The Effect of Gender Composition of Research Teams on Individual Researchers' Performance in China†

Ying Ma*

Abstract

This article explores the relationship between the gender composition of scientific research teams and scientists' individual performance. The gender composition of research teams is an important feature of workplace settings and influences the way people interact and communicate; however, previous research has not directly examined its relationship with scientists' individual performance. Drawing on data collected on university faculties in China in 2016, this article tests several hypotheses about individual's performance in teams with different gender compositions. The results show that team gender composition has a clear gendered effect on scientists' individual performance. The effects of tokenism for women in men-majority teams is proven, but men in women-majority teams appear to be unaffected by tokenism. Moreover, the theories claiming that homogenous teams are more conducive to better individual performance than mixed teams are supported for men but not for women. The findings of this research suggest that recruiting more women into the scientific workforce may improve their performance and thereby help diminish the gender gap in performance. It also indicates that the Chinese preferential policies towards women in science formulated in recent years have had positive impacts. However, considering that more than half of the researchers in the survey are working in men-majority teams, the task of narrowing the gender gap in performance remains a challenge. Further work is needed to explore the tensions and benefits of working with the opposite gender.

Keywords
team, gender composition, performance, individual, China

+ The author thanks Prof. Elaine Weiner for her guidance and valuable comments in the writing of this article and Robert Jones for his editing work. The opinions reported herein are those of the author alone. This research was supported by the Science and Technology Policy Institute (STEPI) (Korean) fellowship, the China Scholarship Council, the Merit Scholarship Program for Foreign Students (2016) (Quebec, Canada), National Natural Science Foundation of China (NSFC)'s Young Scientists Fund "The impact of the Youth Science Fund project to scientists' academic and occupational development" (71503239), and NSFC's policy project, "Policy evaluation on NSFC's policy to promote the development of women scientists" (L422062).

* Postdoctoral fellow, McGill University, Montreal, Canada; Associate professor, Chinese Academy of Science and Technology for Development, Beijing, P. R. China, may@casted.org.cn.

1. INTRODUCTION

In recent years, promoting women's participation in science has become an important policy concern for leading countries in science and technology. More trained scientists and technologists are needed to accelerate discovery and innovation, and there is also a serious gender gap favoring men in scientific performance and achievement. This gap is a major obstacle to attracting women into a career in science as well as retaining them.

A lot of research has been done to identify the factors contributing to the gender gap in science, but few researchers have looked into the possible effects of the gender composition of research teams. Given that so many scientists work in teams, it is necessary to understand how a team's gender composition is related to scientific performance and productivity. There is now some research available on the effect of gender composition on team performance among research and development (R&D) teams, but few studies target individual performance. Moreover, there has been no research on the effect of teams' gender composition on individual scientists' performance in an academic setting.

This study examines the effect of team gender composition on individual scientists' performance and serves the following purposes:

I. Provides a new explanation for the gender gap in scientific production and achievements—a puzzle that has not been adequately explained. For example, a research team aims to jointly produce scientific output, but it may have disparate returns for members of different genders. This is a perspective not explored by previous research, and this article hopes to fill the gap.
II. Explores how the increase in the number of women in science transforms the gender dynamics in science. As more women enter into the scientific field, the prejudice, discrimination, tension, conflict, cooperation, and sharing between men and women may take on new forms. Will the increase in female scientists change the male-dominant culture in science? How does gender composition relate to scientists' individual performance and success? Those questions have both policy and theoretical implications and need to be addressed.
III. Provides policy suggestions for promoting the development of female scientists. Previous polices supported individual scientists by providing resources or opportunities to women such as maternity leave and special funding. There are also policy measures that seek to transform the institutional culture and structure such as the ADVANCE (Increasing the Participation and Advancement of Women in Academic Science and Engineering Careers) program implemented by the NSF (National Science Foundation) in the U.S. This article provides policy suggestions through highlighting the research team as an intermediate target for policy intervention.

To achieve the aforementioned goals, this article employs data collected in a survey sponsored by National Natural Science Foundation of China (NSFC) in 2016 to evaluate preferential policies towards women. The sample is composed of both male and female scientists. The examination of Chinese scientists not only provides a compelling case for how team gender composition influences Chinese scientists' performance, but also furthers our understanding of the role of gender composition in scientific performance generally.

2. LITERATURE AND HYPOTHESES

2.1. Gender Gap in Scientific Production and Team Gender Composition as a Contributing Factor

The gender gap in scientific production and achievement—i.e., the fact that women produce fewer scientific publications and attain lower positions in science than men—has been widely acknowledged and is viewed as a major indicator of gender inequality in science. A substantial amount of scholarly work has been done to explain the causes and mechanisms of this phenomenon. Individual characteristics such as intelligence, age, marital status, collaboration with a mentor, prestige of one's PhD program, and prestige of one's current institution have all been examined (Long & Fox, 1995). Researchers have also looked at structural positions and resource variables such as availability of research funding and research assistance, teaching load, and child care responsibilities (Xie & Shauman, 1998). However, it is still unclear why characteristics that are conducive to academic production differ between men and women.

Some researchers have highlighted the characteristics of academia as an important institutional setting in explaining the gender gap in science (Allison & Long, 1990; Fox, 1991; Fox & Mohapatra, 2007). As Fox (1991) stated,

> …men share traditions, styles, and understandings about rules of competing, bartering, and succeeding. They accept one another, support one another, and promote one another. As outsiders to this milieu and its bartered resources, shared influence, and conferred self-confidence, women are shut out of ways and means to participate and perform. (p. 194)

Following this acknowledgment of men's predominance as an important feature of science, a method to measure this feature is needed. Research in the field of industrial-organizational science has developed to look at gender composition of workplaces as an organizational feature that conditions the attitudes and behaviors of its members (Pfeffer, 1985; 1991; Tsui, Egan, & O'Reilly III, 1992). This approach provides a more tangible way to measure male dominance in science while providing room for theories focused on group dynamics to offer explanations.

Contemporary science is rarely performed by a solo scientist. Since World War II, with the transformation from little science to big science (Price de Solla, 1963), teamwork has become the "new paradigm" for the organizational structure of research (Beaver, 2001, p. 368). Beaver notes that this paradigm allows scientists to complete tasks in a more efficient and timely manner and cover a broader array of research problems while energizing and exciting participants with a multiplicity of viewpoints. He also notes that teamwork in scientific research reduces risks, as multiple projects and experiments are performed simultaneously, and that the accuracy and visibility of research findings can be increased by virtue of the different perspectives and activities.

According to Beaver (2001), it is easy to perceive a research team as a structure or context for science in that the team provides both the material resources, such as equipment, data, and money, and the colleagues to communicate and exchange ideas. Previous studies on the gender gap in science examined institutional considerations such as the departments and found a significant relationship between departmental prestige and scientists' production (Allison & Long, 1990; Nakhaie, 2002). However, a research team is more than just an organizational structure or merely an environmental context. A team affects the way individual scientists conduct research in the sense that they have to rely upon and collaborate with other team members to achieve a common goal in research. This underlines the importance of intra-group relationships. Previous research focusing on the relationship between scientists' collaboration and production mainly focused on co-authorship. Collaboration is found to increase the quality and frequency of publications, but not the quantity of papers fractioned by number of authors (Hollis, 2001; Lee & Bozeman, 2005).[1] In a study on collaboration in political science, Fisher, Cobane, Vander Ven, and Cullen (1998) found that cross-sex collaboration is the predominant form of authorship for women, while the two most common forms of authorship for men are single authorship and same-sex collaborative authorship. The authors question how cross-sex collaborative articles would be evaluated and how credit is distributed among male and female authors, and they caution that these collaborations may not favor women's academic development. Indeed, their concerns are validated by a recent study on academic economists showing that co-authorship has a penalty for women but not for men (Sarsons, 2015). Considering that most contemporary scientific findings are a product of teamwork, it is important to now consider how research teams involve the interplay between structure and intra-group relationships. This is a necessary step to understanding the gender gap that is present in scientific achievement to date.

2.2. Gender Composition of the Workplace and Its Effect on Group and Individual Performances

As more women enter into workplaces that are traditionally male-dominated, it is important to know how this change affects performance, especially for organizations that rely on superior performance to survive in the market. Research answering this question often uses the term "gender diversity," as the targeted workplaces are usually transforming from male-homogenous to gender-mixed/diversified.

There are both permissive and optimistic views on the effect of gender diversity. They draw on dif-

[1] There are two ways to count the quantity of publications. One way is to count any paper with the author's name as "1," and the other is fractioning by the number of authors. For example, if one paper has two authors, it is counted as 0.5 papers for each author.

ferent theories and focus on different meanings attached to the characteristic of gender. The permissive view is based on similarity-attraction theory and social identity theory (i.e., self-categorization theory) (Byrne, 1971; Newcomb, 1961; Rosenbaum, 1986; Tajfel & Turner, 2004; Turner, 1985). It postulates that as gender is a salient and distinctive characteristic, it can be easily provoked in the process of self-categorization, which means gender diversity is often related to negative group processes such as low cohesion, conflict, and high rates of turnover (Jehn, Northcraft, & Neale, 1999; Pelled, 1996; Mannix & Neale, 2005). The optimistic view, also called the "value in diversity" perspective, is based on the information-processing approach (Ancona & Caldwell, 1992; Mannix & Neale, 2005; Winquist & Larson, 1998). It argues that because gender diversity is used as a proxy for diversity in backgrounds, networks, information, skills, and experience that are meaningfully related to cognitive processes of the group, gender diversity enhances a group's ability to perform tasks, such as solving problems and generating more profits and earnings (Cox, 2001; Mannix & Neale, 2005; Williams & O'Reilly, 1998).

Some researchers have tried to reconcile these two models. Herring (2009) showed that gender diversity is positively related to organizational sales revenue, number of customers, and profit in a representative sample of U.S. profit-making work organizations. He suggests that although diversity may be associated with more group conflict, this conflict may actually be an indication of contestation of different ideas, more creativity, and superior solutions to problems, which can lead to better organizational performance (Herring, 2009).

Some researchers emphasize the importance of context to understanding the different effects of gender diversity (Joshi & Roh, 2009; Mannix & Neale, 2005). Joshi and Roh (2009) tackled the inconsistency in the effect of team diversity on team performance by examining the moderator variables between the two; they performed a meta-analysis on 8,757 teams from 39 studies conducted between 1992 and 2009. Their analysis found that team characteristics, occupational demography, and industry setting moderated the effect of team gender diversity. Specifically, team gender diversity had more negative effects on performance in occupations dominated by men[2] than in gender-balanced occupations, and also had a negative effect in high-tech industry settings. The authors argue that extra-organizational context is important in shaping diversity outcomes because of the stereotypes associated with underrepresented groups and an "expertise advantage" linked to the dominant group. Regarding team characteristics, gender diversity showed a negative impact on team performance in moderate and high interdependence teams[3] and in long-term teams,[4] but a positive impact in low-interdependent and short-term teams. Although Joshi and Roh's (2009) meta-analysis was conducted on teams in organizational settings, the conditions associated with a negative impact of gender diversity all apply to teams in academic settings, which are characterized by a male-dominated occupation, the performance of interdependent tasks, and significant longevity.

---

[2] Measured by the occupation's female percentage being below the overall mean of female percentage for all U.S. occupations (46.3%).

[3] In their definition, R&D teams are high interdependent.

[4] Two years is considered a long time by their definition.

Another approach used to identify the effect of gender diversity is to distinguish different types of performances. For example, in a study examining the innovation of R&D teams in Spanish firms, Díaz-García, González-Moreno, and Jose Sáez-Martínez (2013) distinguish two types of innovation—incremental innovation and radical innovation—and find that gender diversity has a positive effect on radical innovation but no effect on incremental innovation. However, using the same database, Garcia Martinez, Zouaghi, and Garcia Marco (2016) found almost opposite results. They demonstrated that gender diversity is positively related to both incremental and radical innovation in manufacturing firms and that the effect on incremental innovation is larger than on radical innovation. Although the contradictory results may be due to the different research designs they employ (one is cross-sectional and the other uses lagged-variables models), the inconsistent findings warrant further exploration.

Compared to the wealth of research on the effects of gender diversity on group performance, there is a relative dearth of research on the effects of gender diversity on individual performance. Among studies that do examine individual performance, many have been conducted in non-academic settings. One study on female employees in a U.S. federal bureaucracy found that women receive less support from men as their proportion in the work group increases (South, Bonjean, Markham, & Corder, 1982). Indeed, as numerical minorities in a group, women are subject to performance pressure and often perform less well than they would have otherwise. This phenomenon, known as "tokenism," has been observed in several situations: For example, in female employees in a Fortune 500 industrial firm, female litigators in law firms, and female students in two law schools (Kanter, 1977; Pierce 1995; Spangler, Gordon, & Pipkin, 1978). However, other research has found that tokenism only applies for token women in men-majority situations; in fact, token men in women-majority situations not only tend to escape penalties but may even benefit from preferential treatment from coworkers (Alexander & Thoits, 1985; Williams, 1992).

In sum, the bulk of the research on the effects of gender diversity are in the field of organizational studies and are centered on the effects of gender diversity on group performance, with the group defined as R&D teams, organizational branches/departments, or even the organization itself. The reason for the focus on group performance may be due to their salience to market competition. When individuals are considered in gender diversity studies in companies and firms, the focus is more on behavioral and social-psychological consequences such as job satisfaction, organizational attachment, and turnover (Allmendinger & Hackman, 1995; Fields & Blum, 1997; Martin & Harkreader, 1993; Tsui, et al., 1992; Warton & Bird, 1996) than individual performance.

Scientific academia is distinct from many organizational settings in terms of the goals being pursued, rules (i.e., how things should be done), and the evaluation criteria. Most importantly, individual scientists' achievements are considered the key to scientific advancement (Merton, 1973). To explore the effect of gender diversity on individual performance in the setting of academic research teams, we need research that directly addresses this issue.

2.3. The Academic Setting and the Individual Academic Performance

Scientific academia is a field that is traditionally male-dominated, and there are few studies that have examined the consequences of gender composition in this setting. Henderson and Herring (2013) looked into the effect of gender composition of faculty and students on departmental rankings in a sample of 5,000 doctoral programs at various U.S. research universities. After controlling for relevant variables, they found that for every unit increase in the faculty gender diversity indicator, program rankings increase by 14.44 points.[5] They concluded that the results supported the value-in-diversity perspective. However, because they did not show the actual percentage of women in the faculty—only percentage changes—and because they only tested the linear relationship, it is not known if the relationship between percentage of female faculty and program ranking would change beyond some threshold.

Tolbert, Simons, Andrews, and Rhee (1995) examined the relationship between departmental gender composition and faculty turnover in 50 academic departments form 1977 to 1988. They found that before the proportion of women in a department reached a threshold of about 35% to 40%, turnover among women increased but then began to decline after that threshold. The proportion of women had a negligible or negative impact on turnover among male faculty. Taking turnover as an indicator of inter-group relations, they concluded that the results supported the argument that women's growing representation in work groups can lead to an increasingly negative environment for them. Their results also indicated that men and women were affected differently by the increasing proportion of women in the work place. However, as they did not distinguish between voluntary and involuntary leave, it is unclear how the increasing proportion of women affects group relations and what its consequences for men and women may be.

At the team level, Hinnant, Stvilia, Wu, Worrall, Burnett, Burnett, Kazmer, and Marty (2012) performed an analysis on research teams conducting experiments at the National High Magnetic Field Laboratory (MagLab) in the U.S. They investigated whether team diversity and network characteristics affect team productivity. They found no significant relationship between gender diversity and team productivity. However, female scientists comprised only 9% of the total number of scientists in their sample.

The dearth of studies that examine the effect of gender composition in the academic setting focusing specifically on individual performance is important because academia values individual credentials and achievements. The core academic performance indicators, such as number of publications, number of patents, and number of grants, are crucial for individual researchers' career outcomes, and they are vital to understanding why it appears women's performance lags behind men's. In a workplace setting such as an academic team, there are explicit and implicit rules and politics on things like how tasks are prioritized and distributed, how people talk and relate to each other, how

[5] Rankings range from 0 (lowest ranking) to 100 (highest ranking).

consensus is achieved and disagreements are expressed, how leaders can incentivize team members, and how team members can request benefits. All these practices and behavioral patterns may have gendered effects, considering the prevailing gender stereotypes in science and society. As these gendered effects are usually taken for granted and accepted unconsciously, people tend to ignore their role in sustaining the gender gap in scientific academic performance. In addition, as teamwork is the preponderant way of working in science, only those able to adapt to teamwork may survive and thrive. In this way, the doctrine of teamwork is self-perpetuating, making it an obscure yet compelling force in shaping the map of gender disparity in science.

2.4. The Case of China

China has become an important actor in the international science arena. With globalization and internationalization, the practices of Chinese science are increasingly in line with international practices. Gender relations in science in China also resemble the situation in many other countries, whereby science is a male-dominated world and women continue to shoulder a disproportionate share of family responsibilities (Zhu & Ma, 2015).

In recent years, Chinese policy makers have started to focus more on supporting the development of female scientists. In 2011, the Ministry of Science and Technology and the National Women's Federation jointly enacted a policy document entitled, *On Promoting the Construction of Talented Women in Science and Technology*. This was the first policy document ever formulated in China to promote the development of female talent in science and technology. The document recommended that science and technology administrative departments and women's federations at all levels support women's participation in employment, training, development, and administration in the fields of science and technology. More concrete measures to promote the development of women scientists were also undertaken by the National Natural Science Foundation of China (NSFC). As early as 2010, the NSFC recommended in their project review guideline that "in all types of project review meetings, pay attention to the principle of 'female priority under the same conditions'." In 2011, the NSFC extended the maximum age to apply for the "Young Scientist Fund" from 35 to 40 for women (the maximum for men is 35), stated in its project guidelines that female scientists may extend their project term due to pregnancy and giving birth, and asked for a gradual increase in the number of female scientists on review panels. In a new fund set up by the NSFC in 2011, the "Excellent Young Scientists Fund," the maximum age for men is 38 and 40 for women (Ma, 2017). After several years of implementation, the effect of these policy measures requires evaluation, and new evidence is needed to further formulate policy measures.

2.5. Hypothesis

A research team in an academic setting is generally located in a male-dominated occupation with interdependent team members working together for a sustained period of time. As concluded by Joshi and Roh (2009), in teams with these characteristics, gender diversity may have negative effects on team performance. However, the studies reviewed in this article showed negative, positive, and neutral relationships between gender composition and performance. To formulate the hypotheses of this article, I draw on theories employed in the studies on gender composition in workplaces.

Self-categorization theory—previously known in the literature as social identity theory—states that maintaining high levels of self-esteem and a positive self-identity is important for individuals. Individuals gain their social identity by categorizing themselves into a social group. Then, in an effort to enhance their self-image, individuals tend to value those who share their identity and discriminate against those who do not (Byrne, 1971; Tajfel & Turner, 2004). Researchers have tested this theory with respect to gender categorization and found that people are more satisfied when working with the same gender, and that tensions and negative consequences rise with an increase in people of the opposite gender (Garcia Martinez, et al., 2016; Tsui, et al., 1992).

Self-categorization theory proposes that there are benefits to working in a team that is gender homogeneous. It also suggests that working in a team with heterogeneous gender composition is not good for performance (Garcia Martinez, et al., 2016). This line of research portrays intergroup relationships from the point of view of conflict: the minority is seen as a threat to the possession of scarce resources by the majority, and as the number of minorities increase, so does the threat and the resultant intergroup conflict (Blalock, 1957; South, Bonjean, Markham, & Corder, 1983). Following these theories, I propose that an individual's performance will decrease as team gender heterogeneity increases, regardless of that individual's gender. *H1. For men and women, working in teams with their own gender as the majority is more conducive to their individual performance than working in teams of mixed gender.*

According to intergroup contact theory, with an increase in the numbers of the minority, there will be more cross-group interactions and thus less stereotype and prejudice, which would lead to more harmonious intergroup relations (Blau, 1977; Tolbert, et al, 1995). The minority would be under less performance pressure and encounter less hostility, which would enhance their performance. Kanter's (1977) theory of tokenism is the most referenced in describing the minority gender situation. She proposed that when women are in a minority role, higher visibility, contrast, and assimilation lead to women being stereotyped, facing higher interactional pressure, and having access to fewer resources and prospects for achievement and career advancement. Men, however, take advantage of their dominant position, and the situation becomes self-perpetuating. Her theory underlines the importance of numerical representation in a group, and thus tokenism is "gender neutral" and can be applied to both men and women. Given Kanter's theory of tokenism, I propose that: *H2. For men and women, working in a team where they are a minority is less conducive to individual performance than working in a more gender-balanced team.*

However, some researchers find that Kanter's theory of tokenism may not apply for men. Williams (1992) demonstrated that in female-dominated occupations, although men do face prejudice and discrimination from the outside world, their work environment is not oppressive. They are fairly treated by their bosses and colleagues and sometimes even have a subtle advantage in their career advancement, which Williams coined as the "glass escalator." Rather than explain her work as an exception to Kanter's theory of tokenism, Williams attributes her findings to the gendered nature of the structure of work—something that Kanter failed to account for. The gendered workforce bestows men with advantages regardless of the work environment they are in, and it also puts a higher value on work done by men. Budig (2002) examined the effects of different gender compositions of U.S. occupations on wage and wage growth in 1982–1993. His findings support Williams' claims about gendered work forces in that he found that both male-dominated occupations and men have advantages over mixed and female-dominated occupations and women. He did not find the effect of tokenism for either men or women. He argued that tokenism was a result of improper comparison in previous studies that only looked into settings with tokens while forgetting to consider situations where the concerned group is not a token. That being said, he limited his criticism of tokenism to occupational pay and wage growth. This article's third hypothesis is: *H3. Men are not disadvantaged when they are a minority on a team.*

3. DATA AND METHODS

3.1. Data

The data used in this article comes from a survey conducted by the Chinese Academy of Science and Technology for Development in 2016, sponsored by the NSFC. The NSFC is the largest funding agency in China that supports the development of basic sciences. In 2015, it supported more than 30,000 projects with over 22 billion RMB. The NSFC collaborates with its cooperating work units (similar to proposing organizations of NSF in the U.S.), which mainly includes universities and research institutes, to organize project application and project management. There were over 2,000 cooperating work units in 2015, which covered almost all universities and research institutes in China that had the ability to conduct innovative basic science research. The targeted population in this survey was all the researchers in the cooperating work units, which included project awardees, applicants, and potential applicants to the NSFC.

The survey employed a two-stage cluster sampling design. First, among work units cooperating with the NSFC, the survey team selected 135 work units following the probability proportional to size (PPS) method. Second, they selected three departments or schools from work units with more than 500 researchers. In work units with less than 500 researchers, there was no second stage sampling. The survey team sent links to an online questionnaire via email to all researchers in the selected departments, schools, and work units. In total, 12,040 questionnaire links were sent and 5,802 questionnaires were completed.

As the selected work units were composed of research institutes and universities of different levels, this article includes only the "985" universities.[6] The "985" universities originated from a Chinese

[6] There are 39 "985" universities in China in 2016. The category "985" university used in this article includes 12 "985" universities and the University of Chinese Academy of Science (UCAS). UCAS started recruiting undergraduate students in 2014.

government project to build up world-class universities in 1998, and they comprises the best universities in the country. The selection of only the "985" universities in this article is to create more homogeneous evaluation criteria and competition environments for the researchers in the analysis. In addition, the analysis only included researchers who worked in a stable team and were not team leaders. The reason for this is that team leaders have much more power over the distribution of team resources and output relative to other team members. In the sample, there are 447 team member cases.

Dependent variables
This article uses two variables to measure scientists' individual performance: the number of SCI and EI[7] articles published and the number of grants obtained for each scientist in the last three years before the survey. The number of articles published is an important and commonly used indicator for measuring scientists' productivity. In the Chinese scientific community, there is a tendency to value articles published in international indexed journals as more important and credible than articles published elsewhere. Therefore, this article uses the number of SCI/EI articles to indicate scientists' academic production. Obtaining grants is also an important part of academic work. It is closely related to publication, as previous publication increases the chances of getting grants, and in return, receiving grants funds publishable research. Both publications and grants are crucial for the career development of scientists, especially in universities with a strong focus on research and academic competence. In the sample, the average number of articles published is 7.53 (S.D.=9.89), and the average number of grants obtained is 3.41 (S.D.=2.80).

Independent variables
In the survey, each respondent was asked, "Do you work in a stable research team?" Those who answered, "Yes, I am a team member," were asked about the number of persons the team had and how many of them were women. The gender composition of the team is calculated as the percentage of women in the total number of team members. If less than 30% of the members are women, the team was defined as "men-majority," if 31% to 60% are women, it was defined as "mixed," and if 61% and above are women, it was defined as "women-majority." 54% of the teams were men-majority, 35% were mixed, and 11% were women-majority.[8]

Controlled variables
I. Gender: women were coded as 1 and men were coded as 0. In the sample, 33% of researchers are women.

---

7 SCI articles are those published in Scientific Citation Index journals, while EI articles are those published in the Engineering Index journals.

8 In previous research, different percentages (such as 30%, 20%, 15%) have been used to make the thresholds for majority-minority categorization. In this article, 70%+ of the members being men was classified as men-majority team and 60%+ of the members being women was classified as women-majority team. The reason is that there are few scientists in the teams with large percentages of women. In this sample, only 11% of the scientists worked in teams with more than 60% women, which amounted to only 50 cases. Among the 50 scientists, only 4 were male and 46 were female. There were no male scientists in teams with at least 70% women. The author tested the thresholds of 30%, 20%, and 15% using the same regression. The results were similar with the current threshold when there were valid cases.

II. Age: Previous studies have found that scientists' production has a life cycle and is closely related to their age and years of experience in work. As age and years of experience directly correlate, age is used in this article to control the effect of age and years of experience. The average age in the sample is 38.59 with a standard deviation of 6.09.

III. Doctoral degree: Those with doctoral degree were coded 1, and others were coded 0. 95% of the researchers have a doctoral degree.

IV. Type of institute of highest education: In China, universities and research institutes directly administered by the central government are considered to be of higher prestige. Those who graduated from more prestigious universities and research institutes were coded 1, and others were coded 0. 83% of the researchers graduated from highly prestigious universities and institutions.

V. Size of the team: The size of the team may relate to production (Fox & Mohapatra, 2007). The total number of team members in the multivariate analysis is taken as a controlled variable. The average size of the team is 7.93 persons.

VI. Children: Family responsibility has been an important factor influencing the gender gap in science (Xie & Shauman, 1998). Many articles indicate that it is children rather than marriage that impede women's scientific performance. In this article, the child variable is coded 1 if the researcher has one or more children and 0 if they do not. 83% of the researchers have children.

VII: Discipline: The disciplines of the researchers were categorized into math (6.7%), physics (7.4%), chemistry (9.0%), life science (13.4%), engineering (49.2%), agriculture (2.7), medicine (4.0%), management (2.9%), and other (4.7%). Math is the reference group.

Compared with the non-"985" university researchers who worked as a team member in the survey, the "985" researchers were more likely to have doctoral degrees and to have graduated from high-prestige universities. They also published more SCI/EI articles and obtained more grants. Meanwhile, a smaller percentage of them were women, and the teams they belonged to were more likely to be male-majority than female-majority. All these differences are statistically significant at the 0.001 level. The above comparison shows clearly that "985" university researchers had stronger academic backgrounds and better academic performances. However, they were also more likely to be men, which is an indication of the gender disparity in academia in China.

3.1. Methods

To examine the relationship between gender composition of teams and scientists' performance, this research utilized three steps. First, the bivariate relationship between team gender composition and individual scientists' performance will be examined. This will indicate what type of team is associated with better performing scientists. Second, the multivariate relationship between team gender composition and individual scientists' performance will be examined to see whether the pattern detected in the bivariate relationship still holds when compounding factors are controlled. Third, in the negative binomial model, the interaction between group composition and an individual's gender will be added to see whether men and women performed differently in environments with different team gender composition. STATA 13 software was used for performing the statistical analysis.

4. RESULTS

4.1. Bivariate Analysis

A one-way ANOVA was conducted to determine if scientists' individual performance was different in teams with different gender compositions. The differences are small or negligible. For the number of SCI/EI articles, there was a statistically significant difference at the 0.10 level between groups as determined by one-way ANOVA ($F(2,442)=2.90$, $p=.06$). A Tukey post-hoc test revealed that number of articles was higher (significant at the 0.10 level) in the men-majority team compared to the mixed team (2.31± 1.01 articles, $p=.06$). However, there were no statistically significant differences between the other pairs of groups. For the number of grants, there was no statistically significant difference between teams of different gender compositions as determined by one-way ANOVA.

TABLE 1. Mean Differences in the Number of SCI/EI Articles and Grants Between Teams of Different Gender Compositions.

	Number of SCI/EI Articles	Significance(*p*)	N	Number of Grants	Significance (*p*)	N
Men-majority (1)	8.58	1 vs. 2: 0.06	238	3.61	1 vs. 2: >=0.10	217
Mixed (2)	6.27	2 vs. 3: >=0.10	157	3.12	2 vs. 3: >=0.10	142
Women-majority (3)	6.52	3 vs. 1: >=0.10	50	3.43	3 vs. 1: >=0.10	44

4.2. Multivariate Analysis

The negative binomial regression model was used to examine the effects of team gender composition on the number of SCI/EI articles published and the number of grants, as the distribution of the two are over-dispersed (Hilbe, 2011) (articles: mean=7.53, S.D.=9.89; grants: mean=3.41 S.D.=2.80). The regression results are shown in Table 2. Model 1 and Model 3 are the results of the regressions for the number of SCI/EI articles and grants, respectively. The likelihood-ratio tests of alpha for both models are significant, indicating that the negative binomial regression is more appropriate than Poisson regression.

The pattern found in the bivariate analysis was replicated—scientists working in mixed teams published significantly fewer articles than those working in men-majority teams. Using the margin command in STATA 13, the results show that the average number of published articles for researchers in mixed teams is 6.27, while for researchers in men-majority teams and women-majority teams the figures are 7.91 and 7.98, respectively; the difference is more than 1.5 articles. After controlling for the difference in education, age, type of institution graduated, size of the team, children, and discipline[9], scientists working in women-majority teams published about the same as those working

[9] The author controlled the effect of the gender of the team leader in an earlier version of the analysis. The results are similar to those presented here.

in men-majority teams. In terms of the number of grants received, there is no significant difference between teams of different gender composition. On average, researchers working in men-majority teams received 3.43 grants, 3.27 grants in mixed teams, and 3.84 grants in women-majority teams.

The results show that female scientists published significantly fewer articles than male scientists, and those with a doctoral degree published more than those without a doctoral degree. For the number of grants, except for the discipline dummy variables, only the doctoral degree variable is positively related to the number of grants obtained.

TABLE 2. Negative Binomial Regression on the Number of SCI/EI Articles and Grants

	Model 1	Model 2	Model 3	Model 4
VARIABLES	Number of Articles	Number of Articles	Number of Grants	Number of Grants
Mixed	-0.232* (0.112)	-0.351** (0.125)	-0.047 (0.087)	-0.168 (0.099)
Women-majority	0.009 (0.191)	-0.591 (0.500)	0.112 (0.145)	-0.570 (0.420)
Female	-0.277* (0.125)	-0.684*** (0.206)	-0.143 (0.096)	-0.586** (0.179)
Mixed#female	–	0.574* (0.260)	–	0.604** (0.217)
Women-majority#female	–	1.013 (0.550)	–	1.144* (0.465)
Age	0.008 (0.008)	0.007 (0.008)	0.007 (0.006)	0.005 (0.006)
Doctoral Degree	0.996*** (0.242)	0.990*** (0.241)	0.418* (0.200)	0.396* (0.199)
Type of Institute of Highest Education	-0.021 (0.128)	-0.010 (0.129)	0.069 (0.104)	0.057 (0.103)
Size of the Team	-0.004 (0.007)	-0.001 (0.007)	0.008 (0.005)	0.011* (0.005)
Children	-0.221 (0.119)	-0.210 (0.118)	0.180 (0.106)	0.164 (0.105)
Physics	0.581* (0.230)	0.576* (0.229)	0.385 (0.226)	0.361 (0.225)
Chemistry	0.857*** (0.221)	0.860*** (0.220)	0.806*** (0.209)	0.796*** (0.208)
Life Science	0.167 (0.210)	0.165 (0.208)	0.548** (0.205)	0.549** (0.204)
Engineering	0.237 (0.180)	0.265 (0.179)	0.846*** (0.183)	0.859*** (0.182)
Agriculture	-0.728* (0.342)	-0.714* (0.340)	0.639* (0.286)	0.618* (0.283)
Medicine	0.251 (0.287)	0.292 (0.286)	0.790** (0.252)	0.809** (0.249)
Management	-0.262 (0.312)	-0.265 (0.311)	0.563* (0.277)	0.575* (0.276)

Other	0.575* (0.258)	0.564* (0.256)	0.812*** (0.233)	0.799*** (0.231)
Constant	0.794 (0.467)	0.829 (0.464)	-0.387 (0.387)	-0.289 (0.387)
/lnalpha	-0.439*** (0.082)	-0.455*** (0.083)	-1.817*** (0.182)	-1.881*** (0.188)
Observations	444	444	402	402
Log Likelihood	-1314.534	-1311.517	-849.780	-844.366
Pseudo R2	0.028	0.030	0.032	0.038

Standard errors in parentheses
**** p<0.001, ** p<0.01, * p<0.05*

4.3. Interaction

To test if men and women perform differently in teams of different gender composition, an interaction term was added between team gender composition and the gender of the researcher in the regression. The results are shown in Model 2 and Model 4 in Table 2. In both models, the interaction terms are significant. Using the margin plot command in STATA, the interaction effect in Figure 1 and Figure 2 is drawn for a better representation of the interactive relationships between team gender composition and gender.

FIGURE 1. Interaction Effect of Team Gender Composition and Individual's Gender on the Number of SCI/EI Articles

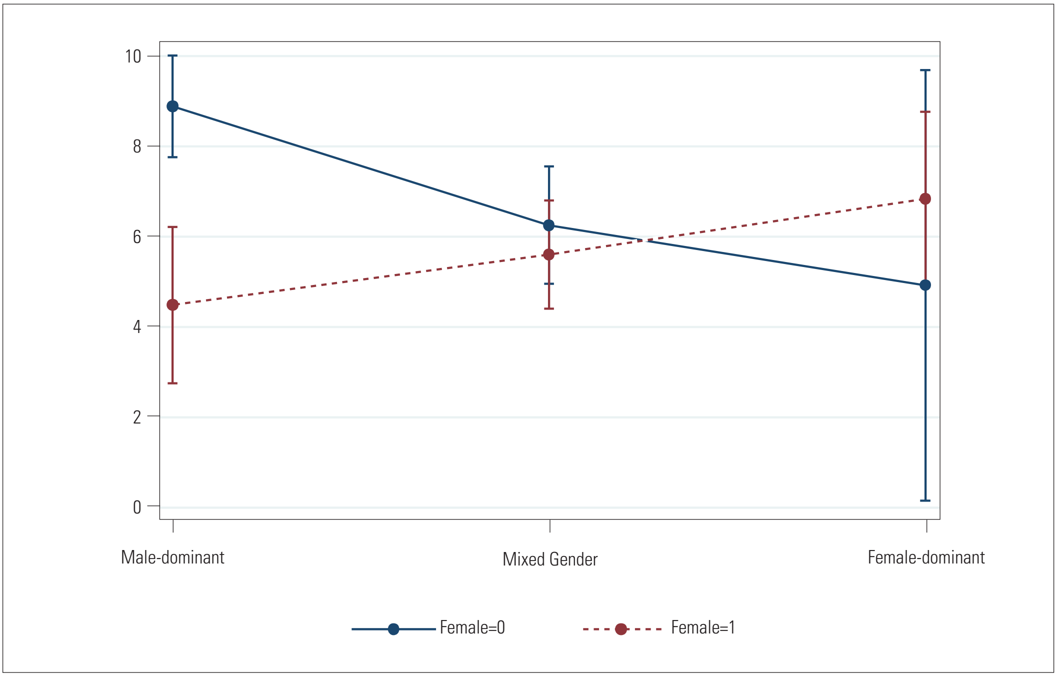
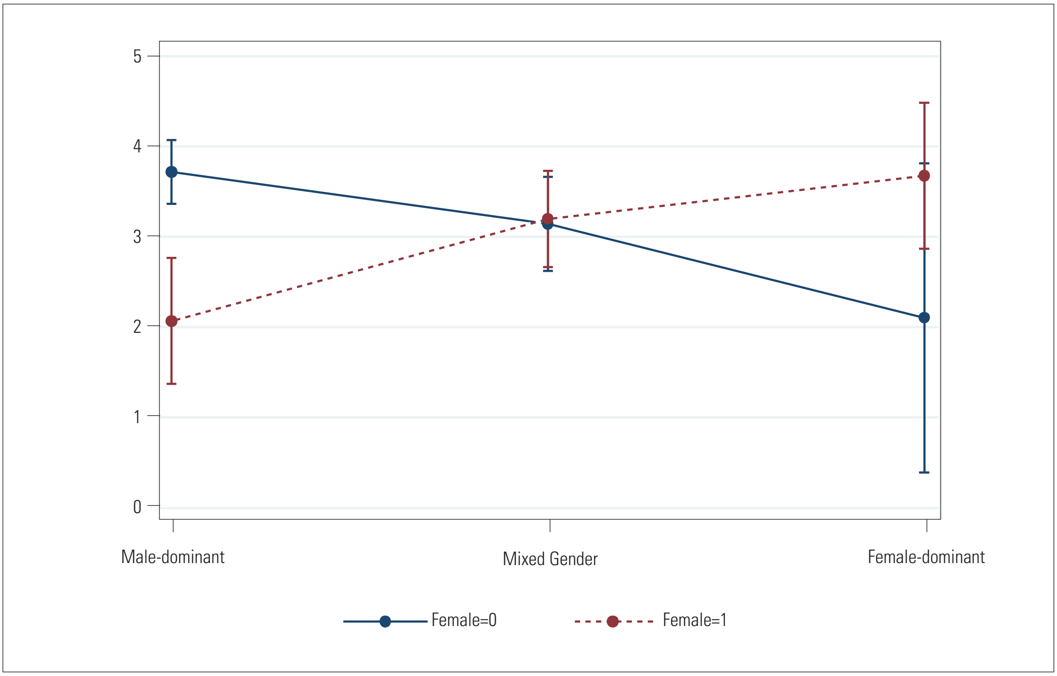


FIGURE 2. Interaction Effect of Team Gender Composition and Individual's Gender on the Number of Grants



As shown in both Figure 1 and Figure 2, there is a typical pattern of interaction in that with the increasing percentage of women on the team, the number of articles and grants for men decreases while for women increases.

The lincom command was used in STATA 13 to test if the differences between groups are significant or not. For women, only the difference between those working in women-majority teams and those working in men-majority teams is significant at the 0.10 level (p=0.09). Women working in women-majority teams published 1.53 times more articles than those working in men-majority teams. The differences between those working in mixed teams and those working in men-majority and women-majority teams were not significant.

Figure 1 shows a declining line for men's productivity as the percentage of women in teams increases. However, the significance tests show that only the difference between mixed teams and men-majority teams is significant. For men, those working in mixed teams published 30% less than those working in men-majority teams, significant at the 0.01 level (p=0.005). The differences for other pairs are not significant. In these three types of gender composition teams, only in men-majority teams is there a significant gender difference in the number of articles produced. Among those working in men-majority teams, women published 50% less than men (p=0.001). In both mixed and women-majority teams, there is no significant difference between men and women in the number of articles produced.

Figure 2 shows a similar pattern as figure 1, and the results of the significance test are also similar. Women working in both mixed teams and women-majority teams obtained more grants than those working in men-majority teams, women working in mixed teams obtained 1.55 times more grants than those working in men-majority teams (p=0.02), and women working in women-majority teams obtained 1.78 times more grants than those working in men-majority teams (p=0.006). The difference between mixed teams and women-majority teams is not significant. Among men, only the difference between mixed and men-majority teams is significant at the 0.10 level (p=0.09). Those working in mixed teams obtained 15% less grants than those working in men-majority teams. As for gender difference, only in men-majority teams is there a significant difference between men and women, where women received 44% less grants than men (p=0.001). In both mixed and women-majority teams, there is no significant difference between men and women in the number of grants obtained.

Hypothesis 1 is mostly supported for men, as they are better off when they are working with men than they are with women, while for women, the difference between working with women and working in mixed teams is not significant. Predictions of self-categorization theory on the advantages of working in homogenous groups are partially proved for men. H2 is supported for women, as for both articles and grants, women working in men-majority teams performed worse than those in other groups. Tokenism is definitely at play for women. For men, however, tokenism does not apply. This is shown by the lack of any significant difference between men working in women-majority teams and other teams, as well as there being no significant gender difference in both indicators of performance in women-majority teams. Thus, H3 is supported here and this data confirms that the effects of tokenism are mediated by gender.

5. DISCUSSION AND CONCLUSION

This article examined the effect of team gender composition on scientists' individual performance in academia. The results supported self-categorization theory for men, in that they performed better in homogeneous groups and less so in mixed groups. These results resonate with other findings on men's more negative subjective responses than women—evaluation on group performance and member's relationships, their own motivation and satisfaction, levels of psychological attachment, absence, and intent to stay—to the increase in women's proportion in traditional men-majority workgroup (Allmendinger & Hachman, 1995; Tsui, et al., 1992). The theory of tokenism is supported only for women. Women are disadvantaged when they are in men-majority teams, and as the percentage of women in teams increases, they perform better; but men in women-majority teams were not disadvantaged compared to women, and to men in other teams.

The gendered work force claim is supported with some reservations. Men have never been disadvantaged compared to women, even in situations where they are supposed to perform worse according to tokenism theory. However, their advantage is not preponderant: It is only significant when they are the majority. When more women enter into the group, men's advantage over women diminishes and disappears.

These results help us to understand the gender gap in science. Considering that more than half of the researchers in this study are working in men-majority teams, the persistence of the gender gap in scientific production and performance is more understandable. Meanwhile, the diminishing gender gap in performance in mixed and women-majority teams suggests that as the percentage of women increases, there will be a change in the intra-group relations between men and women. Although this research did not look into the exact mechanism behind this, it suggests that as the number of women increases in science, there will be a positive culture change in science towards gender equality that allows female scientists to more easily compete with their male counterparts.

Policy-wise, these results support efforts to recruit more women into the scientific workforce, both as a way to achieve gender equality and as a goal itself in gender equality. In addition, two policy measures to further the development of women scientists are proposed based on the current findings. First, considering the increasing number of women and changing gender relationships in science, education and training on gender relations are needed. In China, the notions of gender stereotypes and gender discrimination are still new to many scientists. Increasing scientists' gender awareness and advocating the value of gender equity is work that needs to be done. Training and educating male scientists is particularly important, as they may experience more negative results because of the increasing number of female scientists. Second, the significant gendered effect of team gender composition on individual performance indicates that intervention through research teams is an effective way to promote the development of women in science. Project-granting and policy-making agencies may make team gender composition a criterion in project applications.

The finding that women-majority teams are not worse off than men-majority teams in terms of being conducive to individual performance is worthy of attention. Part of the reason may be due to the fact that this article is not measuring team performance but individual performance and that the team leader is excluded from the analysis. Another possible reason might be that women scientists are catching up with men as a consequence of the preferential policies by the Chinese government and the NSFC.

This article also has the following limitations. Firstly, it does not touch upon the relationship between individual performance and team performance. These two are related but different, and are not addressed in this article, as measurement of team performance was not collected in the survey. Reconciling the results of individual performance with group performance would be necessary for a more complete assessment of the effects of team gender composition on scientific performance. Secondly, it does not identify the causal relationship between team gender composition and individual performance. The data used in this article are cross-sectional. Although academic research teams are usually quite stable over time, we do not know for sure, based on current data, if the gender difference in performance is a result of team gender composition or a reason for team gender composition. A longitudinal research design is needed to tackle this issue. Thirdly, this article does not provide evidence on how changes in team gender composition influence individual performance. Is it through more contacts so that stereotype and discrimination is reduced, or does it relate to group powers between the minority and majority? Further research is needed to answer these questions.

REFERENCES

Alexander, V. D., & Thoits, P. A. (1985). Token achievement: An examination of proportional representation and performance outcomes. *Soc. F., 64*(2), 332-340.

Allmendinger, J., & Hackman, J. R. (1995). The more, the better? A four-nation study of the inclusin of women in symphony orchestras. *Social Forces, 74*(2), 423-460.

Allison, P. D., & Long, J. S. (1990). Departmental effects on scientific productivity. *American Sociological Review, 55*(4), 469-478.

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

Beaver, D. D. (2001). Reflections on scientific collaboration (and its study): Past, present, and future. *Scientometrics, 52*(3), 365-377.

Blalock, H. M. (1957). Percent non-white and discrimination in the South. *American Sociological Review, 22*(6), 677-682.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure* (Vol. 7). New York: Free Press.

Budig, M. J. (2002). Male advantage and the gender composition of jobs: Who rides the glass escalator? *Social Problems, 49*(2), 258-277.

Byrne, D. E. (1971). *The attraction paradigm* (Vol. 11). Academic Press.

Cox, Jr., T. (2001). *Creating the multicultural organization: A strategy for capturing the power of diversity*. Jossey-Bass.

Díaz-García, C., González-Moreno, A., & Jose Sáez-Martínez, F. (2013). Gender diversity within R&D teams: Its impact on radicalness of innovation. *Innovation, 15*(2), 149-160.

Fields, D. L., & Blum, T. C. (1997). Employee satisfaction in work groups with different gender composition. *Journal of Organizational Behavior, 18*(2), 181-196.

Fisher, B. S., Cobane, C. T., Vander Ven, T. M., & Cullen, F. T. (1998). How many authors does it take to publish an article? Trends and patterns in political science. *PS: Political Science & Politics, 31*(4), 847-856.

Fox, M. F. (1991). Gender, environmental milieu, and productivity in science. In H. Zuckeman, J. R. Cole, & J. T. Bruer (Eds.), *The outer circle: Women in the scientific community* (pp.188-204). N.Y.: W. W. Norton

Fox, M. F., & Mohapatra, S. (2007). Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *The Journal of Higher Education, 78*(5), 542-571.

Garcia Martinez, M., Zouaghi, F., & Garcia Marco, T. (2016). Diversity is strategy: The effect of R&D team diversity on innovative performance. *R&D Management, 47*(2), 311-329.

Henderson, L., & Herring, C. (2013). Does critical diversity pay in higher education? Race, gender, and departmental rankings in research universities. *Politics, Groups, and Identities, 1*(3), 299-310.

Herring, C. (2009). Does diversity pay?: Race, gender, and the business case for diversity. *American Sociological Review, 74*(2), 208-224.

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.

Hinnant, C. C., Stvilia, B., Wu, S., Worrall, A., Burnett, G., Burnett, K. Kazmer, M. M., & Marty, P. F. (2012). Author-team diversity and the impact of scientific publications: Evidence from physics research at a national science lab. *Library & Information Science Research, 34*(4), 249-257.

Hollis, A. (2001). Co-authorship and the output of academic economists. *Labour Economics, 8*(4), 503-530.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative science quarterly, 44*(4), 741-763.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal, 52*(3), 599-627.

Kanter, R. M. (1977). *Men and Women of the Corporation*. New York City: Basic books Inc.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social studies of science, 35*(5), 673-702.

Long, J. S., & Fox, M. F. (1995). Scientific careers: Universalism and particularism. *Annual Review of Sociology, 21*(1), 45-71.

Ma, Ying. (2017). Policies and measures towards promoting women's development in S&T fields ——International experiences and China's status. *Forum on Science and Technology in China, 3*,180-184.

Mannix, E., & Neale, M. A. (2005). What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest, 6*(2), 31-55.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.

Martin, P. Y., & Harkreader, S. (1993). Multiple gender contexts and employee rewards. *Work and Occupations, 20*(3), 296-336.

Nakhaie, M. R. (2002). Gender differences in publication among university professors in Canada. *Canadian Review of Sociology/Revue canadienne de sociologie, 39*(2), 151-179.

Newcomb, T. (1961). *The Acquaintance Process*. New York: Holt, Rinehart & Winston. doi:10.1037/13156-000.

Pelled, L. H. (1996). Demographic diversity, conflict, and work group outcomes: An intervening process theory. *Organization Science, 7*(6), 615-631.

Pfeffer, J. (1985). Organizational demography: Implications for management. *California Management Review, 28*(1), 67-81.

Pfeffer, J. (1991). Organization theory and structural perspectives on management. *Journal of Management, 17*(4), 789-803.

Pierce, J. L. (1996). *Gender trials: Emotional lives in contemporary law firms*. Berkeley; LA; London: University of California Press.

Price de Solla, D. J. (1963). *Little science, big science*. New York: Columbia University Press.

Rosenbaum, M. E. (1986). The repulsion hypothesis: On the nondevelopment of relationships. *Journal of Personality and Social Psychology, 51*(6), 1156.

Sarsons H. (2015). *Gender differences in recognition for group work*. Harvard University Working Paper. Also Available at https://scholar.harvard.edu/sarsons/publications/note-gender-differences-recognition-group-work

South, S. J., Bonjean, C. M., Markham, W. T., & Corder, J. (1982). Social structure and intergroup interaction: Men and women of the federal bureaucracy. *American Sociological Review, 47*(5), 587-599.

South, S. J., Bonjean, C. M., Markham, W. T., & Corder, J. (1983). Female labor force participation and the organizational experiences of male workers. *The Sociological Quarterly, 24*(3), 367-380.

Spangler, E., Gordon, M. A., & Pipkin, R. M. (1978). Token women: An empirical test of Kanter's hypothesis. *American Journal of Sociology, 84*(1), 160-170.

Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T. Jost & J. Sidanius, (Eds.), *Political psychology: Key readings, key readings in social psychology* (pp. 276-293). New York, NY: Psychology Press.

Tolbert, P. S., Simons, T., Andrews, A., & Rhee, J. (1995). The effects of gender composition in academic departments on faculty turnover. *ILR Review, 48*(3), 562-579.

Tsui, A. S., Egan, T. D., & O'Reilly III, C. A. (1992). Being different: Relational demography and organizational attachment. *Administrative Science Quarterly, 37*(4), 549-579.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and Research* (pp. 77-122). Greenwich, Connecticut: JAI Press.

Wharton, A., & Bird, S. (1996). Stand by your man: Homosociality, work groups, and men's perceptions of difference. *Re-*

search on Men and Masculinities Series, 9, 97-114.

Williams, C. L. (1992). The glass escalator: Hidden advantages for men in the "female" professions. *Social problems, 39*(3), 253-267.

Williams, K. Y., & O'Reilly III, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. *Res Organ Behav, 20*, 77-140.

Winquist, J. R., & Larson Jr, J. R. (1998). Information pooling: When it impacts group decision making. *Journal of Personality and Social Psychology, 74*(2), 371.

Xie, Y., & Shauman, K. A. (1998). Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review, 63*(6), 847-870.

Zhu, Y. & Ma, Y. (2015). Gender, Time Allocation and Scholarly Productivity among Chinese University Faculty. *Collection of Women's Studies, 4*, 24-49.