Athena Rising?
Surfacing a Female Advantage in Horizontal Allocation to Work Groups

Adina D. Sterling
Stanford University
Graduate School of Business
adinad@stanford.edu

Jill Perry-Smith
Emory University
Goizueta Business School
jill.perry-smith@emory.edu

May 17, 2018

Draft: Do Not Cite or Circulate Without the Permission of the Authors

Abstract

While many studies investigate dynamics of teams and groups once they are already formed, lesser studied is horizontal allocation, or the assignment of employees to work groups and teams in organizations. In this paper the authors investigate the allocation patterns of male and female employees and their assignment to work groups whose members vary in social prominence. Expecting ex ante that women are at a disadvantage, the authors find the opposite using archival data on group allocation: women have higher returns to their performance in being allocated to groups with socially prominent members versus men. After surfacing this unanticipated pattern, the authors take an abductive approach to develop theory about how gender, performance, and numerical representation in organizations affect horizontal allocation. They test their arguments in an experiment with 640 working adults and find a female advantage in horizontal allocation that depends on the numerical representation of women in organizations. The authors discuss the contributions of this study for literature on gender inequality, organizational demography, and team allocation as well as how the quantitative abductive approach used here can be used more broadly to advance organizational scholarship.
INTRODUCTION

Even though women have made significant inroads into work organizations over the last few decades, women can nonetheless receive differential allocation outcomes versus men upon entry (Baron, 1984; Reskin and McBrier, 2000; Castilla and Benard, 2010; Huffman, Cohen, and Pearlman 2010; Merluzzi and Phillips, 2016). Organizations can separate employees into positions, roles, and jobs that permit opportunities to differ by gender (Acker, 1990; Chan and Anteby, 2015). Rates of promotion tied to jobs can differ according to whether they are largely occupied by women or men (Baron, Davis-Blake, and Bielby, 1986; Cohen and Broschak, 2013), and hurdles for advancement can also differ for these jobs (Gorman and Kmec, 2009; Haveman and Beresford, 2012). And yet, while studies yield important insights into how managers allocate employees into jobs or positions that differ along a hierarchy or within a division-of-labor, a number of conditions including the global environment and the life cycle of markets have altered the structure of organizations (Osterman, 1994; Smith, 1997; Vallas, 2003). Flat, horizontal organizations with teams or work groups have become increasingly prevalent in the organizational landscape (Kalev, 2009; Purnam, Alexy, and Reitzig, 2014; Barley, 2016). As a result, the decision many managers face is not how to allocate people into formal positions, jobs or roles, but how to allocate employees horizontally into teams and work groups.

From a theoretical standpoint horizontal allocation raises questions about what, if any, differential sorting may occur in organizations of this type. In horizontally-arranged organizations and subunits, the same or similar jobs are held by employees with similar background and training, such as scientists, consultants, or software engineers. These
employees collaborate in the joint-production of work (Zellmer-Bruhn, 2003; Hargadon, 2005; Uzzi and Shapiro, 2005; Taylor and Greve, 2006). A main building block for extant theories of gender inequality is that jobs generate differences in salaries, promotions, and resources across employees: outcomes for men and women differ because they actively sort, or are sorted into different jobs (Baron and Bielby, 1986; Strang and Baron, 1990; Baron, Mittiman, and Newman, 1991; Cohen and Broshak, 2013; Chan and Anteby, 2015). For this reason, it may be reasonable that horizontal organizations avoid differential sorting, merely because this option is not available to managers: jobs are essentially identical so sorting on this basis is a moot point. Nonetheless, even when jobs are the same other aspects of work may not be. Teams and work groups vary in the social prominence of their members. Some groups and teams offer employees a chance to affiliate with individuals with greater social prominence (Merton, 1968; Roberts and Sterling 2012; Hu and Randel 2014). Other groups contain less prominent members that have lower levels of visibility and fewer resources (Baron and Pfeffer 1994; Ely, 1994; Petersen, Saporta and Seidel 2000; Huffman, Cohen and Pearlman 2010). It could be that few differences exist within the same job but that individuals nonetheless vary in the social resources available in their work groups. Along these lines, we ask does gender affect the assignment of employees to work groups, and if so, how?

Our baseline expectation is that gender affects allocation because women get assigned to work on teams with less socially prominent members. To the degree performance is used to inform horizontal allocation, beliefs about the skills and abilities of women could lead them to be disadvantaged and men to be placed in groups with more prominent members (e.g. Correll and Benard, 2006; Castilla, 2008; Castilla and Benard, 2010). We explore our intuition using an
archival data set from a firm we term ChemCo. After constructing measures on performance, social prominence, gender, and other factors that might influence allocation with that data, we found that our measure of performance did increase the likelihood that employees would be allocated to teams with prominent work group members. However, much to our surprise women had higher not lower returns to their performance than men (see Figure 1 below).

Although the data and analyses we performed could not definitively “rule in” the surprising pattern of results, we did a number of tests that compelled us that this relationship deserved further scrutiny empirically and theoretically. After surfacing this surprising pattern in the data, we undertook a quantitative abductive approach to examine it further and then to develop theory (Lave and March, 1975). For example, from an empirical vantage Figure 1 shows the relationship between gender and allocation after propensity score matching—i.e. or after the sample of men and women are effectively matched on observable characteristics (Rosenbaum and Rubin, 1983). Across this analysis and in using the full sample of data across various model specifications, we found the empirical relationship between work group allocation and gender—i.e. that women had positive returns to performance—remained. From a theoretical vantage, we used this initial pattern in the data to develop arguments about the conditions under which women may be advantaged over men in horizontal allocation. We engaged in a “creative inferential process” to generate theory on a female advantage and the boundary condition for such an effect based on the demographic composition of organizations (Timmermans and Tavory, 2012). After developing our arguments we return to the field to test our predictions in an experiment among 640 employed adults that allocated two hypothetical
employees—Frank and Sarah—to two types of work groups, and find support for our theory. Taken together, the two studies through a quantitative abductive approach lend support to a female advantage in horizontal allocation, and also reveal the conditions under which such a female advantage is unlikely to persist.

THE ASSIGNMENT OF EMPLOYEES TO WORK GROUPS AND TEAMS

Horizontal allocation refers to the decisions allocators make regarding how to assemble teams and work groups. In horizontally-structured organizations or subunits, managers must decide how to allocate individuals to work groups to carry out joint-production. While there may be supply-side reasons employees join a work group, here we concentrate on how allocators sort employees into groups, not supply-side behavior. Explicitly, we investigate the assignment of employees to work groups as a type of allocation decision similar to the assignment of employees to supervisors or subordinates (e.g. Briscoe and Kellogg, 2011). While a great deal of research investigates how groups and teams that are already assembled interact, divide up work, recognize expertise, and engage in group interactions, to the best of our knowledge very little attention has been paid to the horizontal allocation decision itself—i.e. how groups and teams are assembled. Yet it is important to do so given the way work groups vary. Some work groups are provided more resources and receive greater visibility, largely due to the social prominence of the members of these work groups (Merton, 1968). Other work groups may have less prestigious members, receive poorer visibility and have fewer resources.

A key problem facing allocators is how to assemble groups to meet the challenges of
organizations and the activities at-hand. Some projects require members of groups to solve more complex problems and develop radical solutions while other projects require more incremental inputs (March, 1991; Perretti and Negro, 2006). For example in professional service firms, some client projects are more complex or important in terms of profitability, while others are less so (Von Nordenflycht, 2010). Allocators concern themselves with evaluating employees’ potential for helping a team achieve high performance (Reagans, Zuckerman, and McEvily, 2004). We expect the most highly-skilled, well-performing employees are likely channeled into the same work group to focus on challenging, complex tasks. These are the work groups likely made up of the most prominent members with high levels of visibility and resources in the organization.

The notion that organizations match individuals’ perceived skills and abilities as signaled by their work performance (Spence, 1973) to technical project requirements is an assumption of much of the literature on the performance implications of diversity (e.g. McCain, O’Reilly, and Pfeffer 1983; O’Reilly, Caldwell, and Barnett 1989), as well as research on group effectiveness (e.g. Reagans, Zuckerman, and McEvily 2004). For example, at IDEO, the toughest problems are allocated to engineers who have a history of solving uncertain, poorly defined problems (Hargadon and Sutton 1997). It is likely that those with a good track record of performance are assigned to work with one another (Chandler 1962; Amburgey and Dacin 1994; Okhuysen and Eisenhardt 2002). Lesser-performing individuals are likely to be assigned to work in groups with less important projects and less socially prominent members.

From the standpoint of employees and their behavior, it also seems likely that better performance would lead employees to be allocated into groups with more socially prominent
members because those with social prominence desire to work with high performers (Hinds et al., 2000). In academia for example, prominent members—i.e. those well-regarded with substantial acknowledgement and deference given to their work—may desire to work with the PhD students they perceive to have the most promise, as indicated by assessments of their ability or test scores, precisely because they feel they would be better collaborators and recipients of training. For the same reasons, employees with high social prominence may have sway deciding the composition of teams, and advocate for those with high performance to join them. We expect the following.

Hypothesis 1. As the performance of an employee increases the social prominence of those with whom s/he is assigned to work in groups also increases.

HORIZONTAL ALLOCATION AND A FEMALE-ADVANTAGE

To the degree the social prominence of work group members varies across work groups, it may be that horizontal allocation operates similarly to vertical allocation. In vertical allocation work performance and individual characteristics—i.e. gender—influences the allocation of employees (Elvira and Graham, 2002; Castilla, 2008; Huffman, Cohen, and Pearlman, 2010). By and large, studies of vertical allocation suggest women are disadvantage relative to men. Evidence indicates men receive higher returns to their performance than women with respect to vertical allocation including promotions (Cohen, Broschak and Haveman 1998; Castilla and Benard, 2010), raises (Merluzzi and Dobrev, 2015), and assignment to technically-challenging jobs (DiTomaso et. al., 2007). Men can receive greater rewards for a given level of performance due to cultural beliefs about men possessing superior abilities compared to women; when women perform well it may be attributed to chance or
outside factors, while for men it is attributed to innate ability or superior skills (Ridgeway, 1991; Ridgeway and Correll, 2004). Our initial speculation about women being disadvantaged followed the hypo-deductive process that exists in normal science (Kuhn, 1962), meaning that we began with the expectation women would be disadvantaged in horizontal allocation. Our initial speculations did not seem to pan out. Using archival data, the pattern of results (see Figure 1) suggested women were advantaged relative to men (see the analyses that follow). After assessing these initial, albeit tentative findings, we were compelled to develop new theory on why women might be advantaged in assignment to work groups and the contingencies of such an effect and take these arguments to an empirical test. This is the theory we present herein.

GENDER, PERFORMANCE AND HORIZONTAL ALLOCATION

To the degree allocators consider the performance of employees, it is reasonable to expect gender to influence horizontal allocation. On one hand, it may be reasonable that women are penalized when it comes to being allocated to work in groups with prominent members (as we initially suspected). To the degree that allocators assign employees purely on the basis of performance, cultural beliefs and stereotypes may lead women to be associated with lower levels of competency; therefore they may be more apt than men to be assigned to groups with lesser prominent employees (Joshi, 2014).

Nonetheless the allocation of employees into work groups and teams is apt to involve factors beyond individual-level performance. On teams individuals must not be able to only perform well on their own, they must be able to get along and work well together. Thus, analytical problem-solving skills are needed, but so too are skills that enable people to get
along with others, as well as skills that help integrate others ideas by getting people to work collaboratively (Drazin, Glynn, and Kazanjian, 1999; Hargadon and Sutton, 1997). Allocators may value the ability of employees to get along on teams and to help bring people together to solve problems (Katzenbach and Smith, 1993). If this is the case, otherwise equally performing women may have an advantage over men in this regard. Cultural beliefs exist about which skills and skill-sets are categorically more masculine versus feminine. In Western societies, women are viewed as being more relationship-oriented and helpful than men (Rosette and Tost 2010). Women, simply put, may be viewed as “playing nice” on teams in ways that advance a team’s goals of helping to drive collaborative inputs, or “collective work-products” (Katzenbach and Smith, 1993). That is, women may be viewed as being better team members than men in the sense that they can get along with others. This means all else equal (in terms of performance), women may have an advantage over men when being allocated to work in prominent groups.

Moreover, some evidence indicates women are believed to have skills that allow them not only to get along better with others, but that might help lead groups to more collaborative behavior (Eagly 2013). In teams some aspect of leadership involves handling conflict between team members. Conflict is an emotionally-heavy task, and these tasks in generally are typically viewed as better handled by women than men (Hochschild 1983). Whether or not women actually behave in these ways is irrelevant (Eagly and Johannesen-Schmidt and van Engen 2003; Eagly 2013). Gender as a macro social status provides an “implicit, background identity” (Ridgeway 1997: 231) consistent with this stereotype of women being able to “play nice” in groups and even help bring people together in group settings (Ridgeway and Correll, 2004).
For this reason, allocators may look to gender to develop an assessment of who will get along well with others on teams, and even who might lead the team in this regard. There is some evidence that in certain contexts, women are seen as more natural leaders. In the C-suite leadership literature, while studies indicate that stereotypes of leaders are consistent with masculinity, some evidence suggests that when organizations utilize more democratic and participatory decision-making processes, including delegation and team-based decision-making approaches, the view of women as leaders emerges more easily (Avolio and Gardner, 2005; Eagly, 2007). Studies of transformational leadership which deal with communality and collaboration across individuals find women are viewed more positively that leadership styles that emphasize agency (Bass and Avolio, 1994; Dezso and Ross, 2008). For this reason, somewhat paradoxically, despite men being viewed as leaders generally, they might be viewed as lacking collaborative skills needed to bring a team or work group together relative to women, all else equal. When conflict arises in a team or work group, women face expectations to get involved. Along these same lines, there is some evidence that women on teams may be viewed in familial terms as a “team mom” that leads a group by facilitating interactions among members of a group (Fletcher, 1999). In a signaling framework (Spence, 1973), high-performing women may signal they have performance-enhancing skills, and their femininity may serve as a social cue that indicates they also have skills beneficial to groups— that they get along well with others, and that they may help others get along, too. Men may suffer a penalty in horizontal allocation owing to their lack of this signal. All else equal, given two equally-performing men and women, allocators may favorably bias women in allocation. Overall, we predict the following.
**Hypothesis 2.** Women receive higher gains from their work performance than men with respect to being assigned to work in groups with members of high social prominence—i.e. there is a positive interaction effect between being female and work performance.

Thus far we have theorized about the relationship between gender and horizontal allocation. We now turn to contingencies of this effect. Namely, we argue the demographic composition of an organization influences female advantage. Specifically, we suggest that an advantage for women exists in allocation when women are rare. Kanter (1977) highlights that being in a low-proportion demographic group (in what is termed a skewed organization) leads to an adverse set of outcomes for numerical minorities due to tokenism. The term refers to the behaviors that ensue from being in a highly skewed situation—i.e. being in a demographic group composing less than 15 to 20 percent of the organization (Wallace and Kay, 2012; Williams, Phillips, and Hall 2016). While generally the tokenism literature has focused on unfavorable outcomes that accrue to the minority group (e.g. Turco, 2012), there is also evidence that favorable allocation outcomes may occur when the minority group is in low proportion. For example, studies on C-suite promotions at the upper levels of organizations suggest leaders face pressure from outside clients and other stakeholders to promote women to the executive ranks (Beckman and Phillips, 2005; Petersen and Saporta, 2004), and some evidence suggests women are promoted to leadership roles faster than men when women at the upper levels of the organization are lacking (Tsui and Gutek, 1984; Spilerman and Petersen, 1999; Dencker, 2008).

When women are fewer in number they receive more attention in ways that amplify their differences versus men. This could be beneficial in horizontal organizations centered around teams. When women are perceived as role-incongruent in male-dominated social
contexts, perceived differences become more salient and become heightened along stereotypical lines. In the case of horizontal organizations or subgroups, when women are rare, this may serve to draw strong attention to expectations of how women might operate differently in work groups than men, in terms of their getting along well with others, or their helping to bring people together. Given their rarity, having women on work teams could become valuable “in its own right.”

Studies indicate that even when women and men work side-by-side and have the same occupations, jobs, or titles in organizations, they can be expected to perform different tasks (e.g. Chan and Anteby, 2015). In terms of work groups, women might be expected to perform compensatory tasks on teams, especially along stereotypical lines. When women are rare in the organization, they may be more apt to be channeled to socially prominent work groups. These are the groups deserving of the most resources, including the perceived ability of women to get along with others and help smooth things over on teams.

Hypothesis 3. Women are more likely to be allocated into work groups with prominent members when women are numerically rare in the organization compared to when women are more prevalent.

Finally, for the subset of women that are strong performers, numerical rarity draws attention to their skills and they be more salient to allocators. Their skills may stand out because they seem more remarkable, and thus allocators may be more apt to take notice. This idea is rooted in double discounting theory (Foschi, 2000, 2006; Rosette and Tost, 2010). According to this logic, when employees’ performance is objective, or where there exists objective criteria for performance, women with strong performance get noticed. The reasons for this have to do with attributions, and the belief that women may have had to overcome more
strenuous obstacles (Sackett, Dubois, and Noe, 1991). When there are lower numerical
proportions of women, beliefs about double standards—i.e. that it is harder for women to
succeed than for men to succeed—are activated across members of organizations (Rosette and
Tost, 2010). When women display strong performance, it may be attributed to them having
succeeded “against the odds” so their performance is remembered and stands out more
compared to men’s.

In the leadership literature this logic and its link to double discounting has been found
to help explain why women leaders can stand out and be viewed as having displayed more
compelling evidence of their leadership capabilities than men with the same performance
(Lyness and Heilman, 2006; Ragins, Townsend, and Mattis, 1998). As long as there is some
objective or clear measure of performance, when rare women clearly demonstrate their
performance they stand draw attention to that performance. High-performing women in the
numerical minority may take on a heroic-like status and come to be seen as “overcoming
unique and relatively burdensome obstacles” in such a way that clear evidence overcomes
dISCOUNTING and indicates that the person has strong levels of ability (Wilson et. al 1999:166).
After having their performance stand out more, they receive favorable allocation outcomes. We
predict the following.

Hypothesis 4. The effect of women receiving higher gains from their work performance than
men in horizontal allocation is amplified when women are numerically rare in the organization
compared to when women are more prevalent.

METHOD: A QUANTITATIVE ABDUCTIVE APPROACH

We use a quantitative abductive approach to develop and investigate our theory, meaning
that we use an initial unexpected pattern of empirical results to inform our theorizing. Abduction is often associated with qualitative research (see Timmerman and Tavory, 2012), yet there is nothing that places abductive reasoning squarely within qualitative rather than quantitative approaches. From a philosophy-of-science standpoint, abduction itself is nothing more than one of three main methods of logic in the production of scientific knowledge, and it goes hand-in-hand with induction and deduction (Peirce 1878). Timmermans and Tavory (2012:170) suggest “in the context of research, abduction refers to an inferential creative process of producing new hypotheses and theories based on surprising research evidence. A researcher is led away from old to new theoretical insights.” Abduction begins with observation(s) and surmises that the cause and effect of such that is hidden from view. It is most confused with induction, but to draw a distinction, induction is a method of seeking empirical facts to form a hypothesis, while abduction seeks a hypothesis to account for a possible empirical relationship. We do not test hypotheses as a part of Study 1. We use it to get an initial indication of what the pattern of relationships might be, though not to necessarily “rule in” the relationship per se (Lave and March, 1975). We test of our hypotheses in Study 2.

Study 1: Archival Data Using ChemCo. Data

The first study is used to inform the pattern of relationships between work performance, gender, and horizontal allocation using data on patents from an organization we term ChemCo. We were compelled to using patent data as a baseline for initial inquiry because like other creative outputs such as books, articles, media, and art, citations of patents can provide a type of deference that can indicate the degree of social prominence of the actors associated with the output (Podolny and Phillips, 1996; Gould, 2002; Bothner, Smith, and White, 2010). The study
proceeded after the authors interviewed an R&D director in the organization. The director suggested patents were a good indication of work group assignments, as the collaborative efforts of scientists in the organization were largely documented in patent applications. The organization had an IP strategy to patent any improvement in technology, no matter how incremental, and most groups continued working on projects until they had a patentable technology before being reassigned, and represented the most comprehensive allocation records they had over the period. We asked the R&D director if scientists got to choose with whom they worked, and the director indicated that they did not solicit employee requests. Rather, employees were allocated to teams based on the needs of the business and the trajectory of the technology.

Our unit of analysis is the scientist and we took several steps to prepare the data before analyzing it. We began with the patent records of patents applied for at ChemCo from 1975-2000 (Hall, Jaffee, and Trajenberg 2001). As a first step we identified individual scientists at ChemCo. This required us to clean and identify unique names in the patent data. On patent records the same person may appear with slightly different names, initials, or surnames. In order to surface unique scientists we applied a name-matching algorithm to identify focal scientists in the sample and assign them unique IDs using Soundex (Trajenberg, Shiff, and Melamed 2006) ¹.

After gleaning a first-cut of potential same-individuals, we further inspected these initial matches according to a pre-set heuristic. Scientists at ChemCo were located in multiple cities and towns within the U.S. After consulting prior studies, we assigned scientists the same unique

¹ To do this we used Soundex. Soundex is a SAS program that matches names phonetically. This program is commonly used to identify errors and name misspellings in patent data. For example, Soundex identifies Phoebe and Phobe as being the same name phonetically.
identification number based on their name, city, and frequency of their name in the overall U.S. population (Fleming, Mingo, and Chen 2007). Specifically, our heuristic was to assign scientists the same unique inventor ID if either: 1) scientists had the same first, middle, and last names and city/town or 2) the scientists had the same first, middle, and last names but were located in different cities/towns and the frequency of their first and last names occurred in less than 0.01% of the population. There are 3,575 unique scientists in our data set.

The data set was constructed by linking each scientist in the organization to each of his or her patents. The patents were linked to the scientist in chronological order in terms of the patent application date from 1975 to 2000. Based on prior literature and the discussion with the R&D director, the *number of patents* that accrued to an employee prior to the placement of him or her in a group was used as the performance variable, because in this organization, there was attention paid to patent successes, in terms of overall patent counts. At ChemCo. being able to bring a technology to the point in which it was patentable was considered a core aspect of employee performance. As previously intimated, it is for this reason our performance variable is the *number of patents* that have accrued to an individual prior to the placement of him or her on a team. The demographic variable *female* is dichotomous, equal to one if an employee is a women, and zero otherwise. Consist with prior studies of name classifications (e.g. see the

---

2 Frequency of name usage in the U.S. was assessed using the website http://names.mongabay.com/most_common_surnames.htm

3 Some prior studies have also stipulated that unique IDs are assigned to individuals only if they are also working in the same patent technology classes or have the same co-scientists (i.e. Fleming et. al. 2007). In our case, we knew that in this organization, managers may assign scientists to work with other scientists in a variety of technological arenas.

4 To account for the possibility of left-truncation of those scientists, we included only scientists that did not have any patents prior to the observation window.
USPTO 1998 Patent Report) we consulted name books (Rule 1966; 1985) to classify scientists by gender that were relevant for our period of inquiry. In robustness checks we used different measures of performance (i.e. rate of patenting rather than overall patent counts, number of solo authored patents). The pattern of results remained the same.

In terms of the dependent variable, social prominence refers to the social esteem or deference afforded to an individual based on their output (Podolny and Phillips 1996; Gould, 2002; Bothner, Smith, and White, 2010). For example, in academia performance can be measured in the number of publications that individuals have with some weight given to the outlets in which those publications appear, while prominence has to do with how many citations the publications have received because it indicates that the work has been acknowledged by others. While of course related, the two constructs are conceptually distinct. Performance is about individual-level output, while social prominence has to do with the recognition that the output is given by some audience. The social prominence of a focal employees’ group members with whom he/she was assigned to work when developing a new technology. That is, we define the prominence of an individual as follows:

\[ D_{it} = \frac{\sum_j C_{ijt}}{L_t} \quad i \neq j \]

where \( D_{it} \) is the prominence of the group member i at time t, and \( C_{ijt} \) is coded a 1 when the group member i is cited by another inventor j within the organization during the interval t, and \( L_t \) is the total number of patent citations accruing to all inventors during the interval t. The restriction \( i \neq j \) removes self-citations. We record social prominence using the citations by other employees within ChemCo in the five years prior to a scientist being placed into a group.
We again used other windows (two and three years prior) and this did not have a substantive effect on the observational results.

Finally, we included several control variables that may affect the relationship between performance and the allocation of employees into groups with members of varying social prominence. We control for observed tenure of employees which we measure as the difference in years between the application year on a patent, and the year that an employee first appears in the sample. This proxies for the extent to which the performance of employees and allocation to prominent members within a group increases given the number of years an employee spends in the organization. Additionally, we include a measure of the average tenure of group members. With tenure citations have longer to accrue.

We additionally control for the frequency of patents within patent class. The number of patents applied for varies across patent classes. The USPTO divides patents into approximately 400 main technology classes. Patents may be more prolific in a tech class because they represent a main priority of the organization. The patent class in which an inventor works influences the likelihood of developing new technologies and the prominence of inventors that get placed on projects. The likelihood of being paired with highly prominent members increases with size, so we control for the number of group members on a patent. Finally, in order to account for variability across time we include year dummies in the models. We use OLS regression with robust standard errors clustered by scientist to assess the relationship between performance and allocation to groups with members that vary in prominence.

---

5 For first employees appearing on patents before 1975 their tenure is calculated by looking at inventors in the three decades prior.
Results

Before being allocated to a group, the average number of patents an employee has produced at the firm is approximately four (3.98). In this sample approximately 6% of the employees are women. Employees were assigned to work in groups with just over two members, not including the focal employee. The measured tenure of the focal employee and his or her group members is approximately 4.5 years (4.53 and 4.45, respectively). The descriptive statistics are provided in Table 1.

[INSERT TABLE 1 ABOUT HERE]

The models with the controls are shown in Table 2. An employee’s work performance had a positive and statistically significant effect on the average group member prominence in which s/he was assigned to work in all models, with the coefficient on performance being statistically significant across each (p < 0.01). In terms of effect size, the results are substantial: the marginal effects calculated at the average level of performance indicate a scientist gets allocated to a team where a member has a maximum prominence level that is 64 percent higher than the team the individual would have been allocated absent the patent. Turning to Hypothesis 2, as indicated previously we found an unanticipated positive and statistically significant coefficient on the female x performance interaction variable (p < 0.01). The results suggest women receive higher returns to their performance than men.

[INSERT TABLE 2 ABOUT HERE]

While providing some evidence that this pattern of relationships is indeed true, we were compelled to do other checks of this pattern (see the appendix for further details including on
other measures of the DV). One that we draw attention to here is that we further considered if our positive interaction effect could be driven by selection out of the organization. In male-dominated fields such as the one examined here, perhaps only women in the right tail of the distribution are in the organization—i.e. superstars—because other women opt-out of working for the firm. Indeed, arguments and support for the female superstar hypothesis have been found for female CEOs (Gayle, Golan, and Miller 2012) and in academic settings where in male-dominated fields such as engineering women tend to have higher GPAs than men because low performing women drop out of engineering programs (Sonnert and Fox 2012; Wolfe and Powell, 2009). In our setting women make up 6% of scientists. This is somewhat lower than reports by the U.S. Commerce Department about the share of patents accruing to women during this period (Button to Biotech Report 1996). If only superstar women remain because others drop out, the underlying quality distribution of men and women in our sample may be driving our results.6

Given this possibility we undertake a propensity score matching approach (Rosenbaum and Rubin 1983). Propensity score analysis attempts to reduce the bias due to non-random sorting into a treatment and control condition. Matching techniques mimic randomization that occurs experimentally by creating a sample of observations that received the treatment that is comparable on all observed covariates to a sample of observations that did not receive the treatment.

We perform propensity score matching (from herein PSM) using nearest neighbor matching with caliper ($\varepsilon < 0.25\sigma_0$) to rebalance the sample (Caliendo and Kopeinig, 2008). Post-

---

6 An analysis of the subsamples indicate that differences exist in the tenure, performance and number of group members across gender subgroups ($p < 0.01, \chi^2$ test).
matching $\chi^2$ tests reveal no meaningful differences in covariates across groups once the sample is balanced. Density plots of the estimated propensity score for the full and matched data are shown in the appendix.

We provide the propensity score matching results in Table 3 that go along with Figure 1 that we showed initially. Overall, we find that our results are substantively the same. In models where the maximum social prominence of group members is used as the dependent variable, a one-standard deviation increase in the number of patents a scientist has leads to a 14.1% standard deviation increase in maximum prominence of group members in the post-PSM balanced models (in the unbalanced models, there is a 8.7% standard deviation increase). The results continue to indicate there is no main effect for gender. Again, we find evidence that women receive higher returns to performance than men, as the coefficient on the interaction is positive and statistically significant ($p < 0.01$).

[INSERT TABLE 3 ABOUT HERE]

In sum, while we could not fully “rule in” the pattern of relationships that we found could not be an artifact of the setting or due to some aspect of the measures, the analyses that we ran led us to believe that there might, indeed, be “something there” in terms of a previously unanticipated positive female advantage in allocation. We were compelled to develop the aforementioned theory based on these surprising findings, and submit them to a test, in order to see if they map onto our findings from Study 1.

Study 2: Experimental Study of Work Group Allocation
An experiment was chosen for the second study. By designing an experiment we could randomly assign the study participants—who we asked to assume the role of allocators—to conditions where employee gender and performance varied. This allowed us to have careful control over aspects of the design. Second, we could more cogently manipulate the conditions under which a female advantage may or may not hold. For instance, if we were to look at the influence of potential positive returns to women based on numerical rarity in actual organizations, it would be hard to not rule out omitted factors that would impinge upon allocation decisions and the numerical proportionality of groups simultaneously. In this way, experimentally manipulating the numerical composition allows us to deal with “omitted variable bias” as an alternative explanation for our findings for Study 1, or allows us to “rule out” suspects in an abductive sense (Peirce 1878).

In Study 2 participants were recruited from a Qualtrics panel of working adults interested in participating in web-based research. The average age was 36.8, and 27% were men. Given the importance of gender to our research question we sought to achieve a balance between male and female participants across all conditions. Given an initial skew in participant gender in which we collected data from more women participants, we collected an additional sample of only men. Overall, the final sample was 52% men, and the average age was 37.3. All participants passed two attention filters and confirmed that they were currently employed. Our final sample contained 640 working adults.

---

7 Participants received a nominal payment for their participation.
Participants were randomly assigned to a 2 (target gender: male, female) x 2 (target performance: high, average) x 2 (proportionality: high proportion male, balanced gender) between subjects factorial design. Participants were told that they would be participating in a study examining how employees are assigned to work groups. We used the term ‘team’ in the actual manipulation because it was a term more familiar to participants and elicited expectations of teamwork, which we wanted to prime in the participants. Participants were asked to assume the role of a manager of a consulting firm that assists client companies with a range of strategic issues. Consulting firms are common types of horizontal bureaucracies. We described employees as working with client companies in groups of 3-5 employees. They were clearly told that one of their primary responsibilities was to assemble effective work groups among available employees.

Their task was to assign an available employee to a client work group in an organization where social prominence varied across work groups. Participants first reviewed a list of all employees in the department. Among listed employees, one was noted as being available for assignment. They were given performance information about the target employee and about average performance in the department. Then they were asked to make their assignment decisions and answered manipulation check and demographic questions.

**Manipulations.** The list of department employees contained seven employees, which indicated each employee’s name, assignment status, and gender.

**Target gender.** Depending on the gender condition to which the participant was assigned, the target employee had either a female or male name (i.e., Frank or Sarah) and their assignment status was noted as “[He/she] is available to be assigned.” In all cases, the last name
column was blacked out for each department employee and noted as “confidential.” All employees, other than the target employee, were noted as “[He/she] has already been assigned.”

**Proportionality.** Depending on the sex composition condition to which they were assigned, participants saw a list of either six male names (high proportion male condition) or a list containing three female and three male names (balanced sex condition). The high proportion male condition included: Dan, Greg, Eric, Andrew, Robert, and Steve. The balanced sex condition included: Jennifer, Andrea, James, Andrew, William, and Susan. We compiled these names after consulting experimental studies that manipulate gender (e.g. Cuddy, Fiske, and Glick 2004; Belliveau 2012). In addition, the gender of each department member was reinforced by using “he” or “she” in the adjacent column noting each member’s availability. Last, the effectiveness of the proportionality manipulation was supported by our manipulation check results (see Results section).

**Target performance.** To understand the effects of performance we created high and moderate performing conditions. During the experiment, after participants reviewed the list of department employees, participants then read additional information about the target employee (i.e. Frank or Sarah). More specifically, they were asked to review data about his/her client ratings over the past year with client comments. They were also told the average client rating for employees in the department was 3.5 (out of 5). In the high-performance condition, they read the following: “[Frank/Sarah]’s rating of 4.9 reflects that [he/she] is a ‘stellar performer’ who effectively manages client accounts to drive client satisfaction and revenue generation. [His/her] performance puts [him/her] in the top 5% of all employees.”
In the average condition, they read the following: “[Frank/Sarah]’s rating of 3.6 reflects that [he/she] is an ‘average performer’ who manages client accounts similar to other employees in the firm. [His/her] performance puts [him/her] among the top 50% of all employees.”

They then read a chart stating the target’s name (Frank or Sarah), his or her average client ratings out of 5 (4.9 or 3.6), and a list of client comments (“driven, exceeds expectations, goes above and beyond, exceptional” for the high-performance condition or “reliable, meets expectations, delivers client needs on time, satisfactory” for the average performance condition).

**Measures**

**Manipulation checks.** The post-task questionnaire contained questions to assess the effectiveness of the manipulations. The questions were designed to check whether participants recognized the sex composition of the organization and the performance level of the target. Two questions related to proportionality were averaged (r = 0.45, p < 0.01): “The employees in the department were mostly men” (reverse coded) and “The department contains a balanced number of men and women.” Two questions, which assessed the target’s performance, were averaged (r=0.24, p < 0.01): “The employee I was asked to assign was an outstanding performer” and “The employee I was asked to assign was a below average performer” (reverse coded).

**Work Group Allocation.** We assessed work group assignment using work group suitability and group assignment. In the experiment participants first selected the work group to which they assigned the target employee. They were given two choices: a high-prominence work group and a low-prominence work group and read a description of each. To avoid confusion we
described the work groups as teams in order to convey the idea that individuals would be working together. For the same reason, we substituted the term status for social prominence.

The high-prominence work group (group A) was described as follows:

This team will serve your biggest and most important client. This is a high-profile team, because in addition to serving your most lucrative client, the other members of the team are some of the highest status employees within the firm. They are well connected, and whomever you place on the team would get an outstanding opportunity to develop new relationships leading to a better network.

The low-prominence work group (group B) was described as follows:

This team will serve one of your typical, mid-range clients. These clients are typical of the firm’s “bread and butter” type of clients. In other words, there are many teams in the firm that serve clients that generate this level of revenue. The other members of the team are employees with mid-level status. Whoever is placed on this team would have the chance to solidify existing relationships in an established network.

After reading this, participants were asked to select the work group to assign Sarah/Frank: Team A or Team B. Next, participants were asked to rate the extent the high-performing work group was the best choice for the target employee. More specifically, participants were asked to rate their agreement with the statement “Team A is the best choice for [Sarah/Frank]” (1=strongly disagree to 7=strongly agree).

Results

We began by assessing the effectiveness of the manipulation. As expected, participants in the balanced gender condition reported the organization’s gender composition was significantly more balanced (M=4.5) than participants in the high proportion male condition (M=3.5),
(F(1,645)=66.35, p < 0.01). Similarly, those in the high target performance condition reported that the target employee was a higher performer (M=6.3) relative to those in the average target performance condition (M=3.5), (F(1,644)=547.53, p < 0.01).

Because of the possibility that it might affect the results, in addition to our primary variables of interest, we also measured participant gender and participant education level. We found there is a non-significant effect of participant gender on high prominence team suitability (F(1,642)=2.59, p > 0.05). We also did not find a significant effect of education on high-prominence work group suitability (F(3,640)=.41, p > 0.05). Given these findings, the analyses reported below do not control for gender or education, although controlling for both did not substantively change the results.

Turning to our predicted relationships, we expect to see a main effect of performance on assignment to work with socially prominent individuals. As Table 1 indicates, we find a main effect of performance (F(1,632)=121.52, p < 0.01) such that highly performing employees are considered significantly more suitable to work with highly prominent individuals in groups (M=5.31) compared to average performing employees (M=3.68). Thus, we find support for Hypothesis 1.

[INSERT TABLE 4 AND FIGURE 2 ABOUT HERE]

Next, we turn to the effect of performance and gender. In Hypothesis 2 we predicted that there would be a positive interaction between being female and work performance. In Table 1 we show the two- and three-way interactions for gender, work performance, and sex composition. While Hypothesis 2 is positive as expected, the effect is not statistically significant, and thus we
find no support for this hypothesis. Turning to Hypothesis 3, we find a positive and statistically significant interaction effect for target gender (female) x department composition (low women), where $F(1,632)=8.65$, $p < 0.01$. In other words, as predicted, women are being allocated to prominent teams when women are numerically rare in the organization. Finally, turning to Hypothesis 4, we find a significant three-way interaction between target gender x target performance x proportionality ($F(1,632)=7.13$, $p < 0.01$). Figure 2 and Figure 3 shows this graphically across the balanced gender condition, and the female-numerically-rare condition respectively. Consistent with Hypothesis 4 Tukey post hoc mean difference tests revealed that high-performing female targets in the low-proportion female condition were rated as significantly more suitable to join the high prominence work group ($M=5.7$) than high-performing male targets in this condition ($M=4.8$). Overall, we find women have more favorable assignment outcomes when they are rare, and as their performance increases (a three-way interaction).

[INSERT FIGURE 3 ABOUT HERE]

In line with our second dependent variable, a logistic regression was performed to ascertain the effects of gender, performance, and sex composition on the likelihood of being allocated to workgroup A or B (results not shown but available upon request). The logistic regression model was statistically significant $\chi^2(7) = 111.75$ ($p < 0.01$). The model explained 22% of the variance (Nagelkerke $R^2$) and correctly classified 69.5% of the cases. Again, consistent with Hypothesis 1, performance has a positive and statistically significant effect of being assigned to a work group with more socially prominent members ($p < 0.01$). Consistent with our prior results we find strong support for the female advantage that we surfaced in Study
1. With the same variables and all of the two-way interactions in the models as shown in the ANOVA, the three-way interaction between gender, performance, and sex composition is positive and statistically significant ($p < 0.05$).

**Qualitative Comments**

The quantitative analysis above provides support for performance affecting allocation to work groups with prominent members, and that women have higher returns to their performance than men. However, this was not true under all conditions, but when women are numerically rare. Further, we find that women are more likely to be allocated to work groups with socially prominent members in general when they are rare. The experiment was not a qualitative study, but we did ask respondents why they decided to allocate Frank/Sarah in the way they did in an open ended question toward the end of the survey. We examine these results speculative here.

The majority of respondents did indicate a reason why they made an allocation decision. After removing “no comment” or nonsensical responses, there were 297 comments made about Sarah and 294 made about Frank indicating a one word or more response about a basis for their allocation decision. Overall, there were 5,712 words written about Frank, and 6,235 words written about Sarah, or an average of 18.7 and 20.2 words per respondent, respectively. In the gender balanced condition Frank and Sarah did not differ much in the average word counts used by respondents, at 19.1 words and 20.2 words, respectively. In the highly-skewed-male condition (across the two performance conditions) Frank had 17.7 words on average written about him per respondent, while Sarah had 20.0. This difference remained consistent when looking at the high performing, male-skew condition, where Frank had 17.5 words per respondent written, while Sarah had 19.8 words written about her, on average.
We read through comments to understand how respondents thought about their allocation decisions. Unsurprisingly, allocators for both Frank (the male condition) and Sarah (the female condition), spoke highly of their abilities when they were in the high performance condition. For example, a quote for Sarah as well as one Frank, were as follows.

*She is described by other clients as a stellar performer, and therefore can be trusted to meet the needs of the most important client.*

*Frank is an exceptional colleague who would thrive with a higher end client.*

One concern we had was that respondents in the highly-male skew condition would view Sarah as having been mistreated somehow (due to her being the only woman to be allocated), and therefore use this to drive their allocation decision. There was practically no evidence this was the case. The only comment along those lines was the following (high performing condition, male-skew).

*Sarah was the only unassigned person, and the only female in the group of employees. She may have been discriminated against by employers and other employees. The existing team is excellent, and can help her raise herself up within the agency and find a way to shine. Client comments indicate she is reliable, which I prize above "flash in the pan." She does good work in a timely manner. She should have a shot at an excellent career move and to earn some respect within the company.*

We look speculatively at qualitative support in line with Hypothesis 1 about performance. We find respondents spoke highly of employees—regardless of whether it was Frank or Sarah—in the high performing condition. Overall, there were similar occurrences indicating that both were viewed as having high quality. Words such as “stellar” and “great” and “top” were used for both. Respondents’ comments reflected the quality of the employee, and was frequently used as a basis
for allocation either to Team A or Team B. That is, in line with Hypothesis 1, the qualitative comments indicated work performance used as a basis for sorting individuals to teams. For instance, in the average performing condition, and the high performing condition, respectively, respondents stated the following about Frank.

*Frank’s score seemed to match better with Team B. He was in the middle range and I feel like Team B is more average.*

*Due to his stellar performance I believe putting him with a team that handles one of the companies biggest clients will not only benefit the company but him as well. It will show appreciation for his hard work and dedication.*

Seeing no quantitative support for Hypothesis 2, we did not explicitly look for qualitative comments to support it. We did explore differences in expectations on teams to see if there was any evidence that respondents spoke different about Sarah and Frank. In comparing word counts, one point of difference were words “lead” and help.” We inspected who respondents attached expectations of leadership and helping behavior to more, Frank or Sarah. Consistent with our expectations, comments about Sarah being viewed as being able to lead teams came up more than they did for Frank. In the experiment, respondents’ answers for Sarah included the word “lead” 24 times, and only 10 times for Frank. Sarah was also associated with “help”-ing a team more than Frank; there were 21 mentions of this for Sarah, while only 14 mentions of this for Frank. For instance, two different comments for Sarah are as follows.

*Sarah clearly has great leadership qualities and would have the opportunity to take these mid-range clients to the next level. You need someone to lead the team. On team A, she would be another high achiever and her skills might not be utilized most effectively. She can help the other employees on Team B to become better and lead them successfully.*
She scored very high and deserves to be at the top. The only reason I would even think to put here in team B is to be a leader.

For both Sarah and Frank, if respondents made statements about leading and helping behavior, they tended to do it more with respect to Team B than Team A. For example, the following are two comments from different respondents about Sarah and Frank, respectively.

**Team A would be the best team for Sarah but putting her on team B to help strengthen it is what would be best for the company. Sarah will be a great leader for team B. I believe she will help them achieve better results.**

**Although Team A is full of high performers, Frank could not only help Team B and their clients earn more or succeed, he could also be a leader and teach the rest of the team the skills they might not have yet to bring every one up a level skill wise.**

To us, favoring Team B with respect to allocating employees with leadership potential, in retrospect seemed natural because respondents had this direct comparison put before them. The team that was most in need of leadership or help was Team B, based on the description of the study. But, it also may have put downward pressure on the allocation of, especially Sarah, to Team B. This may be why we fail to find support for Hypothesis 2.

Finally, to qualitatively explore Hypotheses 3 and 4, we examined how Sarah may have been discussed differently across the low proportion women and balanced women conditions, especially when performance is high. We argued that when women are rare their performance will stand out and be noticed more. In the high performance condition, we coded each comment as whether or not it indicated any of Sarah’s qualities were noticed (i.e even if multiple adjectives were used describing an employee, we coded for the first word only, which is a more
In the balanced-high-performing condition Sarah had 33 comments (out of 84) in which one of more of the words in Table 5 was used to describe her work performance. When women were in low proportion, all else equal, Sarah received more remarks about her work performance. She had 48 comments directly about her skills as an indication of the basis for why she was allocated (out of 79). As a baseline, we show the number of comments Frank’s performance had in the low-proportion-women-high-performing condition: 38 (out of 85). Consistent with our arguments, this provides some suggestive evidence that Sarah’s skills are getting more attention when women are numerically rare than there is more gender balance.

DISCUSSION

The aim of this article has been to develop and test theory on how employees are allocated horizontally, and what if any, influence gender has on this allocation process. We began by speculating that employees’ performance would have a positive influence on allocation, and that there may be differing returns to that performance for men and women when being allocated to groups with socially prominent members. We found, unexpectedly, that those returns to performance did differ, with evidence suggesting that women had an advantage in being allocated to prominent groups at a specific company termed ChemCo. We utilized a quantitative abductive approach to explore the basis of the female advantage and test our theory in a second study in an experiment with 640 working adults asked to perform an allocation task of managers. Overall, we find evidence that performance leads allocators to sort people into

---

8 This is a somewhat conservative comparison, because typically if respondents mentioned one adjective dealing with performance, they mentioned others. This would amplify differences shown in Table 5.
groups with more socially prominent groups, and that this effect is stronger for women under certain conditions, namely when women are numerically rare.

This study makes a number of important contributions. Evidence suggests that in horizontal organizations there is an informal hierarchy of work groups, wherein an employee’s performance influences placement on teams with prominent members. In Study 1 wherein this was a naturally-occurring aspect of the setting, and in Study 2 in which this was incorporated as a part of the design, employees with higher work performance were allocated to work in groups with more prominent members. This suggests that even in so-called flat organizations, we would be remiss to assume that informal hierarchies are not consequential for allocation. Given the rise of these types of organizations, this study suggests the benefits as well as limitations of our existing theories, which have studied allocation primarily into vertical hierarchies, or when horizontal allocation decisions are investigated, they have studied primarily as differences that exist across jobs and roles. Numerous conditions, including changes in the global environment and the life cycle of markets, have altered the structure of post-industrial organizations (Osterman, 1994; Smith, 1997; Vallas, 2003). Flat organizations have long been a part of the organizational typography (Van de Ven, 1976; Mintzberg, 1980; Sundstrom, 1999) but horizontal organizations are increasingly prevalent in the organizational landscape (Purnam et. al 2014; Barley, 2016). Nonetheless, much of the research on allocation continues to focus on vertical sorting within a hierarchy, not the allocation of employees into work groups. Greater attention needs to be paid to horizontal organizations and the ways they structure opportunities (Kalev, 2009; Kacperczyk, Beckman, and Moliterno, 2015).
Along these lines, this is the first paper that has, to the best of our knowledge, investigated the allocation of individuals into work groups with respect to gender. While studies have long investigated internal team dynamics, less is known about how or why individuals are allocated to work groups. Here our focus is on the decisions of allocators that must assemble teams within organizations. We find evidence consistent with the notion that men may be at disadvantage with respect to horizontal allocation, and may have less opportunity to associate with prominent members on teams. Indeed, this effect seems most dominant when women are high performing and rare. This sheds light on a new mechanism that may be at work within flatter, team-based organizations – a female advantage in access to prominent group members—that may have downstream consequences on careers. For example, Kalev (2009) finds that women are more represented in managerial positions when organizations have team-based designs, which she argues is due to the opportunities teams provide individuals to overcome bias and stereotypes. Here, our results are consistent with women having disproportionate access to prominent members compared to men under certain conditions.

To the degree that access to individuals on teams for the development of informal resources and support, our study indicates that team-based organizations may unintentionally create differential opportunities for this access. Scholars have long recognized that opportunities for social interaction are a precursor for the formation of informal networks, which affect the performance of individuals and organizations. However, many questions remains about how social opportunities within organizations arise, including from formal work structures (c.f. Stuart and Sorenson, 2007). Informal networks are particularly important in knowledge work and creative processes, as they necessitate incorporating knowledge from others to gain new insights.
(Perry-Smith and Shalley 2003). These can develop on teams and in work groups as employees focus on interacting around common and shared activities (Feld, 1981). Future studies should investigate if the networks for men and women in flat organizations varies compared to vertical organizations.

This study also makes important contributions to literature on organizational demography. We find that women are favored in their allocation relative to men, and this effect is contingent on the numerical proportion of women in the organization. When women are rare they appear to be more likely to be allocated on teams with prominent members. In this way, this study opens up new questions about organizational demography and managerial decision-making. A great deal of research has focused on numerical composition and its influence on employees’ careers (Kanter, 1977; Turco, 2012). By and large, studies suggest that being in a numerical minority may lead to a lack of benefits for employees. But, to date, studies have focused on numerical composition of organizations as consequential for employee behavior and less so for allocation (see Duguid et. al 2012 and McGinn, and Milkmann 2012 for exceptions). As we motivated our demographic composition argument, we suggested that the demographic composition of the organization not only affects individuals’ performance, but the decisions of allocators. This both coheres to and sheds light on findings in literature on gender and demography. For example, literature focused on “implicit quotas” posits that progress seems to stall for women after a certain point—i.e. most boards only have one woman—in accord with the idea that implicit quotas satisfy concerns for legitimacy regarding diversity (Tinsley, Wade and O’Reilly, 2016; Kalev, Dobbin, and Kelly 2006). Importantly, our research extends these ideas and posits an alternative reason that perhaps implicit quotas emerge. As demographic groups
increase in proportion, when members of these groups achieve high performance it is viewed as less remarkable, and thus may not be rewarded in the same way. Along these lines, there may be a reasonable basis for behaviors whereby women fail to help other women, as perceptions about one’s performance may depend on the numerical representation or the lack thereof of one’s demographic group (Duguid et. al., 2012). Future research should examine this idea.

In the experimental study, we found no evidence that men and women in the experiment allocate employees differently, this nonetheless prompts questions for future research (Cohen and Huffman 2007; Srivastava and Sherman 2015), as this may not be the case under a broad range of circumstances. Finally, future research should inspect how the demographic composition of organizations affects allocators’ decisions, especially when the clarity of employee performance varies. In work groups especially, it may not always be possible to make performance clear and unambiguous. For this reason, the findings that we surface here might be more limited in other settings.

In closing, we also note that we developed more fully a quantitative abductive approach that we feel is quite novel and could be used in a number of ways going forward by other scholars. Here quantitative abduction was a good place to start given our initial pattern of results. This involved starting anew with theory, and taking the theory to a new set of data or a second study. In doing so, we hope we can inspire others of a similar approach to transparency, which can increase trust between researchers and also between researchers and broader audiences (Kerr, 1998; Vazire, 2017).
Table 1. Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Group Member Prominence</td>
<td>0.08</td>
<td>0.30</td>
<td>0.00</td>
<td>8.48</td>
</tr>
<tr>
<td>Average Group Member Prominence</td>
<td>0.05</td>
<td>0.22</td>
<td>0.00</td>
<td>7.58</td>
</tr>
<tr>
<td>Performance</td>
<td>3.98</td>
<td>6.60</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>Female</td>
<td>0.06</td>
<td>0.23</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Patent Class Frequency</td>
<td>2.88</td>
<td>2.58</td>
<td>0.01</td>
<td>8.28</td>
</tr>
<tr>
<td>Scientist Tenure</td>
<td>4.53</td>
<td>5.93</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>Avg Tenure of Group Members</td>
<td>4.45</td>
<td>4.88</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>Number of Other Members</td>
<td>2.15</td>
<td>1.59</td>
<td>1</td>
<td>16</td>
</tr>
</tbody>
</table>

N=8718; Social prominence variables are multiplied by 10^2.
Table 2. Maximum Prominence of Work Group Members

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Patents</td>
<td>0.51 **</td>
<td>0.39 **</td>
<td>0.37 **</td>
<td>0.37 **</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Size of Group</td>
<td>1.88 **</td>
<td>1.88 **</td>
<td>2.45 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.26)</td>
<td>(0.25)</td>
<td></td>
</tr>
<tr>
<td>Patent Class Frequency</td>
<td>0.59 **</td>
<td>0.59 **</td>
<td>0.56 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>Employee Tenure</td>
<td>-0.19 †</td>
<td>-0.18 †</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>Avg Tenure Group Members</td>
<td>1.25 **</td>
<td>1.25 **</td>
<td>1.41 **</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.29</td>
<td>-1.49</td>
<td>-0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.31)</td>
<td>(1.19)</td>
<td>(1.21)</td>
<td></td>
</tr>
<tr>
<td>Female x No. of Patents</td>
<td>0.49 *</td>
<td>0.61 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6.10 **</td>
<td>-3.88 **</td>
<td>-3.85 **</td>
<td>-5.46 **</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.75)</td>
<td>(0.74)</td>
<td>(0.51)</td>
</tr>
<tr>
<td>Year Intervals</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.01</td>
<td>0.07</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>N</td>
<td>8718</td>
<td>8718</td>
<td>8718</td>
<td>8718</td>
</tr>
</tbody>
</table>

Robust standard errors clustered by employee are in parentheses.
All coefficients and standard errors are multiplied by 10^-4.
** ρ < 0.01; * ρ < 0.05; † ρ < 0.10; two-tailed tests.
<table>
<thead>
<tr>
<th></th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>0.53 *</td>
<td>0.28 †</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Female</td>
<td>1.22</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>(1.31)</td>
<td>(1.40)</td>
</tr>
<tr>
<td>Female x Performance</td>
<td></td>
<td>0.64 *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.23)</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.50 †</td>
<td>-5.29 †</td>
</tr>
<tr>
<td></td>
<td>(2.92)</td>
<td>(2.91)</td>
</tr>
<tr>
<td>Scientist Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Patent Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year Dummy Variables</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>1012</td>
<td>1012</td>
</tr>
</tbody>
</table>

All coefficients and standard errors are multiplied by $10^4$; Robust standard errors in parentheses; Significance levels as follows: **p < 0.01; *p <0.05, † p < 0.1  (two-tailed tests)
Table 4. ANOVA of Work Group Suitability

<table>
<thead>
<tr>
<th>Source</th>
<th>F Stat</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>target gender (TG)</td>
<td>2.56</td>
<td>0.00</td>
</tr>
<tr>
<td>target performance (TP)</td>
<td>121.52 *</td>
<td>0.16</td>
</tr>
<tr>
<td>department composition (DC)</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>TG X DC</td>
<td>8.64 *</td>
<td>0.01</td>
</tr>
<tr>
<td>TG X TP</td>
<td>1.97</td>
<td>0.00</td>
</tr>
<tr>
<td>TP X DC</td>
<td>1.36</td>
<td>0.00</td>
</tr>
<tr>
<td>TG X TP X DC</td>
<td>7.13 *</td>
<td>0.01</td>
</tr>
</tbody>
</table>

N=640

$^1 p < 0.10$

$^* p < 0.05$

$^{**} p < 0.01$

Table 5. Comparison of Words Used to Describe Performance of Sarah or Frank

<table>
<thead>
<tr>
<th>Demography: Balanced Group</th>
<th>Demography: Skewed Male</th>
<th>Demography: Skewed Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex: Female</td>
<td>Sex: Female</td>
<td>Sex: Male</td>
</tr>
<tr>
<td>Performance: High</td>
<td>Performance: High</td>
<td>Performance: High</td>
</tr>
<tr>
<td>Top – 2</td>
<td>Performance – 6</td>
<td>Very good – 1</td>
</tr>
<tr>
<td>Qualifications – 1</td>
<td>Quant Score (4.9, 5%) – 6</td>
<td>Top – 2</td>
</tr>
<tr>
<td>Achiever – 1</td>
<td>Rating – 2</td>
<td>Quality – 4</td>
</tr>
<tr>
<td>Rating – 2</td>
<td>Great – 4</td>
<td>Exceptional – 1</td>
</tr>
<tr>
<td>Exceptional – 3</td>
<td>Top – 4</td>
<td>Track Record – 1</td>
</tr>
<tr>
<td>Performance – 4</td>
<td>Exceptional – 4</td>
<td>Rating – 5</td>
</tr>
<tr>
<td>Excellent – 1</td>
<td>High Score (verbal) – 4</td>
<td>Best – 3</td>
</tr>
<tr>
<td>Talent – 2</td>
<td>Best – 3</td>
<td>Stellar – 3</td>
</tr>
<tr>
<td>Stellar – 5</td>
<td>Stellar – 5</td>
<td>Outstanding – 2</td>
</tr>
<tr>
<td>Best – 2</td>
<td>Outstanding – 2</td>
<td>Better – 1</td>
</tr>
<tr>
<td>Driven – 1</td>
<td>Better – 1</td>
<td>Caliber/Quality – 3</td>
</tr>
<tr>
<td>Quant Score (4.9, 5%) – 4</td>
<td>Caliber/Quality – 3</td>
<td>Good – 1</td>
</tr>
<tr>
<td>Strong – 1</td>
<td>Good – 1</td>
<td>Asset – 1</td>
</tr>
<tr>
<td>High Score (verbal) – 2</td>
<td>Asset – 1</td>
<td>Credentials – 2</td>
</tr>
</tbody>
</table>

Total: 33

Total: 48

Total: 38
Figure 1. Effect of Performance on the Work Group Member Prominence
Figure 2. Gender Balanced Condition

Gender-Balanced Condition

Prominent Group Suitability

Avg. Performance  High Performance

Female Target  Male Target
Figure 3. Influence of Performance on Allocation When Women Are in Low Numerical Proportion

Prominent Group Suitability

Low Proportion Women Condition

Avg. Performance

High Performance

Female Target  Male Target
References

Acker, J.

Amburgey, T. and Dacin, T.
1994 “As the left foot follows the right? The dynamics of strategic and structural change.”

Avolio, B. J., and Gardner, W. L.
2005 “Authentic leadership development: Getting to the root of positive forms of leadership.”
The Leadership Quarterly, 16(3), 315-338.

Bass, B. M., and Avolio, B. J.

Barbulescu, R., and Bidwell, M.

Barley, Stephen R.

Baron, J. N.

Baron J. N. and Bielby, W. T.

Baron, J. N., Davis-Blake, A., and Bielby, W. T.

Baron J. N., Mittman, B. S. and Newman, A. E.
Baron, J. N. and Pfeffer, J.

Beckman, C. M., and Phillips, D. J.

Belliveau, M. A.

Bothner, M. S., Smith, E. B., and White, H. C.

Briscoe, F., and Kellogg, K. C.

Caliendo, M. and Kopeinig, S.

Castilla E. J.

Castilla, E. J., and Benard, S.

Chan, C. K., and Antebay, M.

Chandler, A. D.

Cohen, L. E., and Broschak, J. P.
Cohen, L. E., Broschak, J. P., and Haveman, H. A.
1998 “And then there were more? The effect of organizational sex composition on the hiring and promotion of managers.” American Sociological Review, 711-727.

Correll, S. J.

Correll, S. J., and Benard, S.

Cuddy, A. J., Fiske, S. T., and Glick, P.

Dencker, J. C.

Dezső, C. L., and Ross, D. G.

DiTomaso, N., Post, C., Smith, D. R., Farris, G. F., and Cordero, R.


Duguid, M. M., Lewin Loyd, D. and Tolbert, P.S.

Eagly, A. H.

Eagly, A. H.
Eagly, A. H., Johannesen-Schmidt, M. C., and van Engen, M. L.  

Elvira, M. M., and Graham, M. E.  

Ely, R. J.  

Feld, S. L.  

Fleming, L., Mingo, S., and Chen, D.  

Fletcher, Clive.  

Foschi, M.  

Foschi, M.  

Gayle, G., Golan, L. and Miller, R. A.  

Gorman, E.H., and Kmec, J.A.  

Gould, R. V.  
Hall, B. H., Jaffe, A. B., and Trajtenberg, M.

Hargadon, A.
2005 “Bridging old worlds and building new ones: Towards a microsociology of creativity.” In L. L. Thompson and H. S. Choi (Eds.), Creativity and Innovation in Organizational Teams: Lawrence Erlbaum Associates, Inc.

Hargadon, A., and Sutton, R. I.

Haveman, H. A., and Beresford, L. S.
2012 “If you’re so smart, why aren’t you the boss? Explaining the persistent vertical gender gap in management.” The ANNALS of the American Academy of Political and Social Science, 639(1), 114-130.

Hinds, P. J., Carley, K. M., Krackhardt, D., and Wholey, D.

Huffman, M. L., Cohen, P. N., and Pearlman, J.


Kuhn, Thomas.

Joshi, A.

Hochschild, A. R.


Merton R.

Mintzberg, H.

Okhuysen G. A. and Eisenhardt K. M.

O'Reilly, C. A., Caldwell, D. F., and Barnett, W. P.

Osterman, P.

Peirce, C.S.
1878 Deduction, induction, and hypothesis. Popular Science Monthly, 13, 470-482.

Perretti, F., and Negro, G.

Perry-Smith, J. E., and Shalley, C. E.

Petersen, T., Saporta, I., and Seidel, M. D.

Petersen, T. and Saporta, I.

Podolny, J. M. and Phillips, D. J.
Puranam, P., Alexy, O. and Reitzig, M.

Ragins, B. R., Townsend, B., and Mattis, M.
1998 “Gender gap in the executive suite: CEOs and female executives report on breaking the glass ceiling.” Academy of Management Executive, 12, 128–142.

Reagans, R., Zuckerman, E., and McEvily, B.

Reskin, B. and McBrier, D.

Ridgeway, C.


Ridgeway, C. L., and Correll, S. J.
2004 “Unpacking the gender system a theoretical perspective on gender beliefs and social relations.” Gender and Society 18(4): 510-531.

Roberts, P. W., and Sterling, A. D.

Rosenbaum, P. R., and Rubin, D. B.
1983 “The central role of the propensity score in observational studies for causal effects.” Biometrika, 70(1), 41-55.

Rosette, A. S., and Tost, L. P.
2010 “Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders.” Journal of Applied Psychology, 95(2), 221-235.

Rule, L.
Sackett, P. R., DuBois, C. L., and Noe, A. W.

Sonnert, G., and Fox, M.

Smith, V.

Sørensen, J. B., & Sharkey, A. J.

Spence, M.

Spilerman, S., and Petersen, T.

Srivastava, S. B., and Sherman, E. L.

Strang, D., and Baron, J. N.

Stuart, T. E., and Sorenson, O.

Sundstrom E.

Taylor, A. and Greve, H. R.
Timmermans, S., and Tavory, I.
2012 “Theory construction in qualitative research: From grounded theory to abductive analysis.” Sociological Theory, 30(3), 167-186.

Tinsley, C. A., Wade J., Main, B. G., and O’Reilly, C. A.
2016 “Gender Diversity on U.S. Corporate Boards: Are We Running In Place?” Forthcoming, Industrial Labor Relations Review.

Trajtenberg, M. Shiff, G. and Melamed, R.

Tsui, A. S., and Gutek, B.

Turco, C.

Vallas, S. P.

Van de Ven, A. H.

Vazire, S.

Von Nordenflyght, A.

Wallace, J. E., & Kay, F. M.

Williams, J. C., Phillips, K. W., and Hall, E. V.
Wilson, G., Sakura-Lemessy, I., and West, J. P.

Wolfe, J., and Powell, E.

Uzzi, B. and Shapiro, J.

Zellmer-Bruhn, M. E.
2003 “Interruptive events and team knowledge acquisition.” Management Science, 49(4): 514-528
Appendix A – Results and Robustness Checks for Study 1

We use OLS regression analysis to assess the hypothesis. Our OLS results were robust to a number of specifications. In models not shown but available by request, we used the average prominence of group members instead of the maximum prominent of group members as the dependent variable, and the findings remain substantively the same (see Table 1 for the descriptive statistics). Other robustness checks of our results included clustering standard errors on inventors, given that scientists could appear multiple times in our data, and generalized estimating equation (GEE) models. GEEs use maximum-likelihood estimation and permits non-constant variation that corrects for correlation within subjects. GEE models are appropriate for handling correlated data structures (Liang and Zeger 1993; Homish et. al, 2010) that arise from having repeated observations of the same individuals over time. We ran GEE models with robust standard errors (i.e. Huber/White Sandwich Estimators), which permit estimates to be valid even in the event of misspecification of the correlation structure (Homish et. al 2010).

The results for performance and gender were substantively the same at the same levels of statistical significance using GEE models as those we present for OLS (models available upon request). We also did propensity score matching as described in the main text. The kernel density plots are below.
Figure A-1. Kernel Density Plots of the Estimated Propensity Score for Full and Matched Data

Full Data

Matched Data