



## Faculty Publications

---

2022

# Strength in Numbers: A Field Experiment in Gender, Influence, and Group Dynamics

Olga B. Stoddard

*Brigham Young University*, [olga.stoddard@byu.edu](mailto:olga.stoddard@byu.edu)

Christopher F. Karpowitz

*Brigham Young University*

Jessica Preece

*Brigham Young University*

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>



Part of the [Economics Commons](#), and the [Political Science Commons](#)

## Original Publication Citation

Strength in numbers: A field experiment in gender, influence, and group dynamics" (with Chris Karpowitz and Jessica Preece). IZA Discussion Paper 13742.

---

## BYU ScholarsArchive Citation

Stoddard, Olga B.; Karpowitz, Christopher F.; and Preece, Jessica, "Strength in Numbers: A Field Experiment in Gender, Influence, and Group Dynamics" (2022). *Faculty Publications*. 5806.  
<https://scholarsarchive.byu.edu/facpub/5806>

This Working Paper is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

DISCUSSION PAPER SERIES

IZA DP No. 13741

**Strength in Numbers:  
A Field Experiment in Gender, Influence,  
and Group Dynamics**

Olga Stoddard  
Christopher F. Karpowitz  
Jessica Preece

SEPTEMBER 2020

## DISCUSSION PAPER SERIES

IZA DP No. 13741

# Strength in Numbers: A Field Experiment in Gender, Influence, and Group Dynamics

**Olga Stoddard**

*Brigham Young University and IZA*

**Christopher F. Karpowitz**

*Brigham Young University*

**Jessica Preece**

*Brigham Young University*

SEPTEMBER 2020

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9  
53113 Bonn, Germany

Phone: +49-228-3894-0  
Email: [publications@iza.org](mailto:publications@iza.org)

[www.iza.org](http://www.iza.org)

## ABSTRACT

---

# Strength in Numbers: A Field Experiment in Gender, Influence, and Group Dynamics\*

Policy interventions to increase women's presence in the workforce and leadership positions vary in their intensity, with some including a lone or token woman and others setting higher quotas. However, little is known about how the resulting group gender compositions influence individuals' experiences and broader workplace dynamics. In this paper, we investigate whether token women are disadvantaged compared to women on majority-women mixed-gender teams. We conducted a multi-year field experiment with a top-10 undergraduate accounting program that randomized the gender composition of semester-long teams. Using laboratory, survey, and administrative data, we find that even after accounting for their proportion of the group, token women are seen as less influential by their peers and are less likely to be chosen to represent the group than women on majority-women teams. Token women also participate slightly less in group discussions and receive less credit when they do. Women's increased authority in majority-women teams is driven primarily by men's behavior, not homophily or self-assessment. We find that over time, the gap in general assessments of influence between token and other women shrinks, but this improvement does not carry over to task-specific assessments. Finally, predictors of future grades are different for token women than for other participants, and regardless of treatment condition, women's task expertise is incorporated into group decisions less often than men's. Our findings have implications for team assignments in male-dominated settings and cast significant doubt on the idea that token women can solve influence gaps by "leaning in."

**JEL Classification:** J16

**Keywords:** gender, field experiment

**Corresponding author:**

Olga Stoddard  
Department of Economics  
Brigham Young University  
Provo, UT 84602  
USA

E-mail: [olga.stoddard@byu.edu](mailto:olga.stoddard@byu.edu)

---

\* We thank Rachel Fisher-Sandhu, Dani Hogan, Connor Kreutz, and Mandi Eatough for their excellent research assistance and Sandra Black, Zoe Cullen, Jeff Denning, Jennifer Doleac, Laura Gee, Lars Lefgren, Emily Leslie, Emily Nix, Stephen O'Connell, and Richard Patterson for their valuable feedback on earlier drafts of the paper. We express our appreciation to audience members at the Harvard Kennedy School, University of Chicago Harris School, Michigan State University, Notre Dame University, Stockholm University, Uppsala University, Brigham Young University, Advances in Field Experiments Conference, Western Economic Association International Conference, European Conference of Politics and Gender, Science of Diversity and Inclusion Conference, IZA ECBE conference, E-con of Education Virtual Seminar, and the 2019 Utah Experimental Conference. We thank the Brigham Young University College of Family, Home, and Social Science for generous research support. The order of authors' names was determined randomly. All authors contributed to the project equally.

# 1 Introduction

In recent years, concerns about lack of women’s representation in the workplace and decision-making bodies have led to a variety of public and private initiatives to diversify work teams, academic panels, corporate boards, public commissions, and other such groups. But these initiatives vary dramatically in intensity. For example, Norway set a gender quota of 40% women for corporate boards that was implemented in 2008, and several other Western European countries followed.<sup>1</sup> By contrast, the only US state that has instituted a corporate board quota, California, set a much different standard—California’s Women on Boards law requires all publicly held corporations whose principal offices are located in the state to have at least one female director on their Board of Directors. Early evidence suggests that many California corporations have chosen to comply with this law by adding a token<sup>2</sup> woman to their otherwise all-male boards.<sup>3</sup>

Presumably the purpose of these interventions is to bring women’s voices and expertise to the table so they can influence the discussions these bodies have and decisions they make. Yet many critics argue that interventions that place a token woman in a male-dominated setting are likely to fail because women are rarely taken seriously when they are significantly outnumbered by men. The implication is that the intensity of these interventions matters a great deal. This critique poses a number of important empirical questions. Does the number of women in male-dominated settings matter for their ability to influence deliberations? Can token women "lean in" and become full participants in group discussions? Or, do they face barriers to fully contributing their perspectives and talents even when they nominally have a seat at the table?

We conduct a first-of-its-kind, multi-year field experiment with a top-10 undergraduate

---

<sup>1</sup><https://www.economist.com/business/2018/02/17/ten-years-on-from-norways-quota-for-women-on-corporate-boards>

<sup>2</sup>In this paper, we use the term "token" interchangeably with the term "lone" without intending to signal any normative implications.

<sup>3</sup><https://bpd.cdn.sos.ca.gov/women-on-boards/WOB-Report-04.pdf>

accounting program in the United States. The program uses a group-based pedagogical approach designed to emulate work teams in the business world and partnered with us because of its interest in the experience of both women and men in these team settings. The program allowed us to randomly assign students to a gender composition condition and then within that condition to 5-person groups. We collected survey data from team members at multiple points throughout the following two academic years and also observed group functioning in response to incentivized team-building lab exercises.

The aim of our study was to assess whether the gender composition of an individual's work team affects women's ability to influence group deliberation and decision-making. Using laboratory, survey, and administrative data, we find that lone women are significantly less likely to be rated as influential in team deliberations and to be chosen as a spokesperson for their team than women in majority-women teams. In other words, after accounting for the proportion of women in the group, group gender composition is causally related to who is perceived as influential and capable of representing the group. We also find a persistent gender effect in whose laboratory task expertise is incorporated into group decision-making that the majority-women intervention cannot overcome. In our detailed laboratory data, none of these outcomes improve for women over the course of four months.

Our rich data also allows us to explore a number of mechanisms through which these effects may operate. In particular, we find that token women's competence on group tasks is not rewarded with increased influence and that they get significantly lower credit for participating in discussions than all other participants in the study. We further find that changes in perceptions of influence across conditions are driven by men's behavior; our results are not driven by patterns of self-votes, and women in majority-women groups do not differentially vote for other women.

Unlike previous studies, our data also allow us to study how these dynamics change over time. In monthly survey data, we find modest evidence that token women are able to build

general (as opposed to task-specific) influence in their groups over time. And, we find no statistically significant differences for average program grades across the treatment conditions. However, we do find that correlates of future academic performance<sup>4</sup> are quite different for token women than for all other participants, which suggests avenues for future research.

These findings contribute to a new and growing literature in economics and other social sciences on mixed-gender<sup>5</sup> work teams. Prior work finds that evaluations of competence are highly gendered, and the gender composition of an environment and how a task or domain are stereotyped matter a great deal for these perceptions (Bordalo et al., 2019; Karpowitz and Mendelberg, 2014; Preece, 2016). Challenges appear to be especially acute in settings where women have been historically underrepresented (Bordalo et al., 2016; Coffman, 2014). This leads to outcomes such as gender disparities in leadership because women correctly presume they will have less support from their team when men are in the majority (Born et al., 2020), women taking and getting less credit for joint work with men (Isaksson, 2018; Sarsons, 2017; Koffi, 2020), and women doing more of the “non-promotable” tasks in groups (Babcock et al., 2017).

Most of this research has been conducted in laboratory settings, with only a few studies in the field. And to our knowledge, no one has experimentally studied naturally occurring groups over time, despite the fact that many teams, boards, and councils interact frequently and influence is not a single-shot process. Attending to the dynamics of such groups is important because the gender composition of one’s environment seems to have important implications for individuals’ experiences and outcomes (Bostwick and Weinberg, 2020; Booth et al., 2018; Brenoe and Zölitz, 2019; Hill, 2017; Zölitz and Feld, 2018; Kofoed et al., 2019; Anelli and Peri, 2019; Apesteguia et al., 2012; Folke and Rickne, 2020).<sup>6</sup> On a more practical

---

<sup>4</sup>Grade point average (GPA) is a commonly used performance metric in the U.S. education system.

<sup>5</sup>Although we recognize the conceptual distinction between sex and gender, in this paper we use them interchangeably, in keeping with norms in economics.

<sup>6</sup>This is an observation that many social scientists in sociology, psychology, and political science have studied for decades, though causal identification has typically been a challenge. See, for example (Bratton,

level, understanding how to design effective teams is a crucial task for employers because teams can be an efficient way to coordinate production requiring a diversity of skills, talents, and information (Eckel and Grossman, 2005; Lindbeck and Snower, 2000) and have become increasingly common in the workplace (Berg et al., 1996; Boning et al., 2007; Hamilton et al., 2003; Nix, 2020; Lawler and Mohrman, 2003; Lazear and Shaw, 2007; Wuchty et al., 2007). All of this suggests that patterns of gender marginalization are not just normatively concerning, but also have implications for team effectiveness and firms’ productivity.

Our research design allows us to address some of the limitations of previous work by examining group dynamics in randomized groups that meet repeatedly over an extended period of time in a natural and meaningful setting. We find that across a wide variety of outcomes, token women experience this team-based program differently than all other participants. When women are outnumbered, their leadership and participation are significantly under-recognized. And, our examination of mechanisms suggests that women cannot simply “lean in” to rectify these deficits.

## 2 Research Design

We partnered with a top-10 undergraduate accounting program in the US to randomize the gender composition of students’ assigned work teams during the fall semester of their junior year. Students are competitively admitted to the program at the end of their sophomore year. The fall semester is their first in the rigorous program and assigned five-person teams are a critical part of their experience that semester. They attend classes, work on cases and other course assignments, and study for exams with their five-person group. Program administrators have designed these teams to replicate the work environment of many accounting firms, and faculty use these groups to train students on professional interactions. As is typical of teams

---

2005; Kanter, 1977; Krook, 2010; McDonald et al., 2004; Spangler et al., 1978; Yoder et al., 1998; Zimmer, 1988).



in the business world, students are instructed to make decisions together and respect each other's contributions. During the course of the semester, students complete both individual and group assignments, and they interact frequently with each other. The vast majority of groups met together at least 2-3 days a week.<sup>7</sup>

The program enrolled 535 students (384 men and 151 women) in Fall Semesters of 2016 and 2017. Women comprise only about 28% of the overall enrollment in the program, and historically the program responded to this imbalance by placing one or two female students per group. Program directors were, however, interested in understanding the effects of this method of team assignment on their female students' experiences. Hence, in this study, female students were randomly assigned to one of two conditions: 1 woman and 4 men (1F) or 3 women and 2 men (3F). Because there were significantly more men than women in the program, men were randomized into one of three conditions: all five men (0F), 1F or 3F. Once individuals had been randomized into a condition, we randomized students into groups. The program stipulated that the average GPA of each group should be similar and that there should only be one international student per group. Hence, we block randomized on these variables.<sup>8</sup>

Throughout the semester and subsequent two years, we collected laboratory, survey, and administrative data to measure the outcomes of this randomization. Prior to group assignment, students took an extensive baseline survey. After group assignment, students completed monthly surveys during fall semester that focused on their perceptions of group interactions and asked them to evaluate their team members' contributions. At the completion of their second (senior) year of the program, we conducted an incentivized exit survey of all

---

<sup>7</sup>This program is not unique in using assigned teams as a pedagogical and professionalization tool; many accounting programs and most MBA programs also assign students to work teams like these.

<sup>8</sup>We also had a small number of groups with 2 women and 3 men (2F). Because our statistical power to analyze these groups was low, we do not include them in our main analysis; however Appendix C reports our analysis for "Minority F" groups, which include participants in 1F and 2F groups. Results are robust to the inclusion of these 2F groups.

students to collect data on their graduate school and labor market outcomes. We also have administrative data on student grades.

In addition to the survey and administrative data, teams participated in an incentivized laboratory task at both the beginning and end of the Fall Semester. At the beginning of the semester, the task was the “Survival on the Moon” task<sup>9</sup>, in which participants are given a list of 15 items and asked to rank them from most to least useful for survival on the moon. This task was originally conceived as a group-based pedagogical exercise, and expert answers have been provided by NASA. At the completion of the semester, the task was the similar “Lost at Sea” task.<sup>10</sup> These tasks are commonly used as leadership and team-building exercises in the corporate world; detailed task instructions are in Appendix F. Upon arriving at the lab, students first completed the task on their own, giving us a measure of how well each individual independently performed on the task.

Participants were then asked to deliberate with their group members to create a group ranking on the items in the above task. Because participants were fitted with recording equipment, we are able to identify individual participation in discussions with great precision.<sup>11</sup> After they determined their group ranking, each individual privately completed an exit survey about their group deliberation experience in which they voted for a spokesperson to present their results to a panel of judges at a later date.<sup>12</sup> Lab performance was incentivized in several ways: in addition to the \$5 show-up fee, students earned \$50 for having the individual answer closest to the experts, \$50 per group member for having the group answer closest to the experts, and another \$50 for the whole group to share for the best presentation by their

---

<sup>9</sup><https://www.psychologicalscience.org/observer/nasa-exercise>

<sup>10</sup>[http://plcmets.pbworks.com/f/lost\\_at\\_sea.pdf](http://plcmets.pbworks.com/f/lost_at_sea.pdf)

<sup>11</sup>Recording equipment was 6-channel audio recording, which yields a separate, high-quality recording for each member of the group as well as a track the records the group together.

<sup>12</sup>Students were told that five groups would be selected at random to have an opportunity to have their spokesperson present their group ranking to the judges in 1-2 weeks for a chance to earn additional \$50 for the best presentation. The groups were thus incentivized to select the most capable group member as their spokesperson.

chosen spokesperson.

Students were required to participate in the monthly surveys and laboratory exercises for course credit and for the purpose of an internal program evaluation. However, to be included in the analysis in this paper, they had to consent to allow their data to be used for research purposes.<sup>13</sup> Participants were informed that both the internal program evaluation and the research study would examine team dynamics (without any mention of gender), and the experiment did not include any deception.

### 3 Sample and Balance Tests

Admission to this program is highly competitive. Matriculated students typically have high GPAs and considerable leadership experience. As reported in Table 1, the men and women in our sample both came to the program with impressive academic credentials, and there are no statistically significant differences either within genders and across conditions or across genders in academic performance or demographic characteristics. If anything, women came to the study more qualified to exercise leadership—defined as having held a position in high school student government—than the men who participated. Across a wide variety of background characteristics collected prior to group assignment via administrative data and a baseline survey (See Appendix G for baseline survey instrument), we find few differences between the men and women assigned to the various treatment conditions. In other words, Table 1 affirms that randomization was successful. However, we also include covariate-adjusted results in our analysis tables as a demonstration of robustness.

---

<sup>13</sup>Two students did not consent to the use of their data, and their responses are not included in the results below. Subjects provided separate consent for lab and survey data, and we did not receive consent from one female in the lab and one male in the surveys.

Table 1: Randomization Balance Table By Gender and Condition

	1F Female	3F Female	0F Male	1F Male	3F Male
Age	22.5	22.7	23.8	23.9	23.8
White	73.7	86.9	90.8	94.8	91.5
Married	31.6	23.1	39.9	32.5	36.6
GPA	3.79	3.78	3.80	3.78	3.79
Leadership experience	52.6	53.2	45.4	47.7	52.8
International student	20.5	10.2	3.9	5.8	9.7
Parental income category	3.8	3.9	4.3	4.3	4.1
Political affiliation scale	5.2	5.5	5.8	5.8	5.8
Ambivalent sexism index	44.5	40.8	56.1	57.7	53.9
Egalitarianism index	59.6	56.7	51.8	55.5	55.8
Individualism index	63.9	66.0	69.9	73.2	71.0
Pro-sociality index	72.9	73.1	71.2	72.9	73.0
Conflict avoidance index	49.0	48.4	43.5	41.5	40.3
Individual efficacy index	57.3	59.1	70.1	70.4	69.0
Group efficacy index	54.4	53.9	64.1	64.4	63.1
N	38	108	153	154	71

Note: we found no statistically significant within-gender differences across conditions at the 5% level.  
The N accounts for 1 woman and 1 man who did not provide consent.

## 4 Results

### Perceived Laboratory Influence

Perceptions matter in a variety of contexts, including workplace evaluations with salary and career implications. Hence, we explore whether group gender composition matters for perceptions of who is most and least influential in a group. At the conclusion of both the “Lost on the Moon” and “Lost at Sea” tasks, we conducted an exit survey that asked participants to identify the team member who was the most influential and the team member who was least influential in the decision-making deliberations they had just experienced. Figure 1 reports our results. The horizontal dashed line in Figure 1 represents how many votes women in that treatment condition would receive if they were selected in proportion to their presence in the

group—20% for the 1F groups and 60% for the 3F groups. Normalizing in this way allows for easier comparison across treatment conditions. The variable of interest is equivalent to the number of votes per woman in the group—a number that would equal one if influence were distributed randomly. In all our analyses, results are combined for both tasks because we see no meaningful differences across time, but disaggregated results are reported in Appendix A.

As Figure 1 makes clear, members of 1F teams were extremely unlikely to see their female teammate as most influential in their laboratory task deliberations, even relative to their comparatively low proportion of the group. They were also very likely to see her as the least influential person on the team. By comparison, members of 3F teams chose women as most and least influential about as often as expected given their proportion in the group.

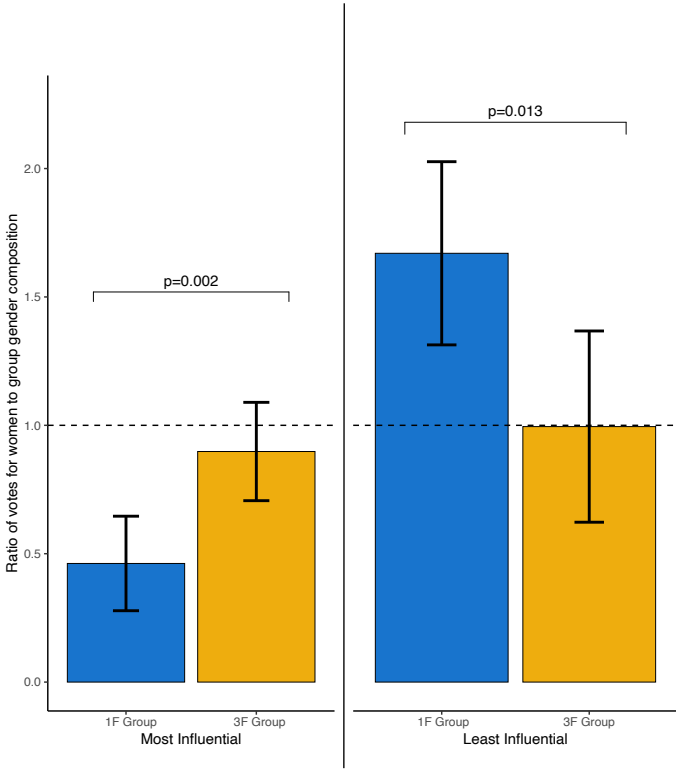
Table 2 confirms that the experimental differences between the 1F and 3F conditions are statistically significant at the  $p \leq 0.01$  level for most influential votes and at the  $p \leq 0.05$  for least influential votes. Specifically, we estimate the following equation using OLS:

$$Y_g = \alpha + 1F\beta + X_g\gamma + \epsilon_g \quad (1)$$

The analysis is conducted at the group level, and the dependent variable is the average proportion of votes for women divided by the proportion of women in group  $g$ .  $1F$  captures the gender composition of the group. The coefficient of interest is  $\beta$ , which represents the difference between the point estimate for the 3F groups and 1F groups shown in Figure 1.  $X$  is a vector of variables that control for GPA, age, race, marital status, and international students. Columns (1) and (2) in Table 2 report the effect of being in a 1F group on women’s votes as *most* influential, while columns (3) and (4) as *least* influential.

We find that token women receive about .44 influence votes per woman less than women in majority women groups. The trend works in reverse for the least influential votes. Token women get about two-thirds *more* least influential votes per woman compared to women in

Figure 1: Lab Influence Votes



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table 2 for full results)

majority women groups. These results remain significant at conventional levels even in the presence of controls for group characteristics, such as the presence of international students, the racial composition of the group, or the average age of group members.

As a robustness check, we also pursued a second analytic strategy in which we explored gender differences in who received the most influence votes within the group by randomly selecting one man in each 1F condition group to compare against the group’s token woman.<sup>14</sup> Averaging across the two labs, 29.5% of the randomly chosen men were chosen as the most

<sup>14</sup>The dependent variable is a dichotomous indicator of whether the group member received the most influence votes. In the case of ties, multiple group members could be "most influential."

Table 2: Perceived Lab Influence

	(1) Most Influential	(2) Most Influential	(3) Least Influential	(4) Least Influential
1F Condition	-0.437*** (0.135)	-0.500*** (0.141)	0.675** (0.264)	0.671** (0.281)
Average GPA (Group)		0.522 (1.128)		0.619 (2.251)
Average Age (Group)		0.125 (0.081)		0.078 (0.161)
International Student in Group		0.106 (0.161)		-0.207 (0.321)
Proportion White		0.814 (0.612)		-0.883 (1.222)
Number of Married Students in Group		-0.023 (0.070)		-0.141 (0.140)
Constant	0.898*** (0.098)	-4.710 (4.885)	0.995*** (0.191)	-2.074 (9.748)
Observations	75	75	75	75
R-squared	0.125	0.183	0.082	0.106

Note: Group-level analysis. Dependent variable is the perceived influence of women, coded by taking the proportion of votes for women divided by the proportion of women in the group.

$$M_{mostinfluential} = 0.67, SD = 0.62, M_{leastinfluential} = 1.35, SD = 1.18$$

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

influential group member, compared to only 8% of token women. This represents a deep influence deficit for the 1F women. By contrast in 3F groups, 19% of randomly chosen women were most influential, compared to 26% of men.

Hence, across multiple measures and analytic approaches, we find strong support for the idea that even though the women in the 1F condition are, on average, equally talented as the women in the 3F condition, they are seen as much less influential in group deliberation. In fact, out of the 39 1F women, *no* women were chosen as most influential in both of their group’s laboratory exercises, compared to 22 out of 154 men in the 1F condition.

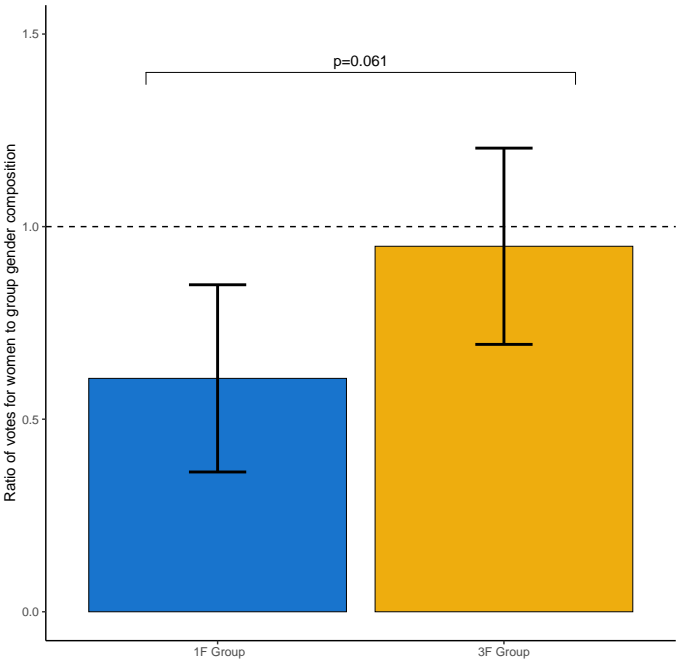
## Laboratory Task Spokesperson Votes

Does this deficit in perceptions of influence manifest itself in behavioral evaluations of competence? Group members did not just vote on the most influential member of the group; they also made an incentivized choice for whom to select as spokesperson for their group.

The spokesperson would have a chance to earn another \$50 for the team by doing a short presentation defending their group’s decision. If the influence votes were perceived as a low-stakes evaluation of group members, the selection of spokesperson was not. With an additional \$50 at stake, each group member was incentivized to vote for the most competent member of the group.

In the spokesperson decision, participants voted individually and without notice or prior deliberation as part of the exit survey. Every group member was listed on the “ballot,” so there was no nomination stage in which people could express that they were not interested in being spokesperson. As with the other measures of influence, we saw treatment effects on the probability of people voting for a woman as spokesperson. Figure 2 and Table 3 report these results.

Figure 2: Lab Spokesperson Votes



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table 3 for full results)



Table 3: Lab Spokesperson Votes

	(1)	(2)
1F Condition	-0.343*	-0.456**
	(0.180)	(0.183)
Average GPA (Group)		-0.073
		(1.469)
Average Age (Group)		0.157
		(0.105)
International Student in Group		0.110
		(0.210)
Proportion White		0.542
		(0.797)
Number of Married Students in Group		0.182*
		(0.092)
Constant	0.949***	-3.202
	(0.130)	(6.362)
Observations	75	75
R-squared	0.047	0.153

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.  $M_{DV} = 0.77$ ,  $SD = 0.79$

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Once again, Figure 2 shows that participants in 1F groups were significantly less likely to choose women relative to their proportion of group membership. Participants in 3F groups were not. Table 3 reports the results of an OLS regression estimating equation 1 at the group level with the proportion of spokesperson votes for women divided by the proportion of women in the group.

Results in Table 3 suggest that including controls sharpens the precision and increases the size of the estimated effect.<sup>15</sup> After adjusting for other factors, token women get about .46 votes less than women in the 3F groups and only about 55 percent of the baseline

<sup>15</sup>The increase in the size of the coefficient suggests some imbalance in demographic characteristics between 1F and 3F groups. In fact, as reported in Appendix Table B5, we find a six-month age difference in the average age of students in 1F and 3F groups. This is not surprising since men are on average slightly older than women in the program. However, the fact that the size of the coefficient in Table 3 actually *increases* when controlling for age, suggests that age differences are not responsible for the differences in spokesperson votes.

expectation. Again, our second analytic approach—randomly selecting other group members for comparison with the token women—also produces evidence of meaningful disadvantage for women in the 1F condition. Averaging across the two labs, only 11% of token women were chosen as spokesperson, compared to 24% of randomly chosen men.<sup>16</sup>

The fact that women were also less likely to be chosen when the choice was consequential and incentivized suggests our attitudinal findings reflect real authority deficits. Insofar as this kind of selection process mirrors opportunities for visibility and promotion in the workplace—for example, presenting findings to a boss or client or being chosen as the “lead” on a project—this result is especially troubling.

## Empirical Influence on the Laboratory Task

Our analysis to this point has focused primarily on group members’ assessments of each other—a process that closely mirrors the way many workplaces assess worker performance. But the structure of our laboratory task also allows to measure empirical influence on group decisions by identifying who was able to pull the laboratory group outcome closest to their initial ranking. We identify which individual had the smallest total difference between their pre-deliberation ranking of items on the Lost on the Moon or Lost at Sea task and the final group outcome. This is a proxy for having the most actual influence in deliberation. We then estimate the following equation:

$$Y_i = \alpha + 1F_i\beta + 1F_i * Female_i\delta + X_i\gamma + \epsilon_i \quad (2)$$

Because we average across both task sessions, the dependent variable in the model takes on 3 values: 0 if the participant was never the most influential member of the group, 0.5 if the participant was most influential once, and 1 if the participant was most influential in both

---

<sup>16</sup>In the 3F condition, 18.5% of randomly chosen women were voted as spokesperson, compared to 26.4% of randomly chosen men.

tasks. For this reason, we present ordered probit estimates below, but these results are robust to other ways of operationalizing empirical influence (see Appendix Tables B1-B3), to OLS estimation (see Appendix Table B4), and to using stacked data instead of averaged data (see Appendix E). As before,  $1F$  is a dummy equal to 1 if an individual was in a 1F group and  $Female$  is a dummy for gender. The interaction between  $1F$  and  $Female$  is our coefficient of interest. The analysis is conducted at the individual level with standard errors clustered at the group level. Table 4 reports our results with and without controls in columns 1 and 2 respectively.

Although we do not find a treatment effect, we do find a durable gender effect. Table 4 shows that token women are just as likely to shape the group decision as 3F women. Or, more accurately, they are just as *unlikely* to shape the group decision. This gender effect persists and even sharpens after controlling for individual task performance and speaking turns. What this suggests is that while group structure may be able to equalize perceptions of influence and likelihood of being chosen to be the group spokesperson, it *does not* equalize actual influence over group outcomes. Regardless of the expertise they bring to the table and how much they “lean in” to the deliberative process, these highly talented women are unable to shape the final decisions of the group as much as men are.

## 5 Mechanisms

We now turn to an examination of some potential mechanisms for token women’s lower levels of influence. Because of the richness of the data we collected, we are able to explore a number of potential mechanisms. However, it is important to note that even though we are able to compare across conditions in these analyses, many rely on measurements of post-treatment behavior, such as talk time and task performance. Hence, we do not interpret these models as causal. Nevertheless, we find these correlations important and believe they provide valuable

Table 4: Determinants of Empirical Influence over Group Decision

	(1)	(2)
Female	-0.366** (0.184)	-0.472** (0.208)
1F Group	-0.085 (0.203)	-0.160 (0.222)
1F X Female	-0.045 (0.324)	0.132 (0.353)
International		-0.143 (0.329)
Age		-0.067* (0.039)
White		0.206 (0.456)
Married		-0.084 (0.175)
Task Performance		2.700*** (0.860)
Speaking Turns		4.284** (1.897)
GPA		0.060 (0.392)
Cut Point 1	0.811*** (0.168)	2.110 (2.042)
Cut Point 2	1.685*** (0.192)	3.050 (2.051)

Note: Cell entries are ordered probit coefficients with clustered standard errors in parentheses. Standard errors clustered by group. Individual group difference is made by taking the absolute value of the difference of each item ranking by the individual and the group. We then divide the minimum difference between the individual and group rankings by the individual ranking. (So, an individual group difference ratio of 1 would be the person with the lowest absolute difference.)

\*\* p<0.01, \* p<0.05, \* p<0.1

insights into observed patterns and mark useful directions for direct testing future research.

## Converting Task Performance to Influence

In an ideal world, proficiency with a task should lead to greater levels of influence in group deliberation about that task. The analysis reported in Table 5 explores how this dynamic plays out among men and women in 1F and 3F groups. Because participants had completed the Lost on the Moon and Lost at Sea tasks on their own before deliberation, we can measure the absolute value of the difference between an individual’s pre-deliberation ranking and experts’ rankings of items. This variable captures deviations from the experts’ answers, meaning that higher scores represent more errors and lower scores are indicative of better task performance in these models. We estimate the following equation using OLS:

$$Y_i = \alpha + TaskErrors_i\beta + Female_i\gamma + Female_i * TaskErrors_i\delta + X_i\theta + \epsilon_i \quad (3)$$

The dependent variable is the individual’s average number of influence votes across the two labs. The interaction between *Task Errors* and *Female* captures whether task performance matters differently for men and women. Table 5 reports our results separately for each condition with and without our standard controls and clustering standard errors at the group level.

In the 1F groups, we find that for men, better individual performance on the task is correlated with being ranked as more influential in group deliberation. The negative coefficient on task errors tells us that for every point farther from the correct answer, the average number of influence votes received by men *decreased* by 0.029. Put differently, a one standard deviation increase in errors reduces influence by approximately one-quarter of an influence vote. Conversely, men who do better on the task are rewarded with similar increased influence. However, women receive none of these positive returns for task performance or penalties for poor performance. The large negative coefficient for women highlights women’s overall low levels of influence relative to men in the 1F condition, as we have already shown.

Table 5: Converting Task Performance to Lab Influence, by Gender and Experimental Condition

	(1) 1F	(2) 1F	(3) 3F	(4) 3F
Task Errors	-0.029** (0.012)	-0.032*** (0.011)	-0.000 (0.013)	0.005 (0.014)
Female	-2.495*** (0.846)	-2.865*** (0.944)	1.340 (1.206)	1.243 (1.164)
Female x Task Errors	0.036** (0.016)	0.043** (0.018)	-0.030 (0.021)	-0.026 (0.021)
International		0.161 (0.323)		-0.156 (0.564)
Age		-0.020 (0.056)		0.115*** (0.038)
White		0.096 (0.270)		-0.170 (0.549)
Married		0.030 (0.152)		-0.018 (0.194)
GPA		-0.613 (0.528)		0.316 (0.503)
Constant	2.594*** (0.580)	5.404* (2.698)	1.157 (0.689)	-2.833 (2.407)
Observations	193	192	180	178
R-squared	0.092	0.098	0.035	0.081

Note: Cell entries are OLS coefficients with clustered standard errors in parentheses. Standard errors clustered by group. Dependent variable is the individual's average number of influence votes. Excluded category is 3F groups.  $M_{DV} = 0.99$ ,  $SD = 1.17$

\*\* p<0.01, \* p<0.05, \* p<0.1

Moreover, the interaction term between female and task performance indicates that task performance matters differently for men than for women. Here the positive interaction term suggests that 1F groups nets out the main effect of task performance and that women who perform better on the task receive no reward. If anything, women who are better at the task might even be *penalized*.

By contrast, there is no correlation between task performance and influence votes in the 3F groups. From an efficient use of human capital standpoint, this may not be ideal. But, the efficient conversion of task performance into influence for men that we see in the 1F

groups must be weighed against the opposite pattern for women on these teams. This has policy implications for organizations that care about the egalitarian treatment of employees.

## Laboratory Speaking Turns

Prior laboratory research has shown a strong effect of group gender composition on women’s participation in deliberation among members of the general population (Karpowitz and Mendelberg, 2014). We explore this potential mechanism by estimating the following equation:

$$Y_i = \alpha + 1F_i\beta + Female_i\delta + 1F_i * Female_i\gamma + X_i\theta + \epsilon_i \quad (4)$$

The dependent variable is the proportion of speaking turns each individual takes, averaged across the two lab sessions. These are individual-level data, and the interaction between the 1F condition and gender highlights the difference-in-differences in women’s speaking turns across conditions. Table 6 reports our results. In addition to the standard controls in column (2), we also include the length of group conversation. We find that women in the 1F condition take a somewhat smaller proportion of the speaking turns than women in the 3F groups, but this effect is substantively small and is no longer statistically significant in the presence of controls. In other words, in both conditions, the women in our sample appear to be “leaning in” by participating actively in discussions.

## Converting Speaking Turns to Influence

What happens when these women participate in deliberations? Is there a relationship between speaking turns and influence, and if so, is that relationship consistent across genders and conditions? We explore this by estimating the following equation:

$$Y_i = \alpha + SpeakingTurns_i\beta + Female_i\gamma + Female_i * SpeakingTurns_i\delta + X_i\theta + \epsilon_i \quad (5)$$

Table 6: Relationship between Experimental Conditions and Speaking Turns

	(1)	(2)
1F Condition	0.013** (0.006)	0.011** (0.005)
0F Condition	0.010* (0.005)	0.005 (0.005)
Female	0.014 (0.009)	0.016** (0.007)
1F x Female	-0.027* (0.014)	-0.020 (0.013)
Length of Group Conversation	-0.000 (0.000)	-0.000 (0.000)
International		-0.047*** (0.012)
Age		0.001 (0.001)
White		0.010 (0.011)
Married		0.011** (0.004)
GPA		0.021 (0.013)
Constant	0.199*** (0.008)	0.100 (0.065)
Observations	527	523
R-squared	0.010	0.097

Note: Cell entries are OLS coefficients with clustered standard errors in parentheses. Standard errors clustered by group. Dependent variable is the individual's average proportion of speaking turns. Excluded category is 3F groups.  $M_{DV} = 0.20$ ,  $SD = 0.05$

\*\* p<0.01, \* p<0.05, \* p<0.1

The dependent variable is the individual's average number of influence votes and we report the results in Table 7 separately by condition and clustering standard errors at the group level.

We find a strong positive correlation between a number of speaking turns and being seen as influential in both 1F and 3F groups. However, in 1F groups, that relationship is driven almost entirely by men, as can be seen in the large and positive coefficient attaching to the speaking turns variable. The large and negative interaction between participant's gender



and speaking turns means that the ability to turn conversational participation into influence evaporates for women. Computing the marginal effect of speaking turns for both men and women, we find that men’s returns for speaking up are nearly six times greater than women’s.

Table 7: Converting Speaking Turns to Lab Influence, by Gender and Experimental Condition

	(1) 1F	(2) 1F	(3) 3F	(4) 3F
Speaking Turns	6.163*** (1.840)	6.585*** (1.758)	8.277*** (2.842)	9.163*** (2.830)
Female	0.392 (0.450)	0.435 (0.601)	0.376 (0.844)	0.453 (0.855)
Speaking Turns x Female	-5.119** (2.208)	-5.116* (2.729)	-3.705 (4.496)	-3.674 (4.549)
International		0.126 (0.314)		0.383 (0.490)
Age		0.017 (0.054)		0.120*** (0.035)
White		0.047 (0.263)		-0.143 (0.432)
Married		0.001 (0.153)		-0.088 (0.180)
GPA		-0.611 (0.513)		0.077 (0.518)
Constant	-0.131 (0.377)	1.640 (2.734)	-0.432 (0.520)	-3.626 (2.211)
Observations	193	192	180	178
R-squared	0.107	0.114	0.092	0.151

Note: Cell entries are OLS coefficients with clustered standard errors in parentheses. Standard errors clustered by group. Dependent variable is the individual’s average number of influence votes. Excluded category is 3F groups.  $M_{DV} = 0.99$ ,  $SD = 1.17$   
 \*\* p<0.01, \* p<0.05, \* p<0.1

The overall relationship between speaking turns and influence in the 3F groups is even stronger than in the 1F groups. But unlike in 1F groups, this relationship is not gendered to the same extent. Although the point estimate for the interaction term between *Speaking Turns* and *Female* is negative, suggesting the possibility that women may also face some disadvantages even when they comprise the gender majority, it is considerably smaller and nowhere near statistically significant.

When these results are combined with the findings about returns to task performance, the message is sobering. In 1F groups, women experience a negative correlation between being good at the task and influence *and* speaking up is not particularly effective. These findings should cast significant doubt on the idea that token women can control the factors that lead to being seen as influential. Instead, those factors appear to be largely structural in our data. As mentioned above, these are post-treatment variables and need to be interpreted with caution. However, these findings suggest that the internal dynamics of how influence is distributed in 1F and 3F groups are quite distinct from each other.

## Source of Influence Votes

What is the source of the difference in votes for women in 1F and 3F groups? Perhaps the results are the effect of lone women questioning their own contributions to group discussions and being less likely to self-promote?<sup>17</sup> If so, the inequality could be ameliorated by training token women to give themselves more credit. Or perhaps men's behavior is changing across the conditions? Table 8 and Figure 3 address these questions.

Table 8 suggests that token women feeling uncomfortable voting for themselves is not responsible for the main effects. Assigned condition has no influence on whether any participant (male or female) ever voted for themselves. Among women, there is no statistically robust difference between 1F and 3F conditions in likelihood of self-voting, and among men, we find the same, even when we include 0F groups. That result implies that the treatment works through others' perceptions of who is influential, not one's perception of their own influence.

---

<sup>17</sup>See for example [Exley and Kessler \(2019\)](https://users.nber.org/~kesslerj/papers/ExleyKessler_SelfPromotion.pdf) study documenting gender differences in self-promotion. [https://users.nber.org/~kesslerj/papers/ExleyKessler\\_SelfPromotion.pdf](https://users.nber.org/~kesslerj/papers/ExleyKessler_SelfPromotion.pdf)

Table 8: Effect of Experimental Conditions on Probability of Voting for Self, by Gender

	(1)	(2)	(3)	(4)
	Women	Women	Men	Men
0F Condition			0.024 (0.146)	0.031 (0.148)
1F Condition	-0.183 (0.269)	-0.134 (0.280)	0.024 (0.161)	0.011 (0.167)
International		0.781 (0.575)		0.303 (0.355)
Age		0.063 (0.043)		-0.010 (0.042)
White		0.996 (0.658)		0.017 (0.297)
Married		-0.340 (0.308)		-0.097 (0.153)
GPA		0.908 (0.664)		-0.061 (0.420)
Constant	-0.896*** (0.152)	-4.266*** (1.192)	-1.025*** (0.163)	-1.029 (1.260)
Observations	147	145	380	378

Note: Cell entries are OLS coefficients with clustered standard errors in parentheses. Standard errors clustered by group. Dependent variable is a dichotomous indicator of whether the participant ever voted for him or herself in either lab task. Cell entries are probit coefficients.  $M_{DV} = 0.22$ ,  $SD = 0.41$   
 \*\*  $p < 0.01$ , \*  $p < 0.05$ , \*  $p < 0.1$

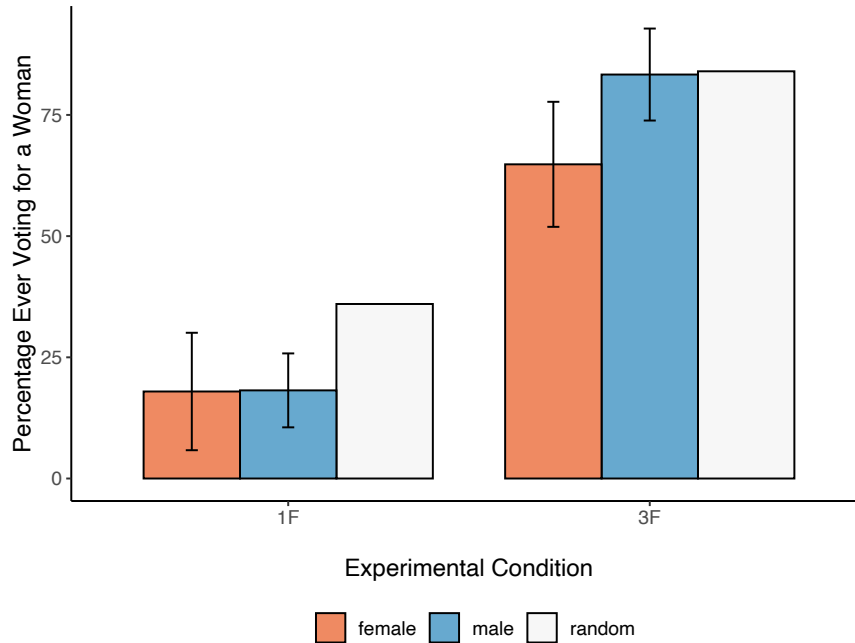
In Figure 3, we explore the possibility that our results are driven by homophily by examining the percentage of participants who ever voted for a woman as most influential in either lab task. This is the most generous way to capture the respondent’s willingness to grant high levels of influence to women. The figure reveals that regardless of the presence or absence of controls, both men and women were much less likely to cast even a single influence vote for women in the 1F condition. The magnitude of the bias against women is startlingly large.<sup>18</sup> If influence votes were allocated at random, we would expect that approximately 36 percent of participants would choose a woman as most influential at least once across the two lab sessions, an expectation that is shown by the white bar in the figure.<sup>19</sup> Instead, only 18%

<sup>18</sup>Results are robust to controls. See Appendix Table B5 for details.

<sup>19</sup>This baseline expectation is generated by computing  $1 - (\Pr(\text{Choosing a Man}) * \Pr(\text{Choosing a Man}))$ . Because men comprise 80 percent of the 1F groups, the expected value is  $1 - (0.8 * 0.8) = 1 - 0.64 = 0.36 * 100 = 36\%$ .

of both men and women actually voted for a woman at least once — only half the expected value.

Figure 3: Percentage Ever Choosing a Woman as Most Influential, by Gender



Note: Bars represent the percentage of respondents ever voting for a woman as most influential in either lab session. (See Table B5 for results with controls.)

As the figure shows, however, results are very different in the 3F condition. In groups with a majority of women, both men and women were substantially more likely to choose a woman as most influential at least once, and we see little evidence that this difference is driven primarily by women. If anything, the opposite is true. Following the same approach we used in the 1F condition, if influence votes were allocated at random in 3F groups, we would expect about 84% of participants to vote for a woman at least once. The actual result among men was 83%. In other words, the structural changes in group gender composition essentially erase men’s unwillingness to vote for women. By contrast, only 65% of women voted for a woman as most influential in 3F groups. With this pattern of results, we see little

reason to conclude that homophily among women drove women’s increased influence in 3F groups. To the extent that men and women differed in the 3F condition, women, not men, were less likely to see women as authoritative.

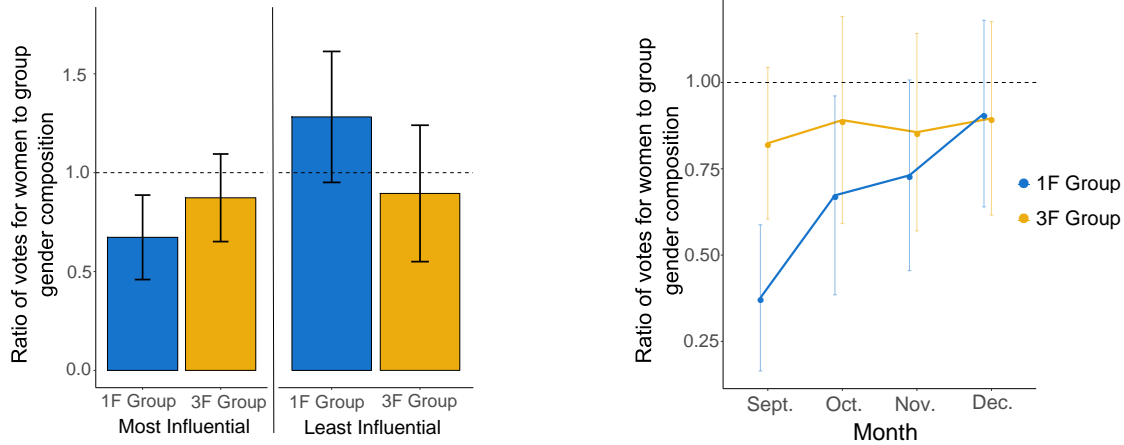
## 6 Medium Term Outcomes and Dynamics

### Survey Results over Time

To this point, we have documented persistent challenges for women’s ability to generate both actual and perceived influence in their groups, especially token women. We do, however, find some good news about women’s empowerment in these groups. In addition to the laboratory tasks in September and December, we collected survey data from the participants over the course of the semester. Once per month, students were asked, among other things, “Who has been the most influential member of your group during group discussion and collaboration?” and were presented with a drop-down menu of their group members (including themselves) to choose from.

The monthly survey data echoes much of what we find in the lab, but with a twist. As the first panel of Figure 4 shows, the same basic pattern shows up in the surveys: relative to baseline expectations, lone women are less likely to be seen as most influential and more likely to be seen as least influential than we would predict given their prevalence in the group and 3F women fare better. But as the second panel shows, these findings are driven primarily by survey responses at the beginning of the semester. In contrast to the lab, where we saw no changes over time, more global assessments of women’s influence in the 1F groups do seem to improve. By December they are indistinguishable from the evaluations 3F women receive. These point estimates are still somewhat below the line that indicates they are getting the same number of votes as their presence in the group, but the confidence interval includes that line. Appendix Tables A6 and A7 confirm these findings.

Figure 4: Survey Influence Votes



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table A6 for full results.)

To our knowledge, the distinction between how global evaluations change over time and how specific task-based evaluations do not change over time has not been previously identified or studied and merits further research. Our best guess is that general assessments of authority can build over time, but that discrete tasks more or less reset the clock. In the discussion section, we discuss the implications of this pattern further.

## Academic Outcomes

We also examine data on academic outcomes from administrative records. As reported in Table 9, we find no statistically significant differences between treatment conditions or between men and women in participants' grade point averages in the program at the conclusion of their Fall Semester or at the conclusion of their junior year. Given the fact that their entering GPAs were also similar and that a significant portion of the individual grade was composed of group-based assessments, this result is perhaps not surprising.

However, we should be cautious about assuming that this average treatment effect means there was no relationship between treatment condition and future academic performance

Table 9: Grades

	1F Women	3F Women	0F Men	1F Men	3F Men
GPA in fall semester	3.54	3.56	3.55	3.55	3.58
GPA in school year	3.56	3.58	3.57	3.55	3.58
N	39	105	152	154	72

Note: We found no statistically significant differences at 10 percent.

for any participants. In Table 10, we explore the relationship between student GPA in each semester’s accounting courses and influence, controlling for academic performance in prerequisite courses.

Table 10: Relationship between Group Influence and Performance

	(1) Women		(3)	(4) Men		(5)
	1F	3F		0F	1F	
	<i>Fall Semester</i>					
Influence Votes Received	0.215*** (0.063)	0.092*** (0.023)	0.076*** (0.015)	0.074*** (0.019)	0.079*** (0.023)	
GPA (Prerequisites)	0.305 (0.333)	0.682*** (0.144)	1.044*** (0.154)	0.814*** (0.129)	0.776*** (0.198)	
R-squared	0.453	0.334	0.410	0.348	0.478	
<i>Second Semester</i>						
Influence Votes Received	0.153** (0.065)	0.078*** (0.025)	0.082*** (0.022)	0.058*** (0.021)	0.061** (0.024)	
GPA (Prerequisites)	0.621 (0.375)	0.702*** (0.182)	0.755*** (0.177)	0.783*** (0.125)	0.883*** (0.133)	
R-squared	0.337	0.319	0.288	0.295	0.469	
Observations	38	107	153	154	71	

Note: Cell entries are OLS regression coefficients. Standard errors clustered by group. Dependent variable is the student’s GPA in accounting classes, by semester.  $M_{fall} = 3.56$ ,  $SD = 0.32$ ,  $M_{second} = 3.57$ ,  $SD = 0.33$ . Influence votes is defined as the average survey influence votes received during the semester. Controls for international student, age, race, and marital status are included in regression models but not shown. Full results available in the appendix.

In these models, we find that survey influence votes over the course of the semester are

correlated with final grades both that semester and the following one. This correlation is much stronger for 1F women, though. And, perhaps even more surprisingly, once controls for influence votes are added to the model, prerequisite GPA has no relationship with future grades for 1F women, despite being a strong correlate for all other participants.

This result is sobering and cause for future research. We show that achieving influence is challenging for token women, and it appears that the ability to achieve such influence may have long term consequences: token women are only able to convert their academic preparation into better program grades when they achieve influence. Put differently, token women who were able to be seen as influential in their groups did exceptionally well, while those who were not seen as influential faltered regardless of how prepared they were for the program as they entered. This pattern persisted after the Fall Semester into later work in the program — in other words, even after token women were reassigned to different groups.<sup>20</sup>

We also find some fascinating contradictions as we look at participants' labor market plans (Appendix Table D4). Token women successfully found internships, and those who applied for jobs received offers at a fairly similar rate as others did. There are overall gender differences in labor market plans, but there are no differences between 1F and 3F women. However, token women are quite a bit more likely to express a plan to attend graduate school than any other group. This difference falls just short of statistical significance, but it is quite striking. What, if anything, to make of this finding is hard to know. It could be that the experiences of being the token woman increased women's ambition. And/or it could be that the experiences of being the token woman caused women to believe that gaining additional credentials would be necessary to be taken seriously in the labor market. These longer-term potential effects of experiencing token status, too, deserve additional research.

---

<sup>20</sup>The influence votes received in the second semester analysis are Fall, not second semester, influence. In the second semester, the coefficient attaching to prerequisite GPA is not statistically significant and still smaller than in the other models, though the differences are less stark than in the fall.



## 7 Discussion

Across multiple indicators, our results show a pattern of devaluing women’s participation and expertise in work teams, especially when they are in the minority. Token women face a steep uphill battle as they try to exert influence and acquire authority, and the typical “lean in” strategies cannot overcome these barriers. It is not difficult to imagine what effect this devaluation might have on women’s opportunities to advance in their workplaces. For example, if women have difficulty shaping group decisions on a project, it will be harder for them to make the case that they ought to get a promotion for their contributions to that project. If women are less likely to be chosen as a team spokesperson, they are less likely to get the kind of face time with bosses and clients that leads to new opportunities and a reputation as a rising star. While we find that peers gradually seem to acknowledge token women’s leadership in the abstract over a period of time, things like salary negotiations often hinge on an employee being able to point to specific contributions to team success. Token women face unique challenges in being able to do this that may compound across settings and possibly throughout their career.

Interestingly, we do not find the same kind of group gender composition treatment effects when we ask participants to identify the most and least supportive members of their group. As Figure D1 and Table D3 in the appendix show, women receive these votes at the rates we would expect in both 1F and 3F conditions. In fact, there is evidence that rather than being seen as influential, token women with high grades entering the program are eventually seen as highly supportive. Given the close cultural tie between authority and masculinity and supportiveness and femininity (Jamieson et al., 1995; Manne, 2017), one wonders the extent to which sexism plays a role in these findings. Our pretreatment measures of sexism suggest that men in our sample are considerably more sexist than women (Table A4), so a 1F environment is simply more deeply saturated in sexist beliefs than a 3F environment.

However, it is important to remember that on average the men in the 3F condition express equal baseline levels of sexism as the men in the 1F. Despite this, their behavior is significantly more egalitarian. In other words, random assignment ensures that the differences across the conditions are not a result of differences in the attitudes men and women bring to their teams. This echoes the argument that others have made that altering structures and institutions is likely to be more effective in reducing sexist behavior than attempting to alter beliefs (Bohnet, 2016).

Do we know anything about the traits of women in the 1F conditions who are able to be influential? Table D2 suggests that a strong preexisting sense of self-efficacy may be an important buffer against the devaluation from men that lone women receive in these groups. In other words, confidence matters for token women in a way that it does not matter for other women or for men. Understanding the source of this self-efficacy is far beyond the scope of this paper, but there is intriguing research that suggests experiences like single sex schooling may paradoxically better prepare girls for later competition with boys (Booth and Nolen, 2012). Hence, interventions that focus on child and young adult gender socialization may be fruitful avenues for future research.

Beyond that, there are a number of other questions that our research cannot answer but which point to excellent extensions and avenues for future study. For example, to what extent are our findings about tokenism versus about gender? Sample size limitations and the relative shortage of women in our study population prevented us assigning groups with token men. Previous laboratory work suggests that men in female-dominated settings do not face the same disadvantages as women in male-dominated settings (Born et al., 2020; Karpowitz and Mendelberg, 2014; Ott, 1989), but to our knowledge this has not yet been tested in a field experimental setting. Furthermore, to what extent are our results driven by the 1F woman being the lone woman versus just being outnumbered? The literature on gender and critical mass is robust, but generally not well causally identified (Sarah and

[Mona, 2008](#); [Dahlerup, 2006](#); [Joecks et al., 2013](#); [Kanter, 1977](#)). We had a small number of groups with two women and replicate our analyses with them grouped in with the 1F women. We find few differences in our results (Appendix C) when these groups are included in the analysis. Nevertheless, much work remains to understand the details of how gender composition changes group behavior. Finally, as the workplace increasingly moves online, questions arise about the role that technology plays in exacerbating or ameliorating gender gaps in deliberative settings. How does participation in a conference call or virtual meeting compare to in-person discussions among colleagues?

Our findings suggest there are strong reasons to investigate further how the gender composition of a workplace contributes to persistent gender gaps in participation, pay and advancement, especially in historically male-dominated sectors. In this paper, we show that even when women are highly qualified and do everything “right,” women’s influence in a group setting is quite clearly shaped by factors outside their control.

## References

- Anelli, Massimo and Giovanni Peri (2019), “The effects of high school peers’ gender on college major, college performance and income.” *The Economic Journal*, 129, 553–602.
- Apesteguia, Jose, Ghazala Azmat, and Nagore Iriberri (2012), “The impact of gender composition on team performance and decision making: Evidence from the field.” *Management Science*, 58, 78–93.
- Babcock, Linda, Maria P. Recalde, Lise Vesterlund, and Laurie Weingart (2017), “Gender differences in accepting and receiving requests for tasks with low promotability.” *American Economic Review*, 107, 714–747.
- Berg, Peter, Eileen Appelbaum, Thomas Bailey, and Arne L Kalleberg (1996), “The performance effects of modular production in the apparel industry.” *Industrial Relations: A Journal of Economy and Society*, 35, 356–373.
- Bohnet, Iris (2016), *What works*. Harvard university press.
- Boning, Brent, Casey Ichniowski, and Kathryn Shaw (2007), “Opportunity counts: Teams and the effectiveness of production incentives.” *Journal of Labor Economics*, 25, 613–650.
- Booth, Alison and Patrick Nolen (2012), “Choosing to compete: How different are girls and boys?” *Journal of Economic Behavior & Organization*, 81, 542–555.
- Booth, Alison L., Lina Cardona-Sosa, and Patrick Nolen (2018), “Do single-sex classes affect academic achievement? an experiment in a coeducational university.” *Journal of Public Economics*, 168, 109–126.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2016), “Stereotypes.” *The Quarterly Journal of Economics*, 131, 1753–1794.

- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2019), “Beliefs about gender.” *American Economic Review*, 109, 739–73.
- Born, Andreas, Eva Ranehill, and Anna Sandberg (2020), “Gender and willingness to lead: Does the gender composition of teams matter?” *Review of Economics and Statistics*.
- Bostwick, Valerie and Bruce Weinberg (2020), “Nevertheless she persisted? gender peer effects in doctoral stem programs.” *NBER Working Paper No. 25028*, 107.
- Bratton, Kathleen A (2005), “Critical mass theory revisited: The behavior and success of token women in state legislatures.” *Politics & Gender*, 1, 97–125.
- Brenoe, Anne Ardila and Ulf Zölitz (2019), “Exposure to more female peers widens the gender gap in stem participation.”
- Coffman, Katherine Baldiga (2014), “Evidence on self-stereotyping and the contribution of ideas.” *The Quarterly Journal of Economics*, 129, 1625–1660.
- Dahlerup, Drude (2006), “The story of the theory of critical mass.” *Politics & Gender*, 2, 511–522.
- Eckel, Catherine C. and Philip J. Grossman (2005), “Managing diversity by creating team identity.” *Journal of Economic Behavior & Organization*, 58, 371–392.
- Exley, Christine L and Judd B Kessler (2019), “The gender gap in self-promotion.” Technical report, National Bureau of Economic Research.
- Folke, Olle and Johanna Karin Rickne (2020), “Sexual harassment and gender inequality in the labor market.”
- Hamilton, Barton H, Jack A Nickerson, and Hideo Owan (2003), “Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation.” *Journal of political Economy*, 111, 465–497.

- Hill, Andrew J (2017), “The positive influence of female college students on their male peers.” *Labour Economics*, 44, 151–160.
- Isaksson, Siri (2018), “It takes two: Gender differences in group work.”, URL <https://www.dropbox.com/s/hafhh70p3li9tij/1120jmp.pdf?dl=0>. [Online; accessed 2019-03-31].
- Jamieson, Kathleen Hall, Kathleen Hall, et al. (1995), *Beyond the double bind: Women and leadership*. Oxford University Press on Demand.
- Joecks, Jasmin, Kerstin Pull, and Karin Vetter (2013), “Gender diversity in the boardroom and firm performance: What exactly constitutes a “critical mass?”.” *Journal of business ethics*, 118, 61–72.
- Kanter, Rosabeth Moss (1977), “Some effects of proportions on group life: Skewed sex ratios and responses to token women.” *American journal of Sociology*, 82, 965–990.
- Karpowitz, Christopher F and Tali Mendelberg (2014), *The silent sex: Gender, deliberation, and institutions*. Princeton University Press.
- Koffi, Marlene (2020), “Innovative ideas and gender inequality.” *Working Paper*.
- Kofoed, Michael S et al. (2019), “The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at west point.” *Journal of Human Resources*, 54, 430–467.
- Krook, Mona Lena (2010), *Quotas for women in politics: Gender and candidate selection reform worldwide*. Oxford University Press.
- Lawler, EE and Susan A Mohrman (2003), “Pay practices in fortune 1000 corporations.” *WorldatWork Journal*, 12, 45–54.

- Lazear, Edward P. and Kathryn L. Shaw (2007), "Personnel economics: The economist's view of human resources." *Journal of Economic Perspectives*, 21, 91–114.
- Lindbeck, Assar and Dennis J. Snower (2000), "Multitask learning and the reorganization of work: From tayloristic to holistic organization." *Journal of Labor Economics*, 18, 353–376, URL <https://doi.org/10.1086/209962>.
- Manne, Kate (2017), *Down girl: The logic of misogyny*. Oxford University Press.
- McDonald, Theodore W, Loren L Toussaint, and Jennifer A Schweiger (2004), "The influence of social status on token women leaders' expectations about leading male-dominated groups." *Sex roles*, 50, 401–409.
- Nix, Emily (2020), "Learning spillovers in the firm." *Working paper*.
- Ott, E Marlies (1989), "Effects of the male-female ratio at work: Policewomen and male nurses." *Psychology of Women Quarterly*, 13, 41–57.
- Preece, Jessica Robinson (2016), "Mind the gender gap: An experiment on the influence of self-efficacy on political interest." *Politics Gender*, 12, 198–217.
- Sarah, Childs and Lena Krook Mona (2008), "Critical mass theory and women's political representation." *Political studies*, 56, 725–736.
- Sarsons, Heather (2017), "Recognition for group work: Gender differences in academia." *American Economic Review*, 107, 141–145.
- Spangler, Eve, Marsha A Gordon, and Ronald M Pipkin (1978), "Token women: An empirical test of kanter's hypothesis." *American Journal of Sociology*, 84, 160–170.
- Wuchty, Stefan, Benjamin F Jones, and Brian Uzzi (2007), "The increasing dominance of teams in production of knowledge." *Science*, 316, 1036–1039.

Yoder, Janice D, Thomas L Schleicher, and Theodore W McDonald (1998), “Empowering token women leaders: The importance of organizationally legitimated credibility.” *Psychology of Women Quarterly*, 22, 209–222.

Zimmer, Lynn (1988), “Tokenism and women in the workplace: The limits of gender-neutral theory.” *Social problems*, 35, 64–77.

Zölitz, Ulf and Jan Feld (2018), “The effect of peer gender on major choice.” *University of Zurich, Department of Economics, Working Paper*.



## Appendix A: Analysis disaggregated by labs and months

Table A1: Lab Influence Votes For September Lab

VARIABLES	(1) Most Influential	(2) Most Influential	(3) Least Influential	(4) Least Influential
1F Condition	-0.390** (0.179)	-0.491*** (0.182)	0.597** (0.284)	0.586* (0.300)
Average GPA (Group)		-0.872 (1.459)		-0.725 (2.401)
Average Age (Group)		0.183* (0.104)		0.098 (0.172)
International Student in Group		0.292 (0.208)		-0.288 (0.342)
Proportion White		1.431* (0.792)		-1.038 (1.303)
Number of Married Students in Group		-0.036 (0.091)		-0.216 (0.150)
Constant	0.852***	-1.449	1.019***	2.879
Observations	75	75	75	75
R-squared	0.061	0.163	0.057	0.098

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Table A2: Lab Influence Votes For November Lab

VARIABLES	(1) Most Influential	(2) Most Influential	(3) Least Influential	(4) Least Influential
1F Condition	-0.481** (0.200)	-0.506** (0.213)	0.750** (0.301)	0.755** (0.322)
Average GPA (Group)		1.911 (1.706)		1.970 (2.582)
Average Age (Group)		0.067 (0.122)		0.058 (0.185)
International Student in Group		-0.077 (0.243)		-0.131 (0.368)
Proportion White		0.192 (0.926)		-0.718 (1.401)
Number of Married Students in Group		-0.010 (0.106)		-0.066 (0.161)
Constant	0.942*** (0.144)	-7.958 (7.388)	0.975*** (0.217)	-7.044 (11.180)
Observations	75	75	75	75
R-squared	0.073	0.098	0.078	0.093

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Table A3: Lab Spokesperson Votes By Lab

VARIABLES	(1) Lab 1	(2) Lab 1	(3) Lab 2	(4) Lab 2
1F Condition	-0.458** (0.177)	-0.569*** (0.181)	-0.216 (0.239)	-0.331 (0.247)
Average GPA (Group)		-0.764 (1.448)		0.630 (1.979)
Average Age (Group)		0.171 (0.104)		0.148 (0.142)
International Student in Group		0.156 (0.206)		0.060 (0.282)
Proportion White		0.466 (0.786)		0.602 (1.074)
Number of Married Students in Group		0.132 (0.090)		0.228* (0.123)
Constant	0.926*** (0.127)	-0.806 (6.269)	0.972*** (0.172)	-5.732 (8.568)
Observations	75	75	75	75
R-squared	0.084	0.176	0.011	0.091

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Table A4: Demographics and Pre-treatment Attitudes By Condition

	0F Groups	1F Groups	3F Groups
Age	23.8	23.6	23.1 <sup>B</sup>
White	90.8	90.6	88.8
Married	39.9	32.3	28.5
GPA	3.80	3.78	3.79
Leadership experience	45.4	48.7	53.1
International student	3.9	8.8	10.0
Parental income category	4.3	4.2	4.0
Political affiliation scale	5.8	5.7	5.6
Ambivalent sexism index	56.1	55.1	46.0 <sup>B</sup>
Egalitarianism index	51.8	56.3 <sup>A</sup>	56.4
Individualism index	69.9	71.4	68.0 <sup>B</sup>
Pro-sociality index	71.2	72.9	73.1
Conflict avoidance index	43.5	43.0	45.2
Individual efficacy index	70.1	67.8	63.0 <sup>B</sup>
Group efficacy index	64.1	62.5	57.6 <sup>B</sup>
N	153	192	179

Note: <sup>A</sup> shows significance at 5 percent level between 0F and 1F groups. <sup>B</sup> shows significance at 5 percent level between 1F and 3F groups.

## 7.1 Speaking Turns

Table A5: Relationship between Experimental Conditions and Speaking Turns

	(1) Lab 1	(2) Lab 1	(3) Lab 2	(4) Lab 2
1F Group	0.009 (0.006)	0.007 (0.006)	0.016** (0.007)	0.014** (0.006)
0F Group	0.006 (0.006)	0.002 (0.006)	0.013** (0.006)	0.008 (0.005)
Female	0.008 (0.010)	0.012 (0.009)	0.020* (0.010)	0.020** (0.009)
1F x Female	-0.016 (0.015)	-0.009 (0.015)	-0.037** (0.016)	-0.030** (0.015)
Length of Group Conversation	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000** (0.000)
International		-0.046*** (0.013)		-0.055*** (0.014)
Age		0.001 (0.001)		-0.000 (0.001)
White		0.012 (0.012)		0.005 (0.012)
Married		0.008* (0.004)		0.012** (0.006)
GPA		0.026* (0.015)		0.015 (0.015)
Constant	0.202*** (0.008)	0.059 (0.077)	0.193*** (0.007)	0.143** (0.072)
Observations	522	518	513	509
R-squared	0.003	0.075	0.014	0.090

Note: Standard errors clustered by group. Dependent variable is the individual's average proportion of speaking turns. Excluded category is 3F groups.

\*\* p<0.01, \* p<0.05, \* p<0.1

Table A6: Survey Influence Votes

	(1) Most Influential	(2) Most Influential	(3) Least Influential	(4) Least Influential
1F Condition	-0.201 (0.158)	-0.230 (0.166)	0.387 (0.243)	0.368 (0.258)
Average GPA (Group)		2.054 (1.331)		-1.000 (2.064)
Average Age (Group)		0.108 (0.095)		0.076 (0.148)
International Student in Group		0.011 (0.190)		0.071 (0.294)
Proportion White		-0.126 (0.722)		-1.120 (1.120)
Number of Married Students in Group		-0.026 (0.083)		0.025 (0.129)
Constant	0.873*** (0.114)	-9.261 (5.763)	0.895*** (0.175)	3.833 (8.935)
Observations	75	75	75	75
R-squared	0.022	0.068	0.034	0.066

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Table A7: Survey Influence Votes By Month

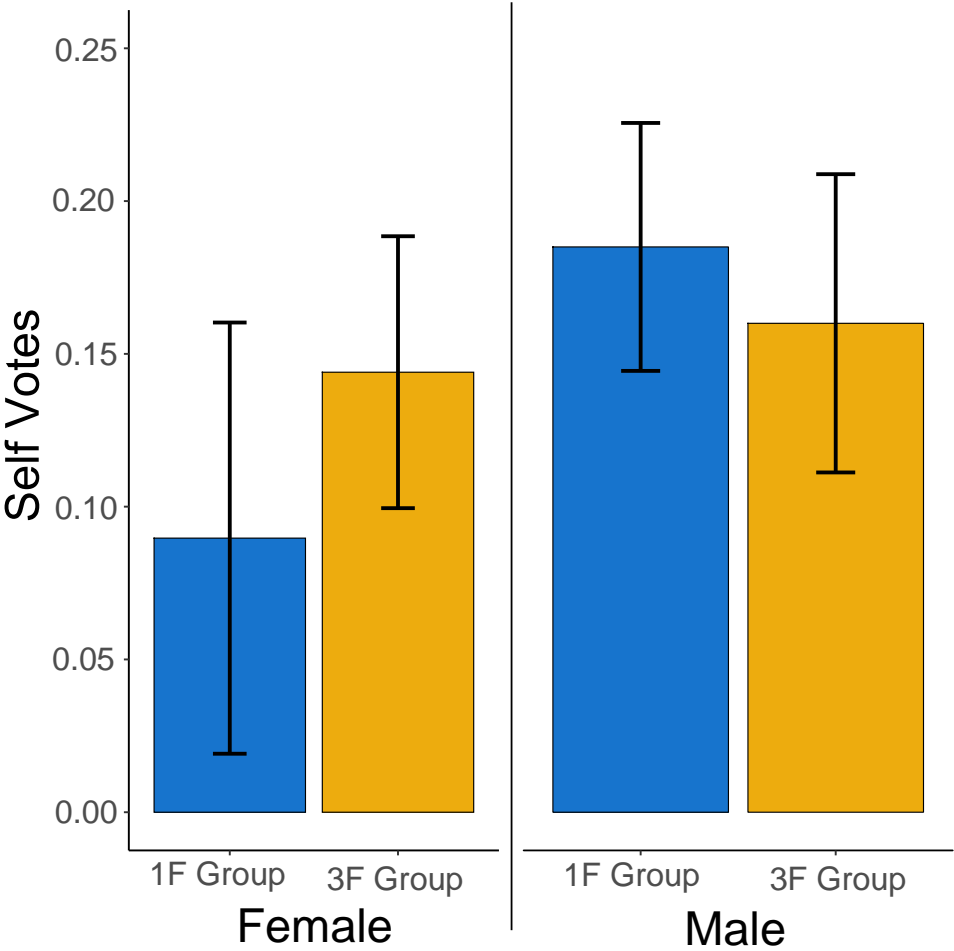
	(1) September	(2) October	(3) November	(4) December
1F Condition	-0.488*** (0.162)	-0.282 (0.222)	-0.115 (0.216)	-0.009 (0.207)
Average GPA (Group)	1.467 (1.297)	2.818 (1.776)	1.895 (1.729)	2.660 (1.660)
Average Age (Group)	0.118 (0.093)	0.155 (0.127)	0.049 (0.124)	0.111 (0.119)
International Student in Group	-0.145 (0.185)	0.058 (0.253)	0.031 (0.247)	0.076 (0.237)
Proportion White	-1.011 (0.704)	0.377 (0.964)	-0.534 (0.938)	0.581 (0.901)
Number of Married Students in Group	0.038 (0.081)	0.027 (0.111)	-0.043 (0.108)	-0.107 (0.103)
Constant	-6.544 (5.616)	-13.770* (7.689)	-6.946 (7.485)	-12.159* (7.187)
Observations	75	75	75	75
R-squared	0.163	0.070	0.032	0.063

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

# Appendix B: Robustness checks

Figure B1: Votes for Self as Most Influential



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table ?? for full results)



Table B1: Most Influential Over Group Ranking

VARIABLES	(1)	(2)
Female	-0.060*	-0.063*
	(0.030)	(0.033)
1F Group	-0.065	-0.071
	(0.050)	(0.050)
1F X Female	0.007	0.017
	(0.054)	(0.052)
International		0.026
		(0.063)
Age		-0.006
		(0.007)
White		-0.004
		(0.053)
Married		0.011
		(0.031)
Task Performance		0.331**
		(0.163)
Speaking Turns		0.768**
		(0.322)
GPA		0.097
		(0.076)
Constant	0.208***	-0.378
	(0.043)	(0.398)
Observations	373	370
R-squared	0.014	0.055

Note: Standard errors clustered by group. Influence is calculated by taking the percentage of individual rankings that are within 3 of the group ranking. Those who had the highest percentage were determined to be most influential on group decision.

\*\* p<0.01, \* p<0.05, \* p<0.1

Table B2: Influence Over Final Group Ranking - Percent of Individual Rankings Within 3 of Group

	(1)	(2)
Female	-0.047*** (0.015)	-0.038** (0.015)
1F Group	0.043 (0.027)	0.029 (0.024)
1F X Female	-0.024 (0.021)	-0.001 (0.022)
International		-0.052* (0.028)
Age		0.005 (0.003)
White		0.004 (0.028)
Married		0.006 (0.013)
GPA		-0.050 (0.035)
Task Performance		0.443*** (0.077)
Constant	0.665*** (0.025)	0.479*** (0.159)
Observations	373	370
R-squared	0.071	0.189

Note: Standard errors clustered by group. Influence is calculated by taking the percentage of individual rankings that are within 3 of the group ranking

\*\* p<0.01, \* p<0.05, \* p<0.1

Table B3: Influence Over Final Group Ranking - Total Difference Between Individual and Group Ranking

	(1)	(2)
Female	-0.050*** (0.018)	-0.027 (0.019)
1F Group	0.001 (0.048)	-0.010 (0.019)
1F X Female	-0.038 (0.025)	-0.003 (0.025)
International		-0.041 (0.028)
Age		-0.001 (0.003)
White		0.022 (0.027)
Married		0.011 (0.014)
GPA		0.006 (0.030)
Task Performance		0.814*** (0.036)
Constant	0.506*** (0.034)	0.042 (0.160)
Observations	373	370
R-squared	0.018	0.702

Note: Standard errors clustered by group. Individual group difference is made by taking the absolute value of the difference of each item ranking by the individual and the group. We then divide the minimum difference between the individual and group rankings by the individual ranking. (So, an individual group difference ratio of 1 would be the person with the lowest absolute difference.)

\*\* p<0.01, \* p<0.05, \* p<0.1

Table B4: Determinants of Empirical Influence over Group Decision

	(1)	(2)
Female	-0.063*	-0.072**
	(0.034)	(0.036)
1F Group	-0.022	-0.036
	(0.038)	(0.040)
1F X Female	0.003	0.028
	(0.045)	(0.048)
International		-0.005
		(0.041)
Age		-0.007*
		(0.004)
White		0.017
		(0.044)
Married		-0.012
		(0.028)
Task Performance		0.399***
		(0.140)
Speaking Turns		0.607**
		(0.301)
GPA		0.019
		(0.055)
Constant	0.132***	-0.149
	(0.032)	(0.256)
Observations	373	370
R-squared	0.014	0.068

Note: Standard errors clustered by group. Dependent variable is whether the individual had the smallest total difference between the individual and group rankings.  $M_{DV} = 0.10, SD = 0.23$

\*\* p<0.01, \* p<0.05, \* p<0.1

Table B5: Effect of Experimental Conditions on Probability of Voting for a Woman, by Gender

	(1) Women	(2) Women	(3) Men	(4) Men
1F Condition	-1.298*** (0.295)	-1.268*** (0.302)	-1.876*** (0.244)	-1.904*** (0.255)
International		0.233 (0.418)		-0.626 (0.467)
Age		0.101* (0.053)		-0.099 (0.073)
White		0.435 (0.376)		-0.352 (0.382)
Married		-0.572** (0.255)		0.152 (0.213)
GPA		0.766 (0.504)		-0.313 (0.574)
Constant	0.380** (0.177)	-5.070** (2.444)	0.967*** (0.194)	4.841 (3.320)
Observations	147	145	226	225

Note: Standard errors clustered by group. Dependent variable is a dichotomous indicator of whether the participant ever voted for a woman in either lab task. Cell entries are probit coefficients.

\*\* p<0.01, \* p<0.05, \* p<0.1

Table B6: Group Level Randomization Balance Table

	1F Group	3F Group
Age	23.6 <sup>A</sup>	23.1
White	90.8	88.6
Married	32.6	28.5
GPA	3.78	3.79
International student	8.8	10.0

Note: <sup>A</sup> shows significance at 5 percent level between 1F and 3F Groups.

## Appendix C: Analysis including 2F groups

Table C1: Randomization Balance Table By Gender and Condition

	Minority F Female	Majority F	No F Male	Minority F Male	Majority F Male
Age	22.5	22.7	23.8	23.9	23.8
White	76.2	86.9	90.8	94.9	91.5
Married	28.6	23.1	39.9	33.1	36.6
GPA	3.80	3.78	3.80	3.78	3.79
Leadership experience	52.3	53.2	45.4	47.8	52.8
International student	18.6	10.2	3.9	5.7	9.7
Parental income category	3.9	3.9	4.3	4.4	4.1
Political affiliation scale	5.2	5.5	5.8	5.8	5.8
Ambivalent sexism index	44.2	40.8	56.1	57.5	53.9
Egalitarianism index	59.4	56.7	51.8	55.4	55.8
Individualism index	65.1	66.0	69.9	73.4	71.0
Pro-sociality index	72.3	73.1	71.2	72.8	73.0
Conflict avoidance index	49.0	48.4	43.5	41.5	40.3
Individual efficacy index	58.1	59.1	70.1	70.4	69.0
Group efficacy index	54.4	53.9	64.1	64.6	63.1
N	43	108	153	158	71

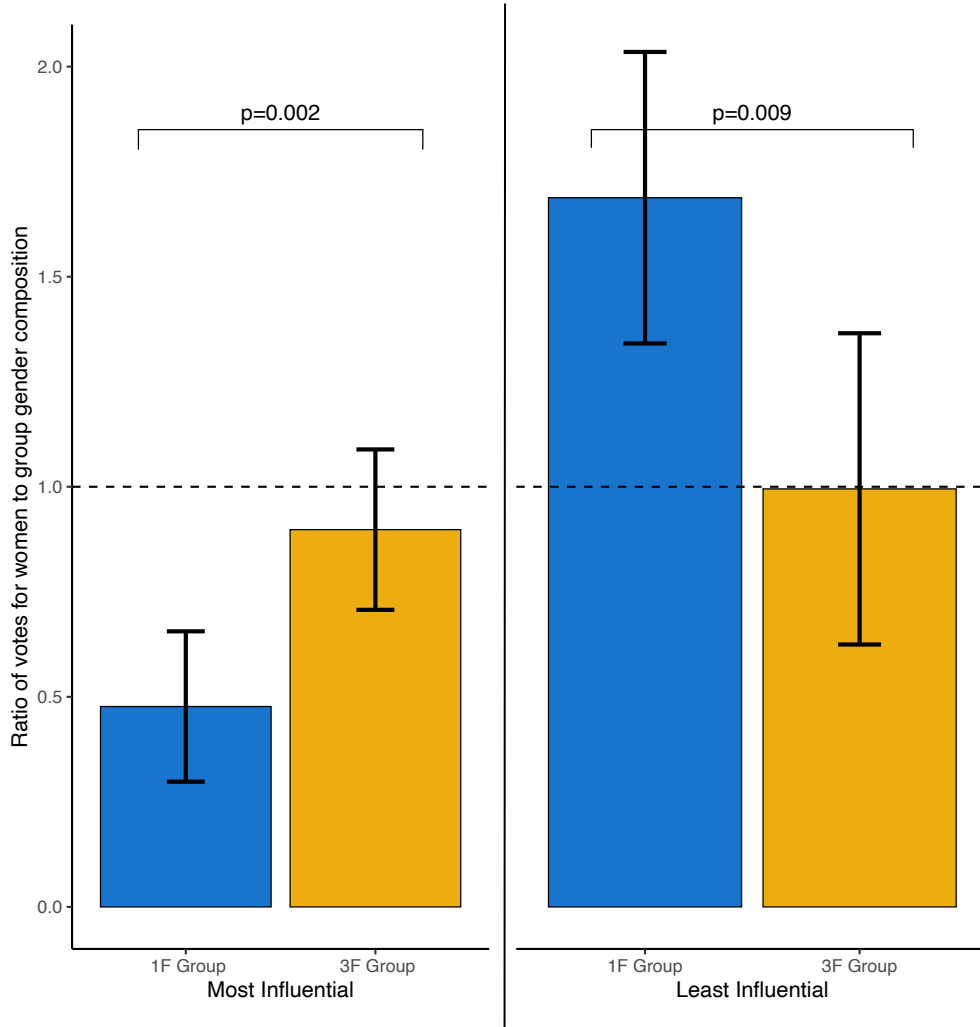
Note: <sup>A</sup> shows significance at 5 percent level between no F and minority F groups. <sup>B</sup> shows significance at 5 percent level between minority F and majority F groups.

Table C2: Randomization Balance Table

	No Female Groups	Minority Female Groups	Majority Female Groups
Age	23.8	23.6	23.1 <sup>B</sup>
White	90.8	91.0	88.8
Married	39.9	32.2	28.5
GPA	3.80	3.78	3.79
Leadership experience	45.4	48.8	53.1
International student	3.9	8.5	10.0
Parental income category	4.3	4.3	4.0
Political affiliation scale	5.8	5.7	5.6
Ambivalent sexism index	56.1	54.7	46.0 <sup>B</sup>
Egalitarianism index	51.8	56.2 <sup>A</sup>	56.4
Individualism index	69.9	71.6	68.0 <sup>B</sup>
Pro-sociality index	71.2	72.7	73.1
Conflict avoidance index	43.5	43.1	45.2
Individual efficacy group	70.1	67.8	63.0 <sup>B</sup>
Group efficacy group	64.1	62.5	57.6 <sup>B</sup>
N	153	192	179

Note: <sup>A</sup> shows significance at 5 percent level between no F and minority F groups. <sup>B</sup> shows significance at 5 percent level between minority F and majority F groups.

Figure C1: Lab Influence Votes



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table C3 for full results)



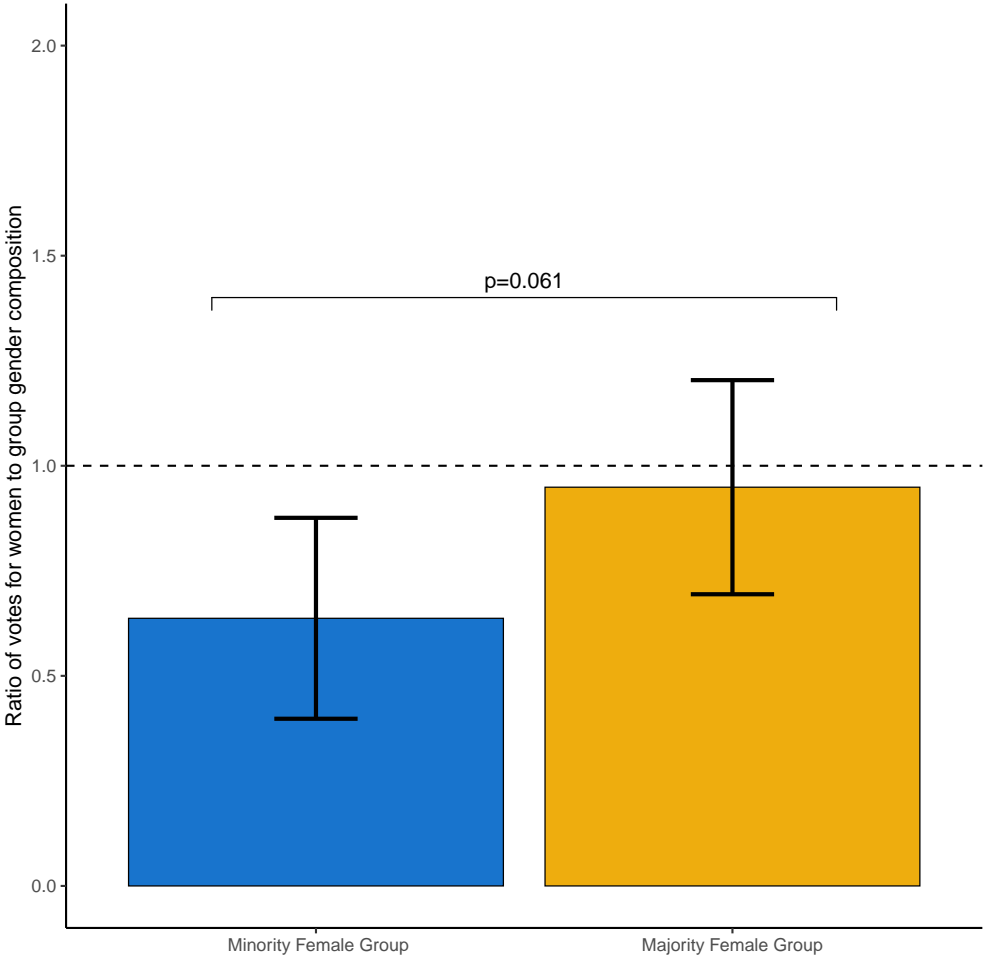
Table C3: Lab Influence Votes

VARIABLES	(1) Most Influential	(2) Most Influential	(3) Least Influential	(4) Least Influential
Minority Female Condition	-0.421*** (0.133)	-0.486*** (0.138)	0.692*** (0.259)	0.691** (0.273)
Average GPA (Group)		0.665 (1.111)		0.724 (2.195)
Average Age (Group)		0.125 (0.080)		0.067 (0.158)
International Student in Group		0.097 (0.160)		-0.213 (0.316)
Proportion White		0.838 (0.609)		-0.852 (1.204)
Number of Married Students in Group		-0.016 (0.069)		-0.152 (0.137)
Constant	0.898*** (0.097)	-5.278 (4.841)	0.995*** (0.189)	-2.231 (9.569)
Observations	77	77	77	77
R-squared	0.117	0.177	0.087	0.113

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Figure C2: Lab Spokesperson



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table C4 for full results)

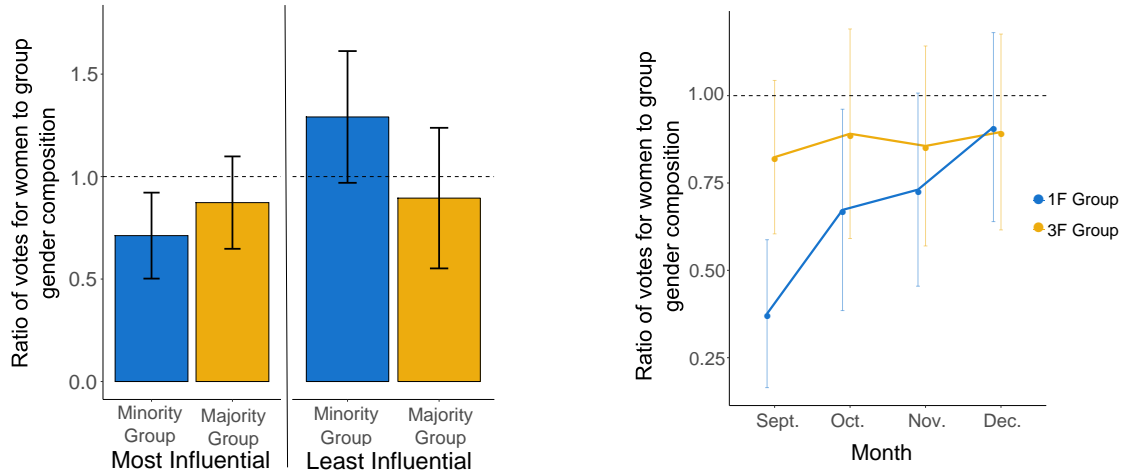
Table C4: Lab Spokesperson Votes

VARIABLES	(1)	(2)
Minority F Group	-0.216 (0.239)	-0.331 (0.247)
Average GPA (Group)		0.630 (1.979)
Average Age (Group)		0.148 (0.142)
International Student in Group		0.060 (0.282)
Proportion White		0.602 (1.074)
Number of Married Students in Group		0.228* (0.123)
Constant	0.972*** (0.172)	-5.732 (8.568)
Observations	381	377
R-squared	0.023	0.028

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Figure C3: Survey Influence Votes



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table C5 for full results)

Table C5: Survey Influence Votes

VARIABLES	(1) Most Influential	(2) Most Influential	(3) Least Influential	(4) Least Influential
Minority F Group	-0.161 (0.157)	-0.191 (0.164)	0.396 (0.239)	0.402 (0.253)
Average GPA (Group)		2.361* (1.316)		-0.837 (2.033)
Average Age (Group)		0.097 (0.095)		0.055 (0.146)
International Student in Group		-0.009 (0.190)		0.062 (0.293)
Proportion White		-0.060 (0.722)		-1.066 (1.115)
Number of Married Students in Group		-0.027 (0.082)		-0.001 (0.127)
Constant	0.873*** (0.115)	-10.209* (5.737)	0.895*** (0.175)	3.717 (8.860)
Observations	77	77	77	77
R-squared	0.014	0.067	0.035	0.062

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Table C6: Academic and Job Placement Outcomes

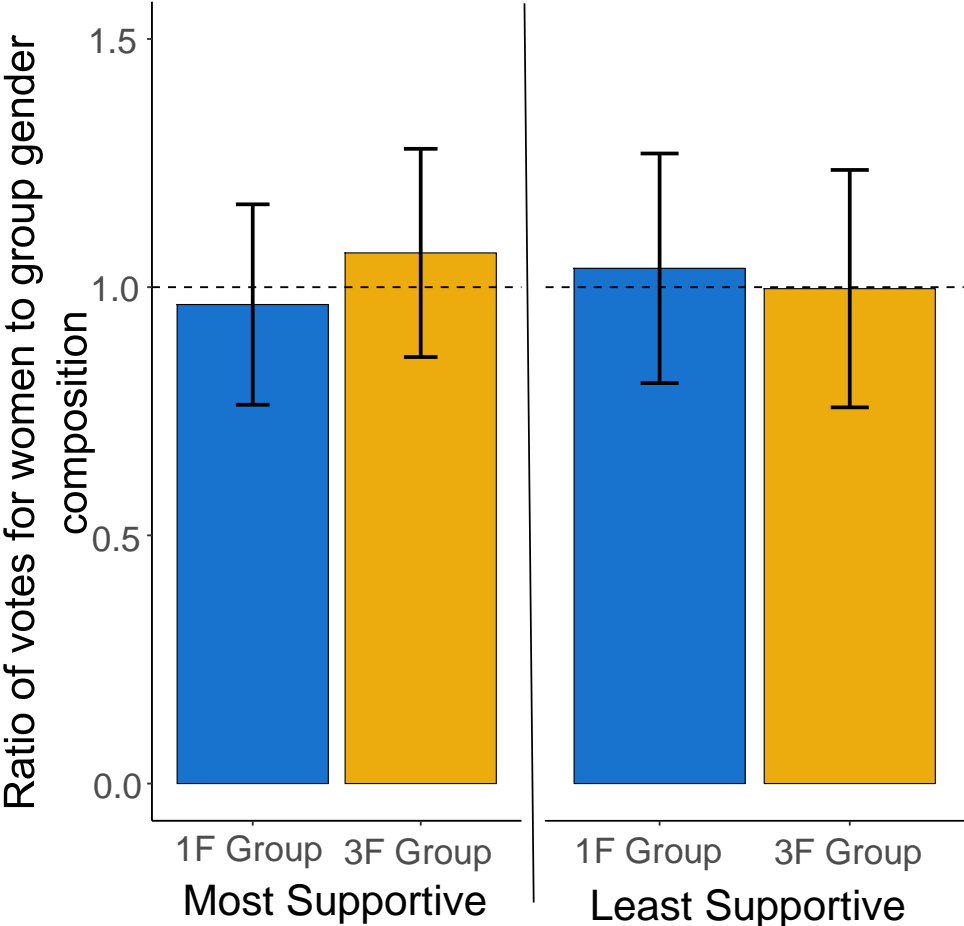
	Minority F Women	Majority F Women	No F Men	Minority F Men	Majority F Men
GPA in fall semester	3.55	3.56	3.55	3.55	3.58
GPA in school year	3.57	3.58	3.57	3.55	3.59
Number of internships	1.12	1.15	1.05	1.04	1.04
Number of job offers	1.5	1.19	1.49	1.49	1.47
Grad school	76.3	58.1 <sub>A</sub>	59.0	58.8	53.1
Full time work plans	31.6	37.6	45.5	43.9	40.6
N	38	93	134	148	64

Note: <sub>A</sub> refers to a p-value < 0.05.

Number of job offers is conditional on applying for jobs N=12, 43, 65, 78, 30

# Appendix D: Additional Analysis

Figure D1: Lab Supportiveness Votes



Note: 95% confidence intervals. Bars represent the average proportion of votes for women divided by the proportion of women in the group. (See Table D1 for full results)

Table D1: Lab Supportiveness Votes

VARIABLES	(1) Most Supportive	(2) Most Supportive	(3) Least Supportive	(4) Least Supportive
1F Condition	-0.105 (0.149)	-0.103 (0.157)	0.042 (0.171)	0.026 (0.182)
Average GPA (Group)		2.213* (1.257)		-1.432 (1.458)
Average Age (Group)		0.037 (0.090)		0.030 (0.104)
International Student in Group		-0.014 (0.179)		0.058 (0.208)
Proportion White		0.137 (0.682)		-0.098 (0.791)
Number of Married Students in Group		-0.041 (0.078)		-0.019 (0.091)
Constant	1.069*** (0.108)	-8.226 (5.444)	0.997*** (0.123)	5.808 (6.312)
Observations	75	75	75	75
R-squared	0.007	0.057	0.001	0.020

Note: Group-level analysis. Dependent variable coded by taking the proportion of votes for women divided by the proportion of women in the group.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Table D2: Correlation between Pre-Treatment Attributes and Influence Votes, by Gender and Experimental Condition

	Women		Men		
	(1) 1F	(2) 3F	(3) 0F	(4) 1F	(5) 3F
GPA	0.342 (0.541)	0.371 (0.644)	0.457 (0.575)	-0.753 (0.680)	0.219 (0.956)
Egalitarianism	-0.399 (0.804)	0.839 (0.840)	-0.840 (0.639)	0.366 (0.884)	1.386 (1.506)
Individualism	0.888 (0.885)	-1.418** (0.626)	0.062 (0.598)	0.319 (0.665)	2.269** (1.056)
Prosociality	-0.077 (1.328)	1.201 (1.210)	1.251 (1.241)	-1.003 (1.067)	1.461 (2.004)
Ambivalent Sexism	1.154 (0.739)	-1.237 (0.819)	-0.381 (0.758)	-0.488 (0.763)	-0.822 (1.648)
Conflict Avoidance	0.906 (1.353)	1.136 (1.083)	-0.100 (0.834)	0.124 (1.103)	-2.192 (1.515)
Comfort with Confrontation	0.809 (1.075)	0.521 (0.545)	-0.302 (0.700)	0.776 (0.791)	-0.722 (1.848)
Opinionation	-0.075 (0.182)	0.267 (0.225)	-0.164 (0.128)	0.150 (0.125)	0.026 (0.182)
Self-Efficacy (Previous Groups)	1.807** (0.698)	0.158 (0.746)	0.514 (0.848)	1.190 (1.018)	-1.065 (1.354)
Group Efficacy (Previous Groups)	-0.474 (0.523)	-0.404 (0.504)	0.226 (0.568)	-0.084 (0.489)	0.991 (0.901)
Risk Acceptance	-0.167 (0.111)	-0.045 (0.072)	-0.040 (0.081)	-0.076 (0.076)	-0.157 (0.130)
High School Leadership	0.079 (0.285)	-0.004 (0.233)	-0.104 (0.214)	0.164 (0.189)	0.282 (0.301)
Constant	-2.661 (3.035)	-1.715 (3.145)	-0.541 (2.854)	3.168 (3.095)	0.238 (5.127)
Observations	38	108	153	153	71
R-squared	0.360	0.156	0.042	0.074	0.143

Note: Standard errors clustered by group. Dependent variable is average influence votes received in the lab tasks.

\*\* p<0.01, \* p<0.05, \* p<0.1



Table D3: Correlation between Pre-Treatment Attributes and Supportiveness Votes, by Gender and Experimental Condition

	Women		Men		
	(1) 1F	(2) 3F	(3) 0F	(4) 1F	(5) 3F
GPA	0.947** (0.440)	-0.356 (0.327)	0.269 (0.336)	0.149 (0.358)	-0.353 (0.615)
Egalitarianism	-0.909 (0.842)	-1.514** (0.668)	0.248 (0.429)	-0.140 (0.495)	1.001* (0.590)
Individualism	-1.124 (0.720)	0.269 (0.606)	0.462 (0.338)	-0.456 (0.428)	0.321 (0.648)
Prosociality	0.537 (1.214)	-0.390 (0.775)	0.590 (0.735)	0.759 (0.589)	0.590 (1.058)
Ambivalent Sexism	-2.025** (0.787)	-0.557 (0.624)	0.489 (0.517)	0.169 (0.580)	0.994 (1.044)
Conflict Avoidance	0.004 (1.002)	0.705 (0.605)	0.843 (0.665)	0.451 (0.604)	0.401 (1.014)
Comfort with Confrontation	-0.426 (0.867)	0.857 (0.574)	0.166 (0.593)	0.017 (0.465)	1.879* (1.074)
Opinionation	0.369** (0.162)	-0.180 (0.150)	-0.057 (0.108)	-0.039 (0.090)	-0.018 (0.153)
Self-Efficacy (Previous Groups)	0.168 (0.832)	-0.595 (0.801)	-0.011 (0.663)	-0.018 (0.680)	-0.433 (1.080)
Group Efficacy (Previous Groups)	0.358 (0.492)	-0.370 (0.355)	0.208 (0.327)	0.272 (0.266)	-0.022 (0.538)
Risk Acceptance	0.198* (0.103)	0.108** (0.042)	-0.031 (0.042)	-0.065 (0.051)	-0.033 (0.079)
High School Leadership	0.093 (0.232)	0.061 (0.173)	0.087 (0.130)	0.352*** (0.118)	0.205 (0.227)
Constant	-2.929 (1.947)	3.181** (1.537)	-1.626 (1.277)	0.104 (1.604)	-0.439 (2.662)
Observations	38	108	153	153	71
R-squared	0.573	0.195	0.065	0.087	0.137

Note: Standard errors clustered by group. Dependent variable is average influence votes received in the lab tasks.

\*\* p<0.01, \* p<0.05, \* p<0.1

Table D4: Professional Outcomes

	1F Women	3F Women	0F Men	1F Men	3F Men
Number of internships	1.2	1.15	1.05	1.05	1.04
Number of job offers	1.36	1.19	1.49	1.48	1.47
Grad school plans	73.5	58.1	59.0	58.3	53.1
Full time work plans	32.4	37.6	45.5	43.1	40.6
N	34	93	134	144	64

Note: We found no statistically significant differences at 10 percent.  
Number of job offers is conditional on applying for jobs N=11, 43, 65, 75, 30

Table D5: Relationship between Group Influence and Fall Semester Performance

	(1)		(2)	(3)		(4)	(5)
	Women			Men			
	1F	3F	0F	1F	3F		
Influence Votes Received	0.215***	0.092***	0.076***	0.074***	0.079***		
	(0.063)	(0.023)	(0.015)	(0.019)	(0.023)		
GPA (Prerequisites)	0.305	0.682***	1.044***	0.814***	0.776***		
	(0.333)	(0.144)	(0.154)	(0.129)	(0.198)		
International	0.015	-0.095	0.051	-0.076	-0.257**		
	(0.148)	(0.075)	(0.081)	(0.067)	(0.099)		
Age	-0.006	-0.006	-0.026**	-0.012	0.005		
	(0.034)	(0.009)	(0.010)	(0.015)	(0.035)		
White	0.096	-0.047	0.056	-0.076	0.082		
	(0.119)	(0.048)	(0.060)	(0.061)	(0.117)		
Married	0.074	0.183***	0.061	0.063*	-0.005		
	(0.080)	(0.058)	(0.041)	(0.036)	(0.057)		
Constant	2.299	1.065	0.060	0.751	0.386		
	(1.613)	(0.635)	(0.610)	(0.702)	(1.197)		
Observations	38	107	153	154	71		
R-squared	0.453	0.334	0.410	0.348	0.478		

Note: Cell entries are OLS regression coefficients. Standard errors clustered by group. Dependent variable is the student's GPA in Fall Semester accounting classes. Influence votes is defined as the average survey influence votes received during the semester.

Table D6: Relationship between Group Influence and Second Semester Performance

	(1)	(2)	(3)	(4)	(5)
	Women		0F	Men	
	1F	3F		1F	3F
Influence Votes Received	0.153** (0.065)	0.078*** (0.025)	0.082*** (0.022)	0.058*** (0.021)	0.061** (0.024)
GPA (Prerequisites)	0.621 (0.375)	0.702*** (0.182)	0.755*** (0.177)	0.783*** (0.125)	0.883*** (0.133)
International	0.202* (0.117)	0.064 (0.119)	0.344*** (0.108)	-0.006 (0.117)	-0.056 (0.160)
Age	-0.015 (0.046)	-0.004 (0.009)	-0.026 (0.019)	-0.018 (0.017)	0.016 (0.016)
White	0.173** (0.082)	0.044 (0.078)	0.044 (0.058)	-0.067 (0.102)	0.316 (0.221)
Married	0.150 (0.119)	0.183*** (0.055)	0.084 (0.051)	0.089** (0.041)	-0.016 (0.051)
Constant	1.239 (1.909)	0.879 (0.790)	1.147 (0.787)	1.000 (0.719)	-0.504 (0.650)
Observations	38	107	153	154	71
R-squared	0.337	0.319	0.288	0.295	0.469

Note: Cell entries are OLS regression coefficients. Standard errors clustered by group. Dependent variable is the student's GPA in second semester accounting classes. Influence votes is defined as the average survey influence votes received during the semester.

## Appendix E: Robustness Checks with Stacked Dataset

Table E1: Robustness Checks for Table B4: Determinants of Empirical Influence over Group Decision

	(1)	(2)
Female	-0.055** (0.027)	-0.056** (0.028)
1F Group	-0.009 (0.036)	-0.020 (0.036)
1F X Female	-0.004 (0.040)	0.016 (0.041)
International		0.000 (0.043)
Age		-0.005* (0.003)
White		0.020 (0.046)
Married		-0.015 (0.026)
Task Performance		0.291*** (0.079)
Speaking Turns		0.507** (0.228)
GPA		0.013 (0.054)
Constant	0.119*** (0.030)	-0.092 (0.255)
Observations	721	702
R-squared	0.008	0.044

Note: Standard errors clustered by group. Dependent variable is whether the individual had the smallest total difference between the individual and group rankings.  $M_{DV} = 0.11$ ,  $SD = 0.31$

\*\* p<0.01, \*\*\* p<0.001, \* p<0.1

Table E2: Robustness Check for Table 5: Converting Task Performance to Lab Influence, by Gender and Experimental Condition

	(1)	(2)	(3)	(4)
	1F	1F	3F	3F
Task Errors	-0.013** (0.006)	-0.014** (0.006)	-0.004 (0.007)	-0.002 (0.008)
Female	-1.288** (0.535)	-1.364** (0.554)	0.046 (0.653)	0.119 (0.638)
Female x Task Errors	0.012 (0.010)	0.014 (0.010)	-0.006 (0.010)	-0.005 (0.010)
International		0.168 (0.307)		-0.159 (0.546)
Age		-0.008 (0.056)		0.123*** (0.038)
White		0.093 (0.257)		-0.172 (0.531)
Married		0.041 (0.154)		-0.030 (0.187)
GPA		-0.479 (0.528)		0.273 (0.496)
Constant	1.796*** (0.302)	3.714 (2.666)	1.368*** (0.416)	-2.490 (2.357)
Observations	386	384	360	356
R-squared	0.054	0.057	0.016	0.053

Note: Standard errors clustered by group. Dependent variable is the individual's average number of influence votes. Excluded category is 3F groups.  $M_{DV} = 0.99$ ,  $SD = 1.17$

\*\* p<0.01, \* p<0.05, \* p<0.1

Table E3: Robustness Check for Table 6: Relationship between Experimental Conditions and Speaking Turns

	(1)	(2)
1F Condition	0.013** (0.006)	0.010* (0.005)
0F Condition	0.010* (0.006)	0.005 (0.005)
Female	0.014 (0.009)	0.016** (0.007)
1F x Female	-0.026* (0.014)	-0.019 (0.013)
Length of Group Conversation	-0.000 (0.000)	-0.000 (0.000)
International		-0.050*** (0.012)
Age		0.001 (0.001)
White		0.008 (0.011)
Married		0.010** (0.004)
GPA		0.020 (0.013)
Constant	0.197*** (0.007)	0.103 (0.064)
Observations	1,035	1,027
R-squared	0.008	0.080

Note: Standard errors clustered by group. Dependent variable is the individual's average proportion of speaking turns. Excluded category is 3F groups.  $M_{DV} = 0.20$ ,  $SD = 0.06$

\*\* p<0.01, \* p<0.05, \* p<0.1

Table E4: Robustness Check for Table 7: Converting Speaking Turns to Lab Influence, by Gender and Experimental Condition

	(1)	(2)	(3)	(4)
	1F	1F	3F	3F
Speaking Turns	5.668***	5.968***	8.544***	9.322***
	(1.504)	(1.417)	(2.389)	(2.334)
Female	0.306	0.348	0.492	0.611
	(0.398)	(0.511)	(0.647)	(0.626)
Speaking Turns x Female	-4.637**	-4.582**	-4.188	-4.347
	(1.864)	(2.191)	(3.493)	(3.427)
International		0.242		0.368
		(0.326)		(0.471)
Age		0.017		0.121***
		(0.054)		(0.034)
White		0.179		-0.130
		(0.289)		(0.425)
Married		0.012		-0.079
		(0.156)		(0.178)
GPA		-0.552		0.092
		(0.514)		(0.514)
Constant	-0.031	1.404	-0.496	-3.749
	(0.308)	(2.733)	(0.437)	(2.228)
Observations	377	375	355	351
R-squared	0.076	0.081	0.080	0.123

Note: Standard errors clustered by group. Dependent variable is the individual's average number of influence votes. Excluded category is 3F groups.  $M_{DV} = 0.99$ ,  $SD = 1.17$

\*\* p<0.01, \* p<0.05, \* p<0.1

Table E5: Robustness Check for Table 8: Effect of Experimental Conditions on Probability of Voting for Self, by Gender

	(1)	(2)	(3)	(4)
	Women	Women	Men	Men
0F Condition			0.062 (0.128)	0.078 (0.128)
1F Condition	-0.278 (0.222)	-0.241 (0.223)	0.099 (0.136)	0.083 (0.139)
International		0.542 (0.362)		0.410 (0.292)
Age		0.037 (0.028)		0.002 (0.035)
White		0.834* (0.457)		0.043 (0.233)
Married		-0.184 (0.299)		-0.126 (0.134)
GPA		0.511 (0.469)		-0.264 (0.398)
Constant	-1.065*** (0.113)	-4.608** (2.164)	-0.996*** (0.109)	-0.064 (1.882)
Observations	294	290	760	756

Note: Standard errors clustered by group. Dependent variable is a dichotomous indicator of whether the participant ever voted for him or herself in either lab task. Cell entries are probit coefficients.

$M_{DV} = 0.16, SD = 0.37$   
 \*\*  $p < 0.01$ , \*  $p < 0.05$ , \*  $p < 0.1$



## Appendix F: Lab Protocol

The groups met twice a semester to complete a lab task (once in September, shortly after groups were assigned, and the second time in December, just prior to the final exams). They selected a time they were all able to meet during a certain week of the semester that did not conflict with Accounting exams or assignment due dates. Groups met in small study rooms in the main business building on campus. Each group member was assigned a random seat at the study table. Each group was randomly assigned a research assistant who explained the instructions and facilitated the session. Each participant received a show-up fee in addition to the specific incentives for each individual stage, as described below.

Each session consisted of three stages and subjects were given instructions for each stage separately. In Stage 1, subjects had 10 minutes to complete the ranking activity individually. In September subjects were given the "Survival on the Moon" ranking activity, and in December they were given the "Lost at Sea" activity. The research assistant remained in the room to prevent discussion, collaboration or cheating. Subjects were told that the person with the highest number of correct answers would be paid 50 dollars after the session. See [F1](#) and [F2](#) below for specific subject instructions for September and December respectively.

At the end of Stage 1, the research assistant equipped each individual with a personal recording device to record subsequent interactions. In Stage 2, the group had 15 minutes to complete the same ranking task, but as a group. Groups were instructed to spend time discussing and completing the task together in collaboration with each other. Subjects were told that the group ranking that had the highest number of correct answers would receive 50 dollars per group member. See [F3](#) for specific subject instructions to this stage.

After this task was completed, each group member was asked to silently complete the exit questionnaire, which included questions about their group interaction during Stage 2 and a spokesperson vote. Specifically, students were asked to select one spokesperson per

group. The spokesperson would present the results of the group ranking to a panel of judges at a later date and would have an opportunity to earn an additional 50 dollars for the group (see F4).

Figure F1: Stage 1 - Individual Lab Instructions for September Lab

**Stage 1 Packet**

**Instructions:**  
In this packet, you will read a scenario and complete the following task *individually*. You have 10 minutes to complete the task. **You are NOT allowed to use your cell phone or any other electronic device at any time during this lab activity. Please complete this task on your own privately without communicating with other members of your group.** It is in your best interest to do this task on your own and not to share answers, because the student in the \_\_\_\_\_ with the answer that most closely matches expert judgments will receive \$50. This is a competition with all other students in the \_\_\_\_\_.

Do **NOT** turn the page until the research assistant signals you to begin.

Read the scenario and rank the items.

*Scenario:*  
"You are a member of a space crew originally scheduled to rendezvous with a mother ship on the lighted surface of the moon. However, due to mechanical difficulties, your ship was forced to land at a spot some 200 miles from the rendezvous point. During reentry and landing, much of the equipment aboard was damaged and, since survival depends on reaching the mother ship, the most critical items available must be chosen for the 200-mile trip. Below are listed the 15 items left intact and undamaged after landing. Your task is to rank order them in terms of their importance for your crew in allowing them to reach the rendezvous point. Place the number **1** by the most important item, the number **2** by the second most important, and so on through number **15** for the least important."

*Items to Rank:*

- \_\_\_\_\_ Box of matches
- \_\_\_\_\_ Food concentrate
- \_\_\_\_\_ 50 feet of nylon rope
- \_\_\_\_\_ Parachute silk
- \_\_\_\_\_ Portable heating unit
- \_\_\_\_\_ Two .45 caliber pistols
- \_\_\_\_\_ One case of dehydrated milk
- \_\_\_\_\_ Two 100 lb. tanks of oxygen
- \_\_\_\_\_ Stellar map
- \_\_\_\_\_ Self-inflating life raft
- \_\_\_\_\_ Magnetic compass
- \_\_\_\_\_ 20 liters of water
- \_\_\_\_\_ Signal flares
- \_\_\_\_\_ First aid kit, including injection needle

\_\_\_\_\_ Solar-powered FM receiver-transmitter

*Please answer the following questions:*  
**Q1.** A group of experts have evaluated this task and have ranked the items. How confident are you that your ranking matches the ranking of the experts? Please indicate your level of confidence in the box below with a percent from 0% to 100%.

%

End of Stage 1. Please sit quietly until RA signals end of task.

Note: Packet also contained extra pages and space for notes.

## Figure F2: Stage 1 - Individual Lab Instructions for December Lab

### Stage 1 Packet

#### Instructions:

In this packet, you will read a scenario and complete the following task *individually*. You have 10 minutes to complete the task. **You are NOT allowed to use your cell phone or any other electronic device at any time during this lab activity. Please complete this task on your own privately without communicating with other members of your group.** It is in your best interest to do this task on your own and not to share answers, because the student in the \_\_\_\_\_ with the answer that most closely matches expert judgments will receive **\$50**. This is a competition with all other students in the \_\_\_\_\_ Program.

Read the scenario and rank the items.

#### Scenario:

“You have chartered a yacht with three friends, for the holiday trip of a lifetime across the Atlantic Ocean. Because none of you have any previous sailing experience, you have hired an experienced skipper and two-person crew.

Unfortunately, in mid Atlantic a fierce fire breaks out in the ships galley and the skipper and crew have been lost whilst trying to fight the blaze. Much of the yacht is destroyed and is slowly sinking. Your location is unclear because vital navigational and radio equipment have been damaged in the fire. Your best estimate is that you are many hundreds of miles from the nearest landfall.

You and your friends have managed to save 15 items, undamaged and intact after the fire. In addition, you have salvaged a four-man rubber life craft and a box of matches. Your task is to rank the 15 items in terms of their importance for you, as you wait to be rescued. Place the number **1** by the most important item, the number **2** by the second most important and so forth until you have ranked all **15** items.”

#### Items to Rank:

- \_\_\_\_\_ A sextant
- \_\_\_\_\_ A mirror
- \_\_\_\_\_ A quantity of mosquito netting
- \_\_\_\_\_ A 25-liter container of water
- \_\_\_\_\_ A case of ready-to-eat meals
- \_\_\_\_\_ Maps of the Atlantic Ocean
- \_\_\_\_\_ A floating seat cushion
- \_\_\_\_\_ A 10 liter can of oil/petrol mixture
- \_\_\_\_\_ A small transistor radio
- \_\_\_\_\_ 20 square feet of opaque plastic sheeting
- \_\_\_\_\_ A can of shark repellent
- \_\_\_\_\_ One bottle of 160 proof rum
- \_\_\_\_\_ 15 feet of nylon rope
- \_\_\_\_\_ 2 boxes of chocolate bars
- \_\_\_\_\_ An ocean fishing kit & pole

#### Please answer the following questions:

**Q1.** A group of experts have evaluated this task and have ranked the items. How confident are you that your ranking matches the ranking of the experts? Please indicate your level of confidence in the box below with a percent from 0% to 100%. Please use whole numbers.

%

Note: Packet also contained extra pages and space for notes.

Figure F3: Stage 2 - Group Lab Instructions for September Lab

### Stage 2 Packet

#### Instructions:

In this packet you will read the same scenario as before and complete the following task *as a group*. You will have **15 minutes** to discuss and complete the task. Your discussion will be audio-recorded. A research assistant will notify you when you have 5 minutes and 1 minute left to complete the task. When the 15 minutes have finished, you will turn in the packet to the research assistant. It is in your best interest to complete this task to the best of your group's ability, because the group with the best answer will receive **\$50 per group member**. It is also important that you **not** share information about this activity with other individuals or groups so that your chance of winning does not decrease. This is a competition with other groups in the

You may begin the task when the research assistant leaves the room.

If you choose, you may use the scratch paper and pens provided. There are five pieces of scratch paper in the packet. Please remain in your seats and do not use the white board.

Discuss the scenario with your group members and rank the items accordingly.

#### Scenario:

"You are a member of a space crew originally scheduled to rendezvous with a mother ship on the lighted surface of the moon. However, due to mechanical difficulties, your ship was forced to land at a spot some 200 miles from the rendezvous point. During reentry and landing, much of the equipment aboard was damaged and, since survival depends on reaching the mother ship, the most critical items available must be chosen for the 200-mile trip. Below are listed the 15 items left intact and undamaged after landing. Your task is to rank order them in terms of their importance for your crew in allowing them to reach the rendezvous point. Place the number **1** by the most important item, the number **2** by the second most important, and so on through number **15** for the least important."

#### Items to Rank:

- \_\_\_\_\_ Box of matches
- \_\_\_\_\_ Food concentrate
- \_\_\_\_\_ 50 feet of nylon rope
- \_\_\_\_\_ Parachute silk
- \_\_\_\_\_ Portable heating unit
- \_\_\_\_\_ Two .45 caliber pistols
- \_\_\_\_\_ One case of dehydrated milk
- \_\_\_\_\_ Two 100 lb. tanks of oxygen
- \_\_\_\_\_ Stellar map
- \_\_\_\_\_ Self-inflating life raft
- \_\_\_\_\_ Magnetic compass
- \_\_\_\_\_ 20 liters of water
- \_\_\_\_\_ Signal flares
- \_\_\_\_\_ First aid kit, including injection needle
- \_\_\_\_\_ Solar-powered FM receiver-transmitter

Note: Packet also contained extra pages and space for notes.

## Figure F4: Stage 3 - Post-Lab Survey for Labs

### Stage 3 Packet

Name: \_\_\_\_\_  
 Net ID: \_\_\_\_\_

(This information will be used to connect your responses to your group for purposes of analysis. No identifying information will ever be shared or published, and your responses will be kept completely confidential at all times.)

**Instructions:**

In this packet you will answer a series of confidential feedback questions about your group discussion and decisions. **Please complete this task on your own privately without communicating with other members of your group.**

For each of these questions, please answer as honestly as possible in regards to how you personally felt about the group interaction.

The answers you provide here will in no way affect your payment or any other part of your experience today. We simply want to know your honest impressions about the conversation you just had with the other members of your group. Your answers will never be shared with other members of your group.

You have 10 minutes to complete the questionnaire. When you have finished, please exit the room with your belongings and Stage 3 Packet. Leave the clipboard and pen on your chair. **DO NOT TAKE ANY LAB MATERIALS WITH YOU.** Please exit quietly so as not to disturb others while they finish the final stage. The research assistant will collect the packet from you and give you your base payment for your participation in the activity. You will be notified of the competition results once every group has had an opportunity to participate in the lab activity.

Please sign below to state that you will NOT share information about this lab activity with other students in the \_\_\_\_\_ program until after all groups have had the opportunity to participate in the activity and the presentations have been given. This is in your best interest, because sharing information will decrease your chance and your group's chance of winning the cash prizes.

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

**Q4.** Of the group members present, select the member who was the *most influential* member of your group during the group discussion. This can include yourself. (Indicate using the letter on the nameplate in front of the group members or refer to the seating chart sheet attached to this packet. Even though you may feel that several members of the group were influential, please mark only *one* option. Choose the person you feel was *most influential*. If you are in a group of four members, do not mark option E.)

- A       D  
 B       E  
 C

**Q5.** Of the group members present, select the member who was the *least influential* member of your group during the group discussion. This can include yourself. (Indicate using the letter on the nameplate in front of the group members or refer to the seating chart sheet attached to this packet. Even though you may feel that several members of the group were not influential, please mark only *one* option. Choose the person you feel was *least influential*. If you are in a group of four members, do not mark option E.)

- A       D  
 B       E  
 C

**Q6.** Of the group members present, select the member who was the *most supportive* member of your group during the group discussion. This can include yourself. (Indicate using the letter on the nameplate in front of the group members or refer to the seating chart sheet attached to this packet. Even though you may feel that several members of the group were supportive, please mark only *one* option. Choose the person you feel was *most supportive*. If you are in a group of four members, do not mark option E.)

- A       D  
 B       E  
 C

**Q7.** Of the group members present, select the member who was the *least supportive* member of your group during the group discussion. This can include yourself. (Indicate using the letter on the nameplate in front of the group members or refer to the seating chart sheet attached to this packet. Even though you may feel that several members of the group were not supportive, please mark only *one* option. Choose the person you feel was *least supportive*. If you are in a group of four members, do not mark option E.)

- A       D  
 B       E  
 C

**Q1.** A group of experts have evaluated this task and have ranked the items. How confident are you that your *group's* ranking from Stage 2 matches the ranking of the experts? Please indicate your level of confidence in the box below with a percent from 0% to 100%. Please use whole numbers.

%

**Q2.** Have you ever done the task – the scenario and item ranking – found in Stages 1 and 2 before in another setting?

Yes     No

**Q3.** Below you will find a series of statements about the group discussion that you may or may not agree with. Please indicate the extent to which you agree or disagree with each of the following statements by clicking on the appropriate button. Remember, there are no right or wrong answers to any of these questions. We are only interested in your own personal opinions.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
My opinions were influential in shaping the group discussion and final decision.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like my voice was heard during the group discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would describe myself as a leader in the group discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The group discussion helped me better understand the different ranking possibilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group work made everything slower and harder to accomplish in Stage 2 compared to Stage 1.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disagreement among group members made our discussion difficult.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A few people dominated the discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
All different perspectives were welcome in our discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Members of my group treated each other with respect and courtesy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Members of my group were too quick to agree with each other.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Q8.** How satisfied or dissatisfied are you with your group discussion?

- Very Satisfied  
 Somewhat Satisfied  
 Neither Satisfied nor Dissatisfied  
 Somewhat Dissatisfied  
 Very Dissatisfied

**Q9.** How satisfied or dissatisfied are you with your group's final ranking?

- Very Satisfied  
 Somewhat Satisfied  
 Neither Satisfied nor Dissatisfied  
 Somewhat Dissatisfied  
 Very Dissatisfied

**Q10.** Which of the following best captures how your group made a decision about the ranking?

- One or two group members made the decision for the group, but in the end, most others weren't sure or would have preferred something different.  
 Three or four group members made the decision, but in the end, there were one or two who weren't sure or would have preferred something different.  
 All members of the group made the decision together, and in the end, no members of the group weren't sure or would have preferred something different.

**Q11.** Is there anything else you would like to tell us about your group's discussion or decision?  
Please enter your comments below.

**Q12.** Finally, we would like you to select one team member who you would like to be the spokesperson for your group. This can include yourself. After all groups have participated in this lab activity, 5 groups will be selected at random for an opportunity to present and defend their group's choice for rankings. Each group must select one member to be their spokesperson. The spokesperson will give a 5-minute presentation explaining why their group chose the ranking order that they did. Whoever makes the most convincing case will earn another \$50 for their group. These "presentations" will occur at a later time after all groups have had an opportunity to participate in the lab activity. The spokesperson will be chosen by majority vote. In the case where no member receives majority vote, a spokesperson will be randomly chosen from the selected group.

Please mark who you would like to select as your group's spokesperson. (Indicate using the letter on the nameplate in front of the group members or refer to the seating chart sheet attached to this packet. Even if you think multiple people could serve as an effective spokesperson, please mark only *one* option. Choose the person you think would be the best spokesperson for your group. If you are in a group of four members, do not mark option E.)

- |                            |                            |
|----------------------------|----------------------------|
| <input type="checkbox"/> A | <input type="checkbox"/> D |
| <input type="checkbox"/> B | <input type="checkbox"/> E |
| <input type="checkbox"/> C |                            |

---

End of Stage 3.  
You may exit the room. Take your belongings. Leave the clipboard and pen on your chair,  
and bring this Stage 3 Packet to RA.

Note: Packet also contained extra pages and space for notes.

# Appendix G: Survey Instruments

Figure G1: August Survey

**Consent Form**

**Introduction**

The [redacted] is working with several faculty members from other [redacted] on a research project that will be used for two purposes: 1) an internal program evaluation to assess different aspects of the [redacted] experience and 2) a scholarly research study to understand the process of decision-making across individuals and in groups. This scholarly research study is being conducted by [redacted].

The program requires participation in the tasks below for the internal evaluation as part of your class credit. However, allowing the data we gather from your participation to be used as part of the research study is completely voluntary and will not affect your grade or standing in the program in any way. Signing this consent form signifies your willingness to allow your data to be used in the scholarly research study. All data will be kept confidential.

**Procedures**

The initial survey, which should be completed before school begins, will take most people about 15-20 minutes to complete.

During the semester, the following will occur:

1. You will receive a link to carry out an online survey through Qualtrics about once a month during the semester. These surveys take most people about 10 minutes to complete. Questions will address your perspectives about the program and your group dynamics. Because we want honest feedback, the content of your answers will not affect your grade or standing in the program in any way. Total time commitment for all of the surveys will be approximately 50 minutes.

If you choose to make your data available for the scholarly study, your responses will always be kept completely confidential and your name will never be attached to your responses or the study results in any way.

**Questions about the Research**

If you have questions regarding this study, you may contact [redacted] for further information.

**Questions about Your Rights as Research Participants**

If you have questions regarding your rights as a research participant contact [redacted].

**Statement of Consent:**

(Completion of surveys is still required for class credit, but data collected from the survey will not be used in scholarly research if you do not consent.)

I have read and understood the above information and consent to participate in this study. I do not wish to participate in the scholarly study and request that any data collected on me be excluded from any published academic articles. I understand that I still must complete the surveys for class credit as part of the internal program evaluation.

**Egalitarianism vs. Individualism**

Welcome to the first survey of our internal program evaluation. This first survey is the longest of the surveys with lots of questions so please bear with the time and length. We advise that you take this survey on a computer or tablet rather than a cell phone. The subsequent surveys are significantly shorter and will not be a heavy time burden for your semester's coursework.

We will ask you a series of questions in which there are no right or wrong answers. We are only interested in your personal opinions. **The answers you provide will in no way affect your grade. You will only be graded for completion of the survey.** Please answer the questions honestly and to the best of your ability so that we can help make your and future students' experience in the program better.

Thank you!

2. In addition to these surveys, you will also participate in a group activity at a location in the [redacted]. You will receive more information and a separate consent form for that activity. If you allow your data to be used in the scholarly research part of the study, **your name will never be published.** Only aggregate, unidentifiable data will be published.

**Risks/Discomforts**

There are minimal risks to you in this research study beyond what you would encounter in your everyday participation in the [redacted] program. The surveys you will take will ask standardized, commonly asked questions. However, it is possible you could encounter some potential emotional discomfort. If answering any specific question causes you emotional discomfort, you may skip it.

**Benefits**

There will be no direct benefits to you. It is hoped, however, that through your participation researchers may learn about group dynamics and decision-making and that this study will inform future [redacted] learn assignment and become an important contribution to the scholarly literature on this topic.

**Confidentiality**

The research data will be kept in a secure location and on password protected computer and only the researchers will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept in the researcher's locked office.

**Compensation**

No compensation will be given for completion of the surveys.

**Participation**

While participation in these tasks is required as part of the internal program evaluation, you do not have to grant access to the data collected during these activities for external scholarly research purposes. You have the right to refuse participation at any time in the external scholarly research study; if you do not consent, your data will not be used in scholarly data analysis nor will it be published in any scholarly article. Refusal to participate in the scholarly study will not jeopardize your class status, program placement, grade, or standing with the university.

In this first set of questions, we present you with a series of statements that you may or may not agree with. Please indicate the extent to which you agree or disagree with each of the following statements by clicking on the appropriate button. Remember there are no right or wrong answers to any of these questions.

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Relative equality of wealth is a good thing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Most people who fail should not blame the system; they have only themselves to blame.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If people work hard they almost always get what they want.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Even if people try hard, they often cannot reach their goals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our society should do whatever is necessary to make sure that everyone has an equal opportunity to succeed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
If people were treated more equally in this country, we would have fewer problems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is not really that big of a problem if some people have more of a chance in life than others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
For some people to succeed, others must fail.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Any person who is willing to work hard has a good chance at succeeding.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This country would be better off if we worried less about how equal people are.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Risk Aversion**

Rate your willingness to take risks in general with 1 being completely unwilling and 10 being completely willing.

	1	2	3	4	5	6	7	8	9	10
Willingness to take risks in general.										

For each of the following statements, please indicate your likelihood of engaging in each activity or behavior.

	Very Likely	Likely	Not Sure	Unlikely	Very Unlikely
Disagreeing with your father on a major issue.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Defending an unpopular issue that you believe in at a social occasion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Taking a job that you enjoy over one that is prestigious but less enjoyable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Arguing with a friend about an issue on which he or she has a very different opinion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I am capable of participating effectively in group discussions about important issues.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In general, I do better on most things than most people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group work makes everything harder than it needs to be.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am frequently frustrated by my inability to express my opinions to others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am confident in my abilities, even when confronting tasks I haven't done before.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like I accomplish more when I work with other people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Group Deliberation**

We want to ask you a little bit about previous group work that you've done. Think back to the most recent time you worked in a group to complete a task. What was the group (e.g. biology lab assignment) ?

Now thinking back to that last group you worked with in response to the previous question, please answer the extent to which you agree or disagree with the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
--	----------------	-------	---------	----------	-------------------

	Very Likely	Likely	Not Sure	Unlikely	Very Unlikely
Admitting that your tastes are different from those of your friends.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Prosociality**

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I try to help others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find it difficult to lend money or other belongings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I share the things that I have with my friends.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am pleased to help my friends/colleagues in their activities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I intensely feel what others feel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I help immediately those who are in need.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I easily put myself in the shoes of those who are in discomfort.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I do what I can to help others avoid getting into trouble.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Efficacy**

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
My opinions were influential in shaping the group discussion and final decision.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Members of my group treated each other with respect and courtesy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt like my voice was heard during the group discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would describe myself as a leader in the group.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group work made everything slower and harder to accomplish.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group discussions helped me better understand the issues.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Demographics**

We're halfway there! Hang in there!

The following are a series of questions for classification purposes and will help us properly analyze responses to this survey. Your answers will always be kept strictly confidential. We only report results for groups of people, not for individuals.

What do you expect your parents' 2016 income to be?

- Under \$50,000
- \$50,000-\$74,999
- \$75,000-\$99,999
- \$100,000-\$149,999
- \$150,000-\$199,999
- \$200,000-\$250,000
- Over \$250,000

In what year were you born?



What is your gender?

- Male
- Female

What is your marital status?

- Single
- Married
- Divorced
- Widowed

What is your most recent overall GPA?

Which of the following best describes your racial or ethnic background? (Select the answer that best applies to you. Select multiple answers if necessary.)

- White/Caucasian
- Black/African American
- Asian
- Hispanic/Latino
- Pacific Islander
- Other

Did you hold a leadership position in high school (e.g., president of a club, student body officer, etc.)?

- Yes
- No

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Women are too easily offended.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for "equality."	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Men should be willing to sacrifice their own well-being in order to provide financially for the women in their lives.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Many women have a quality of purity that few men possess.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Women exaggerate problems they have at work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### Need for Cognition

Some people have opinions about almost everything; other people have opinions about just some things; and still other people have very few opinions. What about you? Would you say you have opinions about almost everything, about many things, about some things, or about very few things? Please indicate which statement best describes you.

- I have opinions about almost everything
- I have opinions about many things
- I have opinions about some things
- I have opinions about very few things

Some people like to have responsibility for handling situations that require a lot of thinking, and other people don't like to have responsibility for situations like that. What about you? Do

#### Hostile and Benevolent Sexism

The following questions will be a series of statements again that you may or may not agree with. Please indicate the extent to which you agree or disagree with each of the following statements by clicking on the appropriate button. **Remember there are no right or wrong answers to any of these questions, but please be honest.**

Below is a series of statements concerning men and women and their relationships in contemporary society. Please indicate the degree to which you agree or disagree with each statement.

you like having responsibility for handling situations that require a lot of thinking, do you dislike it, or do you neither like it nor dislike it?

- Like it a lot
- Like it somewhat
- Neither like nor dislike it
- Dislike it somewhat
- Dislike it a lot

#### Need to evaluate

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
It is very important to me to hold strong opinions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I often prefer to remain neutral about complex issues.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How much of the time do people get what they deserve in life?

- Always
- Most of the time
- About half the time
- Once in a while
- Never

#### Conflict Avoidance

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
When people argue about issues, I often feel uncomfortable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If I'm sure I'm right about an issue, I don't waste time listening to other people's arguments.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I enjoy challenge the opinions of others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I usually find it easy to see issues from other people's points of view.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have no problem revealing my beliefs, even to someone who would disagree with me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I would rather not justify my beliefs to someone who disagrees with me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I do not take it personally when someone disagrees with my views.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I'm in a group, I often go along with what the majority decides is best, even if it is not what I want personally.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel upset after being involved in an argument.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I'm in a group, I stand my ground even if everyone else disagrees with me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Our government would run better if decisions were left up to successful business people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sometimes politics and the government seem so complicated that a person like me can't really understand what is going on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel that I have a pretty good understanding of the important political issues facing us today.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What people call "compromise" in politics is just selling out one's principles.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our government would run better if decisions were left up to non-elected independent experts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Confrontation and Emotion**

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
After a dispute with a neighbor, I would feel uncomfortable seeing him or her again, even if the conflict had been resolved.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I dislike when others have eye contact with me during an argument.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel more comfortable having an argument in person than over the phone.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Sunshine Democracy and Stealth Democracy**

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Openness to other people's views, and a willingness to compromise, are important for politics in a country as diverse as ours.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In a democracy like ours, there are some important differences between how government should be run and how a business should be managed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Most people are too self-interested to agree on solutions that serve the common good.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is important for elected officials to discuss and debate things thoroughly before making major policy changes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Elected officials would help the country more if they would stop talking and just take action on important problems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
If I were upset with a friend, I would discuss it with someone else rather than the friend who upset me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I have a conflict with someone, I try to resolve it by being extra nice to him or her.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I always prefer to solve disputes through face-to-face discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate the extent to which you agree or disagree with each of the following statements.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Everything should be out in the open in an argument, including emotions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It shows strength to express emotions openly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It makes me uncomfortable watching other people express their emotions in front of me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Career Aspirations**

In the next 15 years, how likely do you think it is that you will have a position as a partner or director of an accounting firm or other high executive position in a corporation or firm?

- Very Likely
- Likely
- Neither Likely nor Unlikely
- Unlikely

Very Unlikely

Please tell us a little more about what you expect your career accomplishments will be 15 years from now.

**Political Ideology**

We're almost finished! Just a couple more questions about you!

Generally speaking, do you consider yourself to be a(n):

- Republican
- Democrat
- Independent
- Something else

Would you call yourself a strong Republican or a not very strong Republican?

- Strong
- Not very strong

Would you call yourself a strong Democrat or a not very strong Democrat?

- Strong
- Not very strong

Do you think of yourself as closer to the Republican or Democratic party?

- Republican
- Democratic
- Neither

On most political matters do you consider yourself:

- Strongly conservative
- Moderately conservative
- Neither, middle of the road
- Moderately liberal
- Strongly liberal
- Don't know

**Submit**

Thanks for participating! Don't forget to click "Submit" so that your responses will be recorded and you can receive your grade for completion of this survey. You will receive the next (much shorter) survey in several weeks.

Powered by Qualtrics

Note: Survey was distributed online prior to group assignment.

Figure G2: Monthly Survey

**Introduction**

Welcome to the second survey of our internal program evaluation this semester, Fall 2016. The survey should take approximately 7 to 10 minutes to complete. We advise that you take this survey on a computer or tablet rather than a cell phone.

We will ask you a series of questions in which there are no right or wrong answers. We are only interested in your personal opinions. The answers you provide will in no way affect your grade. You will only be graded for completion of the survey. Please answer the questions honestly and to the best of your ability so that we can help make your experience and future students' experience better in the program.

Thank you!

**Group Deliberation**

Think about your group work *within the past week* and answer the questions accordingly.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Group discussions help me better understand the issues.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like my voice is heard during the group discussion.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group work makes everything slower and harder to accomplish.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My opinions are influential in shaping the group discussion and final decision.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Who has been the **least supportive** member of your group during group discussion and collaboration?

**Leadership Attribute: Competent**

Which member grasps the concepts the **fastest**?

Which member grasps the concepts the **slowest**?

**Leadership Attribute: Helpful**

Which member is the **most helpful** in solving problems when your group faces a challenge or a difference of opinion?

Which member is the **least helpful** in solving problems when your group faces a challenge or a difference of opinion?

Think about your group work *within the past week* and answer the questions accordingly.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I would describe myself as a leader in the group.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A few people tend to dominate the discussions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Members of my group treat each other with respect and courtesy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Members of my group are too quick to agree with each other.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Leadership Attributes: Influential**

On the following questions, we will ask you about attributes you and your group members possess. Please select the member of your current group who best fits the description from the dropdown list. This can include yourself.

Please be honest in your answers. This is for information purposes only. This will not affect anyone's grade. All answers will be kept strictly confidential.

Who has been the **most influential** member of your group during group discussion and collaboration?

Who has been the **least influential** member of your group during group discussion and collaboration?

**Leadership Attribute: Supportive**

Who has been the **most supportive** member of your group during group discussion and collaboration?

**Group Evaluation**

On average, do you speak more, less, or about the same amount as other group members when discussing group assignments?

- More
- The same
- Less

On average, do you come up with more, fewer, or about the same number of ideas as other group members when brainstorming group projects?

- More
- The same
- Fewer

How satisfied or dissatisfied are you with your group interactions?

- Very satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Very dissatisfied

Which of the following best captures how your group typically makes a decision about an assignment, aspects of a project, etc.?

- One or two group members make the decision for the group, but in the end, most others aren't sure or would prefer something different.
- Three or four group members make the decision, but in the end, there are one or two who aren't sure or would have preferred something different.
- All members of the group make the decision together, and in the end, no members of the group aren't sure or would prefer something different.

How often do members of your group have disagreements or differences of opinion?

- Very often
- Sometimes
- Rarely
- Never

Some combination of both

How often do you work on assignments and/or study with Accounting students who do not belong to your group?

- Very often
- Sometimes
- Rarely
- Never

#### Group Evaluation - Beginning/End of Semester

On average, how often do you meet as a group in a week?

- 4+ days
- 2-3 days
- 0-1 days

Please describe the strengths of your group.

Do you feel this is a sufficient amount of time or would you rather meet more or less?

- Sufficient amount of time
- Meet more
- Meet less

Please describe the weaknesses of your group.

On average, how often do you keep in contact digitally (e.g. via text, Google hangouts/docs, Skype, etc.) with your group in a week?

- 4+ days
- 2-3 days
- 0-1 days

Please describe how you spent your time together with your group this past week by checking all boxes that apply below.

- Worked on group projects/assignments
- Studied for quizzes/exams
- Worked on other homework assignments
- Reviewed concepts
- Met virtually (e.g., texting, Google Hangouts) to work on assignments
- Went out to lunch or dinner
- Met in a social setting (e.g., went to a movie or a party)
- Did not meet with my group

When your group gets together, do all members typically attend, or do smaller subsets of the group typically meet together?

- Everyone typically attends
- Smaller subsets of the group typically meet together

Other (please specify):

Thanks for participating! Don't forget to click "Submit" so that your responses will be recorded and you can receive your grade for completion of this survey.

#### Open-Ended Group Evaluation Question

Is there anything else you would like to tell us about your group's discussions or decision-making processes so far? Please enter your comments below.

#### Members Prior Acquaintance

Did you previously know any of the members of your group before entering the Program?

- Yes
- No

How many members of your group did you previously know?

- 1
- 2
- 3
- 4

Please describe briefly the depth of your relationship with the member(s) you already know? (E.g., close friend, former classmate, acquaintance, etc.)

Submit

Note: Surveys were distributed online in September, October, November, December, January, February, March and April. Surveys in October, November, February, and March did not include the sections "Group Evaluation - Beginning/End of Semester". Surveys in September and January included the "Members Prior Acquaintance" section.

# Figure G3: Exit Survey

**Consent Form**

**Introduction**

The [redacted] is working with several faculty members from [redacted] on a research project that will be used for two purposes: 1) an internal program evaluation to evaluate different aspects of the [redacted] experience and 2) a scholarly research study to understand the process of decision-making across individuals and in groups. This scholarly research study is being conducted by [redacted]. The program requires participation in this exit survey for the internal evaluation as part of your class credit. However, allowing these data to be used as part of the research study is completely voluntary and will not affect your grade or standing in the program in any way.

**Upon completion of this exit survey, you will be compensated \$7 via Venmo/Paypal or in person for your time.**

Signing this consent form signifies your willingness to allow your data to be used in the scholarly research study.

- Procedures**
- As part of the exit survey, the following will occur.
- you will receive a link to carry out an online survey through Qualtrics. This survey should take approximately 10 minutes to complete. Questions will address your perspectives about the program and your group dynamics, as well as your professional experiences, such as internships and employment, while being enrolled in the program.
  - In addition to the survey, we will also receive the following information from the [redacted] program:
    - Your final grades in accounting classes;
    - Your internship information;
    - Your final graduate school and/or job outcome, including job title and salary.

**Questions about Your Rights as Research Participants**

If you have questions regarding your rights as a research participant contact [redacted].

Statement of Consent:

I have read and understood the above consent and desire of my own free will to participate in this study.

I wish not to participate in this study and ask that any data collected on me for the internal program evaluation be excluded from any academic publishing.

**Welcome Page**

Welcome to the exit survey of the [redacted]. The survey should take approximately 10 - 12 minutes to complete. We advise that you take this survey on a computer or tablet rather than a cell phone. We will ask you a series of questions in which there are no right or wrong answers. We are only interested in your personal opinions. **The answers you provide will in no way affect your grade. You will be compensated \$7 for your time upon the completion of this survey.**

Please answer the questions honestly and to the best of your ability so that we can help make your experience and future students' experience better in the program. Thank you!

**Internship1**

During your time as a student in the [redacted] Program, did you apply for any on-campus or off-campus internships?

Yes  
No

**Internship2**

- If you allow your data to be used in scholarly research part of the study, your name will never be published. Only aggregate, unidentifiable data will be published.

**Risks/Discomforts**

There are minimal risks to you in this research study beyond what you would encounter in your everyday participation in the [redacted] program. However, it is possible you could encounter some potential emotional discomfort. If answering any specific question causes you emotional discomfort, you may skip it.

**Benefits**

There will be no direct benefits to you. It is hoped, however, that through your participation researchers may learn about group dynamics and decision-making and that this study will inform future [redacted] assignment and become an important contribution to the scholarly literature on this topic.

**Confidentiality**

The research data will be kept in a secure location and on password protected computer and only the researchers will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept in the researcher's locked office.

**Compensation**

You will be paid \$7 through Venmo/Paypal or in person for completing the exit survey.

**Participation**

While participation in these tasks is required as part of the internal program evaluation, **you do not** have to grant access to the data collected during these activities for external scholarly research purposes. You have the right to refuse participation at any time in the external scholarly research study; if you do not consent, your data will not be used in scholarly data analysis nor will it be published in any scholarly article. Refusal to participate will **not** jeopardize your class status, program placement, grade, or standing with the university.

If you choose to make your data available for the scholarly study, your responses will always be kept completely confidential and your name will never be attached to your responses or the study results in any way.

**Questions about the Research**

If you have questions regarding this study, you may contact [redacted] for further information.

How many internships did you apply for?

How many interviews did you have?

How many internship offers did you receive?

How many internships did you complete?

Now, mentally order your internships chronologically (with the first internship being the least recent and the last internship being the most recent). Please respond to the following questions regarding each of the internships.

**Internship3**

Mark **all that apply** to your Internship  $$(\text{m}/\text{Field}/1)$ :

Location Type:	On-campus	Off-campus
Wage Type:	Paid	Non-paid
Work Relevance:	Closely related to what I want to do professionally	Unrelated to what I want to do professionally

Describe **all that apply** to this internship experience.

Great professional experience  
Not a productive use of my time  
Made great professional contacts

Led to a job  
I felt uncomfortable in the work place  
I didn't feel prepared for most of the required responsibilities  
I felt prepared for most of the required responsibilities  
 Other:

**Internship4**

Please select the reason(s) why you did not apply for any internships during your time in the Jr. Core Accounting program. Select all that apply and use the space provided to include others not listed.

- Did not feel the need to have an internship since I was confident I could secure good employment without an internship
- Did not feel the need to have an internship since I planned to attend graduate school post-graduation
- Did not feel the need to have an internship since I did not plan to work or attend graduate school post-graduation
- Did not have time for an internship due to personal/family obligations
- Did not feel that there were good internship opportunities available
- Did not feel qualified/competitive for the internships that I was interested in
- Felt like I needed a break during the spring/summer after an intense school year
- Other reason(s):

In hindsight, do you wish you had applied for/completed an internship while being enrolled in the program?

- Yes
- No

**Plans**

We would now like to ask you some questions about each job offer. For the following questions, please order each job offer chronologically and answer the questions with regard to each job offer separately.

**Jobs3**

Job offer  :

Did you try to negotiate any terms of the job offer/contact that you received?

- Yes
- No

Which specific terms did you try to negotiate? (Select all that apply.)

- Salary
- Other financial terms (eg, signing bonus, commissions, etc)
- Non-financial benefits (medical, leave, schedule, flexible work arrangements, time off, opportunity for promotion, etc)
- Other:

On a scale of 0 to 100, how confident were you in your negotiating skills?  
(0 = not confident at all, 100 = completely confident)

0 10 20 30 40 50 60 70 80 90 100

What was the outcome of that negotiation? (Select all that apply.)

- My requests were met fully
- My requests were met partially; the employer and I met roughly in the middle
- My requests were not met and did not change the terms of the original contract

What are your **immediate** plans after graduation from the Accounting Jr. Core program?  
Select all that apply.

- Work full-time
- Work part-time
- Attend graduate school (beginning this summer or fall)
- Take some time to prepare for graduate school
- Be a homemaker/stay-at-home parent
- Other:

**Jobs1**

You indicated that your immediate plans after graduation are to work. We would now like to ask you some questions about your job search process.

How many jobs have you applied for in the last year?

How many first interviews did you have?

How many of these first interviews led to follow-up interviews?

How many formal job offers did you receive?

**Jobs2**

My requests hurt me and negatively affected the terms of the original contract

- Still negotiating
- Other:

Did you accept the final job offer?

- Yes
- No
- Still considering/negotiating

On a scale of 0 to 100, how happy are you about this job?  
(0 = extremely unhappy, 100 = extremely happy)

0 10 20 30 40 50 60 70 80 90 100

**Jobs4**

Describe your immediate plans at this point in terms of employment. (Select all that apply.)

- I am waiting to hear from other jobs that I have applied for
- I plan to continue applying for more jobs
- I have changed my mind and decided to prepare for/attend graduate school instead
- I have changed my mind and decided to not seek formal employment and be a homemaker instead
- I have accepted another job offer

**Grad School1**

You indicated that your immediate plans after graduation are to attend graduate school. We would now like to ask you some questions about your application process.

How many graduate schools did you apply for?

On a scale of 0 to 100, how confident were you in your ability to get accepted to graduate school?

(0 = not confident at all, 100 = completely confident)

0 10 20 30 40 50 60 70 80 90 100

How many graduate schools did you get accepted to?

Have you committed to attend any graduate program?

Yes

No

Still deciding

### Grad School2

Please indicate the program and university you will attend:

On a scale of 0 to 100, how happy are you about this post-graduation outcome?  
(0 = extremely unhappy, 100 = extremely happy)

0 10 20 30 40 50 60 70 80 90 100

You indicated that you initially planned to attend graduate school. Please tell us why your plans changed. (Select all that apply.)

I changed my mind and decided to seek full/part-time employment instead

I changed my mind and decided to be a homemaker instead

I was not admitted to any graduate school

I decided to wait to apply at a later time

### Open-ended comment

Is there anything else you would like to tell us about your experience in the program?

### Payment

Please indicate how you would like to receive payment for taking this survey and provide the corresponding information:

Information will only be used for payment purposes and all payment will be carried out privately.

Venmo (provide your Venmo username)

Paypal (provide your Paypal email address)

### End Message

Thank you for participating in this survey and the internal program evaluation throughout your time in the Jr. Core! Don't forget to click "Submit" so that your responses will be recorded.

Powered by Qualtrics

### Grad School3

Why did you not commit to any graduate program? (Select all that apply.)

I am waiting to hear from other graduate school(s) that I have applied for

I plan to reapply to graduate school(s) next year

I am still negotiating my graduate school offer

### Alt Plan1

Employment was not one of the options you listed as part of your immediate plans after graduation. We are curious to know if you did have any initial plans to seek full-time or part-time employment?

Yes

No

You indicated that you initially planned to seek employment. Please tell us why your plans changed. (Select all that apply.)

I changed my mind and decided to attend graduate school

I changed my mind and decided to be a homemaker instead

I was not offered a suitable position

I decided to take some time off before seeking further employment

### Alt Plan2

Graduate school was not one of the options you listed as part of your immediate plans after graduation. We are curious to know if you did have any initial plans to attend graduate school?

Yes

No

Note: Survey was distributed online after students at the end of undergraduate program.



## **Appendix H: Ethics Statement**

### **IRB Approval and Informed Consent**

This project was reviewed and approved by the Brigham Young University Institutional Review Board as Project X15228 in 2015. It was later renewed as Project X17352. Participants were fully informed that they were participating in a research study and internal program evaluation and no deception was used in the experiment. They were required to participate in the activities of the project as part of their program of school work for an internal program evaluation, but they could freely grant or deny their consent for their work to be included in the research study.

### **Context of Intervention**

The undergraduate Accounting Program we worked with has a long history of assigning students to work teams of five people to simulate the way the students are likely to work at an accounting firm. They had previously assigned teams in a systematic way to try to balance the average GPAs of the teams, and we continued this by block randomizing by GPA. Women are in the minority in the program—approximately 24% during the years we worked with them. In an effort to ensure that the men in their program had an opportunity to work with and learn from women, program directors had always spread women across as many groups as possible, meaning that most groups had a lone woman with four men.

Their extensive efforts to solicit feedback from students over many years had not indicated that there were significant problems with this approach; men's and women's grades in the program had also always been quite similar. However, when we discussed working together, the program directors were interested in data about whether grouping more women together in teams might support the women in their program in more subtle ways. Specifically, we discussed with them the ways in which group gender composition might shape deliberation.

From their perspective, having majority-woman teams was somewhat costly because of the large number of all-male teams it would create, but they were nevertheless interested in moving forward with the study.

As the project progressed, we found few blatant signs of harm to women in the program (whether in lone woman or majority women groups). Grades across sexes and across conditions remained essentially identical. Participants across conditions overwhelmingly enjoyed working with their groups and reported that decisions were made collaboratively and respectfully. The one exception to this was some early data after the first year that suggested lone women were more likely to plan to go to graduate school than women in majority women groups, but the sample size was too small to draw many conclusions.

After we finished data collection, we found that the self-reported plans for graduate school remained somewhat higher for lone women, though the finding was not statistically significant. We also found the patterns of gender and treatment differences in laboratory task influence and other outcomes that are described in the paper. These differences are real and important, as we argue in the paper, but it is also helpful to note that they take place in a context of extremely high self-reported satisfaction among men and women with group dynamics and the program itself. Nevertheless, in an effort to be as supportive as possible to women students, the program decided that it would no longer assign lone woman groups after seeing the results of the study. Thus, the study had a normatively positive outcome of giving women in future cohorts an increased likelihood of experiencing higher levels of influence and authority within the groups. We expect that this outcome will be beneficial to the women both during their time in the program and subsequently in their careers, though more can be done to investigate this hypothesis.

## Epistemic Justice

Aside from the usual ethical issues like consent and deception, we also believe it is important to consider epistemology in our discussion of the ethics of this project. Women have long expressed frustration working in male-dominated environments, even ones with colleagues they like and appreciate. Often, women themselves struggle to fully identify and explain their feelings, saying things like “I don’t know—something just felt off in that meeting” or “It just didn’t seem like I was being taken very seriously.” For the most part, these concerns have been ignored in educational and work settings; they are sometimes even dismissed with comments like “don’t overreact—everyone has experiences where they are ignored” or “but what specifically was wrong with what happened in that discussion?” Or, of course, “do you have any actual evidence for this?”

In short, the research that we do takes place in a general environment of epistemic injustice (Fricker 2007). Epistemic injustice denotes a “wrong done to someone specifically in their capacity as a knower” (1). It comes in two primary forms: testimonial injustice, “when prejudice causes a hearer to give a deflated level of credibility to a speaker’s word,” and hermeneutical injustice, “when a gap interpretive resources puts someone at an unfair disadvantage when it comes to making sense of their social experiences” (1). Women’s voices are systematically undervalued and their experiences are often not well-explained by existing narratives, theories, and research, leaving them at a loss to convey what exactly is happening to them in a way that is understandable to others. In other words, they face both testimonial and hermeneutical injustice.

We are sensitive to the fact that many before us have pointed out ways in which women’s voices get sidelined in male-dominated environments and that our study—a randomized controlled trial, seen by many as the “gold standard” of science—has the capacity to reinforce the idea that rigorous science is the only way to “know” that this is happening. Further, some may interpret our experimental approach as an endorsement of the idea that research like

this is the only suitable evidence on which to make behavioral decisions. We emphatically disagree with this kind of narrow epistemological approach that so often serves to reinforce existing hierarchies. And yet we still believe that our research has significant ability to reduce the epistemic injustices that women face.

First, our hope is that our research gives women the framework they need to trust their own sense of a situation. When we present this research in seminars, the tagline on our final slide is “Ladies, it’s not just all in your head!” Countless women have expressed deep gratitude that they now have a concrete way to talk about their experiences. When our university featured our research in the alumni magazine, the magazine article garnered more than 200,000 page views in the first week and became the most-shared article ever. Many individuals contacted us to say that sharing the article with coworkers, family members, and friends allowed them to have important conversations that they previously did not know how to have. We see this is an advance toward better resources for women to understand and explain their experiences and therefore greater hermeneutical justice.

We similarly hope that our research—by providing “the receipts” for claims many women have made over the years—increases the credibility of women’s claims about their experiences in the future. What’s more, our specific research questions and results about gendered patterns of influence and authority highlight that the status quo tends toward testimonial injustice, especially when women are outnumbered. Taken together, we see our work as a complement to, not a replacement for, the work that many are doing to create a culture that believes women.

In short, we believe that this study not only meets the ethical guidelines of the Belmont Report’s respect for persons, beneficence, and justice with regard to the costs and benefits of the study, but also fights for the important issue of epistemic justice.