisors’ ratings of subordinate Contextual performance were measured using nine items drawn from the existing literature (e.g., O’Reilly and Chatman 1986). This scale includes items like “makes suggestions to improve work procedures” and “is willing to speak up when policy does not contribute to goal achievement of department.” The supervisors stated the extent to which they agreed that their subordinates exhibited the actions described by the item, on an agree–disagree Likert scale. The Cronbach reliability coefficient for this scale is 0.93.

**Task Performance.** Task performance consists of supervisors’ reports on six items originally described in Tsui et al. (1997), where it was found to be sufficiently distinct from subordinate Contextual performance, with all items loading on their target factor and no cross-loadings. Task performance items focused on quantity, quality, and efficiency of job performance, rated on a Likert agree–disagree scale. Example questions include “employee's quantity of work is higher than average” and “employee strives for higher quality work than required.” The internal consistency reliability of this scale in this sample is 0.96.

### Analyses.

The hypotheses are at the supervisor–subordinate dyadic level; however, these dyads are nested in two other levels: organizations and supervisory raters who evaluated the performance of one, two, or three subordinates. Therefore, hierarchical linear models (HLMs) were used (Hofmann 1997, Kassinis and Vafeas 2006, Raudenbush and Bryk 2002, Singer 1998) with three levels: dyads, supervisors, and organizations. The hypotheses were tested in two steps. First, a null model was estimated that had no predictors at either level to partition the variance for the outcome variables into within-supervisor, between-supervisor, and between-organization components. For example, to test H1A and H1B, we examined the null model for Contextual performance and found that the large percentage of variance (93.52%) lies in dyads within supervisors (i.e., Level 1). A smaller yet still significant percentage (64.8%; \( \chi^2 = 61.79, df = 45, p < 0.05 \)) lies between supervisors within organizations (i.e., Level 2), but the variance between organizations is not significant (\( \chi^2 = 6.92, df = 9, n.s. \)). The intraclass correlation coefficient (ICC) at the organizational level is 0.06 (i.e., 6% of the variance in employee Contextual performance resided between organizations and 94% of the variance resided within organizations), and the ICC at the supervisor level is 0.17. The absence of significant interorganizational variation argues for considering a simpler, two-level analysis of the dyad level at Level 1 and supervisor level at Level 2, ignoring organization.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subordinate gender</td>
<td>3.96</td>
<td>1.26</td>
</tr>
<tr>
<td>Subordinate age</td>
<td>4.16</td>
<td>1.50</td>
</tr>
<tr>
<td>Supervisor gender</td>
<td>3.70</td>
<td>0.78</td>
</tr>
<tr>
<td>Supervisor age</td>
<td>3.80</td>
<td>1.32</td>
</tr>
<tr>
<td>Dissimilarity–Race</td>
<td>1.87</td>
<td>0.70</td>
</tr>
<tr>
<td>Dissimilarity–Gender</td>
<td>1.17</td>
<td>0.45</td>
</tr>
<tr>
<td>Contextual performance</td>
<td>3.83</td>
<td>0.70</td>
</tr>
<tr>
<td>Task performance</td>
<td>3.76</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Notes. \( n = 358 \); \( ^{a} \) = men, \( ^{b} \) = women, \( ^{c} \) = subordinates with dissimilar race supervisors, \( ^{d} \) = subordinates with same-gender supervisors. Larger values denote higher status incongruence.
Second, the conditional models to test our hypotheses were computed. To test H1A and H1B, hierarchical linear models were performed with Contextual performance as the outcome variable. At the dyad level, Dissimilarity–Age and Status incongruence–Age were used as predictors of Contextual performance, with subordinates’ job tenure and racial dissimilarity included as control variables. At the supervisor level, supervisors’ gender and job tenure were included as control variables. As can be seen in Table 1, the younger the supervisors, the lower they tended to rate their subordinates’ task and contextual performance, so we control for age in the gender tests, and gender in the age tests in all of the HLM models, to ensure that effects of categorical demography are not mistakenly interpreted as supporting either of the tested theories. To address the global explanatory power of the models, a pseudo-$R^2$ value was estimated as follows: the error terms in a null model were initially calculated and then replicated in the conditional model with the pseudo-$R^2$ value obtained by calculating the proportion of variance explained by the conditional model (Singer 1998). Finally, to aid in interpretation, all paired-comparison tests of group means were conducted, applying Levene’s tests for inequality of variances to all tests to control for the group-size variance in these samples.

**Results**

All of the “A” hypotheses make predictions derived from status contest explanations for skew in supervisory ratings, whereas the “B” hypotheses make contradictory predictions based on homophily preference explanations. Hypothesis 1 addresses the status and homophily predictions of age-based effects on Contextual performance, and Hypothesis 2 makes those predictions for Task performance. The tests, with unstandardized coefficients, are reported in Table 2.

We can see that the significant negative coefficients support status contest explanations for both Contextual and Task performance, and none of the homophily coefficients was significant, once the status-incongruent dyads were controlled (H1B and H2B). Supervisors rated their older subordinates significantly lower in Contextual performance and Task performance than did supervisors rating subordinates their own age or who were dissimilar but younger, supporting H1A and H2A.

Hypothesis 3 concerns the gender effects on Contextual performance, and Hypothesis 4 concerns the gender effects on Task performance. For Contextual performance the prediction for the status explanation was supported (H3A) with a significant negative coefficient for Status incongruence–Gender ($\beta = -0.68$, $p < 0.05$) but not for Dissimilarity–Gender. Without the status-incongruent dyads, homophily preference has no significant effects (H3B was not supported). Female supervisors rated their male subordinates as having significantly lower Contextual performance ($X = 3.28$) than did men rating either male ($X = 3.75$, $p < 0.05$) or female subordinates ($X = 3.87$, $p < 0.05$). Unfortunately, there were no significant gender-based relational demography effects for subordinates’ Task performance (see Table 2). This means that subordinates’ task performance ratings were not affected by either their own gender, their supervisor’s gender, or any combination, preventing the comparative test of gender-based status contest versus homophily explanations for supervisor ratings skew of Task performance (H4A versus H4B). This could be because task performance is more likely to be assessed objectively and so is less subject to supervisory ratings skew. Neverthelss, for those tests of gender-based performance ratings skew that were possible, the status contest hypotheses were supported and the homophily ones were not.

We also tested two alternative hypotheses for these results. First, DiTomaso et al. (2007b) found that those with high demographic status (U.S.-born men) received higher evaluations and rewards from all evaluators, supporting their argument that shared gender stereotypes, not homophily, drives biased treatment. Likewise, Duguid (2009) found that men were selected more often for teams by both men and women. Thus, it is possible that it is not contested status but simple bias in favor of those with high demographic status that might account for the results. To eliminate this alternative hypothesis, we conducted analysis of variance tests to compare how the supervisors rated Task performance and Contextual performance of both male and female subordinates. There was no significant difference in all supervisors’ ratings of their subordinates’ Contextual performance ($X_{men} = 3.21$, $X_{women} = 3.79$, n.s.) and Task performance ($X_{men} = 3.48$, $X_{women} = 3.80$, n.s.). Note that in these 10 organizations we did not find the bias in favor of men that DiTomaso et al. (2007b) and Duguid (2009) found. For the second alternative hypothesis, it is possible that high demographic status subordinates react so negatively to working for low demographic status supervisors that they become alienated and so actually perform more poorly. It is well established in social psychology that when one’s negatively stereotyped identity is made salient, stereotype threat can lead to poor performance (Steele and Aronson 1995). Therefore, it is possible that older employees emphasized their experience when interacting with younger supervisors and men working for women excessively interrupted and dominated more in interaction, becoming genuinely poor performers. We did not have employees’ perceptions of their performance or fine-grained data on employee actions in these 10 organizations. However, we had a measure of employees’ trust in their supervisor. In Table 1 we can see that employees’ trust in their supervisor was not associated with their own gender or age, nor with their
supervisors’ gender or age, nor with status incongruence or dissimilarity in gender or age. Thus, although we do not have any objective measures of employees’ performance, we can see that working for a younger or woman supervisor does not, in itself, lead to more distrust or alienated subordinates.

**Discussion and Conclusion**

We developed a status contest explanation that better accounts for the existing empirical data on when and why raters will tend to have demographic similarity-skewed ratings of their subordinates. The status explanation was tested against the currently dominant universal homophily explanation by placing predictions from the homophily explanation in head-to-head strong inference tests against status contest explanations. We found that, for both gender and age, status significantly accounted for the supervisory demographic skew in ratings, whereas the homophily explanations received no support.

We found that those subordinates whose higher ascriptive demographic status served to contest supervisors’ own higher hierarchical status were rated systematically comparatively lower in contextual performance and in task performance (when task performance skew could be tested). Consistent with our expectations, contextual performance ratings provide more supervisory discretion and thus more opportunity to protect and enhance supervisory status by denigrating the competence of demographically higher-status subordinates. We also found that these effects were independent of the race, gender, age, and job tenure of the supervisors. The tests were conducted on two demographic categories (age and gender), lending confidence that this may be a general phenomenon.

These tests confirm the claim that status is a powerful explanatory variable in organizational behavior. As such, it contributes to the growing attention to the role of status in understanding organizational behavior. Supervisors and others in organizations exist in social environments and are as concerned with enhancing and maintaining their status as are people in other social settings. Supervisors’ authority is fragile (Sayles 1989), and they can be quite alert to potential threats to their dominant status positions. Organizations are rife with status striving and defense, and theories of status promise to be a fruitful source of explanatory power for a variety of organizational behaviors.

Regarding homophily explanations of demography effects, these tests demonstrate that much previous relational demography research was conducted in a way that masked the more theoretically complete and empirically supported reason for demography effects—threats of relative status loss. The fact that none of the homophily hypotheses was supported strongly suggests that previous assumptions that homophily is the best explanation for relational diversity effects, whether among team members or in supervisor–subordinate relationships, appear to be misplaced. As noted earlier, homophily effects are well established; they just may not be the most important factors accounting for supervisory ratings biases or any of the other many documented demography effects in organizations. Status effects may
predominate over homophily in many workplace settings, as, for example, many may be happy to participate in high-status project teams and boards even when they are demographically different from most other participants. We are certain that the data used in other relational demography studies could be reanalyzed to conduct strong inference tests of status incongruence versus homophily as was done here.

We find it interesting to note that in testing for alternative hypotheses, we found that our results were not consistent with the DiTomaso et al. (2007b) and Duguid (2009) findings that both men and women were biased in favor of higher-status men. Here, we found no overall difference in supervisors’ ratings of men and women. We believe that extrapolating from SDT can help with these conflicting findings. In numerous other settings, SDT scholars have found that those with low status do not consistently favor their own kind—sometimes they do and sometimes they do not (Sidanius et al. 2004). These results suggest that learning the circumstances under which both men and women favor men in their ratings, and when they do not, would prove a fruitful area for future research.

We hope that this generalization of social dominance and status characteristics theories to an understanding of supervisors’ demographic skew of their ratings of their subordinates’ performance will help inform readers in the organizational sciences about the usefulness of these ideas. Work in organizational behavior is beginning to successfully use the individual difference measure of social dominance orientation (see Umphress et al. 2007, Shao et al. 2009), and we hope this study has helped to demonstrate that SDT is much more than a theory of individual differences. Social dominance theory’s explicit theorizing links institutions and individual action, making it particularly useful to the study of a wide variety of organizational behavior phenomena. Likewise, SCT, with its focus on others’ expectations and reactions to potential status loss, is a fruitful source of ideas for addressing the complex social dynamics of organizational behavior. As explanations of how people respond to the experience of dominating, being dominated, and status loss, these theories hold great potential in organizational behavior.

This study is not without limitations. It is possible that those subordinates receiving lower task and contextual performance ratings could possibly have had actual lower real performance. Older subordinates or men may have been passed over for promotion in favor of those with lower demographic status because of their poor performance. Another limitation is that neither psychologically experienced status incongruence nor homophily were directly measured in this study; we tested for patterns of data that would be predicted from each of these explanations. Given the difficulty of getting honest self-reports of status seeking and similarity preference, this study, like others in the literature, did not directly obtain these measures. A laboratory study in which objectively measured employee performance is manipulated could help clarify the underlying causal mechanisms proposed here. Another limitation is that we tested only two potential factors contributing to demography-driven status contests (age and gender), and we would expect many other achieved and ascriptive differences, such as race, social class, and educational prestige, to potentially undermine a supervisor’s hierarchical status. These arguments may generalize to other hierarchically ordered social categorizations, and testing these arguments for other groups could prove to be fruitful future research. We expect these and other boundary conditions on status effects, and hope this research, leads to theory development and research on the role of status and managerial practices.

This study introduces the idea of status contests to the study of supervisory ratings and managerial practices more generally, and we would hope it might spur a broader awareness of status contests in organizational behavior. Social dominance and status characteristics theories provide a theoretically coherent account for these, and possibly many other, demographic effects. We hope this support for a more complete theoretical account of the growing evidence that homophily is not the best explanation for supervisory demographic skew in their subordinates’ ratings will help those studying workplace bias and discrimination to develop more complete, and thus more useful, theoretical accounts. When demographic skew is misunderstood as universal homophily, practical interventions focused on bias eradication too often focus on increasing awareness of homophily as a means of reducing it. This research suggests that such training would be useless in reducing demographic skew in supervisory ratings: supervisors appear to be more concerned with their relative status than they are averse to someone who is demographically different from themselves. Social dominance and status characteristics theories emphasize that differences have status implications. Ignoring the role of status defense and assertion underlying organizational behaviors is a prescription for training impotence.

In conclusion, these results help to demonstrate what has long been apparent to sociologists and social psychologists—social status is important to people, and the seeking of higher status and avoidance of status loss drives much behavior in the workplace as in other social settings. Sociologists and social psychologists have long considered status a central explanatory variable, and the status-loss effects demonstrated here suggest that it could have a prominent place in a wider range of diversity and other organizational behavior theories.
Acknowledgments
Data collection was funded by the National Science Foundation [Grant SES-892123]. A previous version was presented in the Gender and Diversity in Organizations Division’s Program at the Academy of Management, Philadelphia, PA. The authors thank Joerg Dietz, Kristin Behfar, Greg Bigley, Kenji Klein, and the anonymous Organization Science reviewers for their comments.

References


Jone L. Pearce is the Dean’s Professor of Leadership and Director of the Center for Global Leadership at the Paul Merage School of Business, University of California, Irvine. She studies the effects of social processes on organizational behavior. She received her Ph.D. from Yale University and has taught at Carnegie Mellon University, the International Management Center (Budapest, Hungary), Budapest University of Economic Sciences, and the University of Washington.

Qiumei Jane Xu is currently an assistant professor of management at the College of Business and Management of Northeastern Illinois University. She received her Ph.D. in management from the Paul Merage School of Business, University of California, Irvine. She has conducted research on diversity, organizational citizenship behavior, and employee development.