Can You Hear My Voice This Time: Gender Discrimination in the Consideration of Ideas

Clio Bryant Flikkema
Wellesley College, cflikkem@wellesley.edu

Follow this and additional works at: https://repository.wellesley.edu/thesiscollection

Recommended Citation
Flikkema, Clio Bryant, "Can You Hear My Voice This Time: Gender Discrimination in the Consideration of Ideas" (2017). Honors Thesis Collection. 430.
https://repository.wellesley.edu/thesiscollection/430
Can You Hear My Voice This Time?  
Gender Discrimination in the Consideration of Ideas

Clio Bryant Flikkema

Submitted in Partial Fulfillment  
of the  
Prerequisite for Honors  
in Economics  
under the advisement of Olga Shurchkov

April 2017

© 2017 Clio Bryant Flikkema
Acknowledgments

First, an enormous thank you to my incredible advisor Olga Shurchkov. Your faith in my ability and your support has been unfailing, even while you pushed me to take this project to a level I never imagined I could. This has truly been an epic journey, and you were the best navigator I could ever ask for.

All my thanks to the staff at CLER, and particularly Marema Gaye for being such an incredible help with Z-tree.

Also thanks to all of my Wellesley friends and professors, as well as the staff at El Table for keeping me well-supplied with encouragement and coffee. To the all the wonderful people in the Wellesley College Russian Department who have been hearing about the ups and downs of this project for two years: Ура, заработало! И спасибо большое! To my parents, thanks for answering the phone at all bizarre hours of the day and night to listen to my struggles and triumphs.

Corey, thanks for forgoing practice to sit with me in the library while I brought this project to a finish. Thanks also for all your interesting questions, which helped me constantly think about this research in new ways. In our partner-to-partner discussions about dance I doubt I will ever need to apply any of the techniques I uncovered to make myself heard.

Lastly, to Professor David Lindauer, thank you for telling me so long ago that you believed I was capable of succeeding in this field. You gave me the courage to start down this path, and I am happy and grateful every day that I did.
# Table of Contents

Chapter 1: Introduction and Review of Literature .......................................................... 1

Chapter 2: Analysis of Observational Data .................................................................... 6
  2.1. Overview of Observational Data ........................................................................ 6
  2.2. Regression Model for Observational Data ......................................................... 9
  2.3. Discussion of Observation Data Regression Results ........................................... 11

Chapter 3: Experimental Approach ............................................................................. 16
  3.1. The Task and Baseline Ability ........................................................................... 16
  3.2. Group Discussion and Selection of the “Representative” .................................. 18
  3.3. Treatments ....................................................................................................... 20
  3.4. Payoffs ........................................................................................................... 24
  3.5. Procedures ...................................................................................................... 25

Chapter 4: Analysis of Experimental Data ................................................................... 26
  2.1. Overview of Experimental Data ....................................................................... 26
  2.2. Regression Model for Experimental Data ......................................................... 29
  2.3. Discussion of Experimental Data Regression Results ...................................... 31

Chapter 5: Conclusions and Further Research ............................................................. 37

Appendix: Experiment Instructions ............................................................................. 39

References .................................................................................................................... 52
Chapter 1: Introduction and Review of Literature

For one year, I served on an executive board with six men and no other women. Immediately in our first meeting I discovered that regardless of my confidence in my ideas or their merit, whenever I raised a point or made a contribution to the discussion I was systematically ignored, interrupted or had my ideas restated and attributed to one of the men in the room. Short of pounding on the table and shouting, I could find no way to make myself visible to the rest of the group, although they conscientiously considered the contributions of their male colleagues and even took the trouble to ask each others’ opinions. The questions we were discussing were not outside my area of expertise, nor was I intimidated. I was simply annoyed, and never again volunteered to serve on a board for that particular organization.

The phenomena known colloquially as ‘manerrupting’ and ‘bropropriating’ are garnering increased attention in the media and academic literature, as scholars continue to search for evidence to explain the absence of women in leadership positions in business, law, government and other fields. Although today women make up 50.8% of the US population, and earn almost 60% of undergraduate and masters degrees, they are not represented proportionally at the highest levels of many professions (Warner, n.d.). In this study, I investigate whether gender discrimination during the discussions that take place behind closed conference room doors may be impacting women’s careers.

A conflicting literature exists examining whether men and women with comparable characteristics experience different rates of promotion within their field. Studies using a variety of different data sets and techniques have found a bias in promotion and hiring that
favors men, or that favors women, or that no such effect exists favoring either gender. However, more specialized studies have revealed a plethora of potential explanations for the absence of women in the upper echelons. Researchers have used both observational data and experimental approaches to examine how gender-based behavioral differences may contribute to this outcome.

Existing research points to greater risk aversion and reluctance to compete among women, although results appear to be dependent on the circumstances used to elicit these behaviors (Shurchkov & Eckel, 2018). If women are indeed more risk averse than their male counterparts, they may be less willing to work in high-pressure leadership positions. Women are less likely to negotiate for raises or promotions, possibly as a result of corporate culture labeling assertive women as combative and unpleasant (Babcock, Laschever, Gelfand, & Small, 2003; Shurchkov & Eckel, 2018). Women are also less likely to engage in mixed-gender competitions, which may cause them to avoid highly competitive work environments (Gneezy, Niederle, & Rustichini, 2003; Niederle, 2016; Niederle & Vesterlund, 2007). Additionally Marianne Bertrand, Claudia Goldin and Lawrence F. Katz (2010) point to career breaks and working shorter hours due to childcare obligations to explain the gap in career advancement between male and female MBA students working in the finance industry.

This literature points to the troubling impact of societal conventions which encourage women to avoid behaving in ways that may appear aggressive and demanding, and to take on the majority of child-rearing and household tasks. However from the perspective of the firm, these explanations for not promoting women to upper level

\[1\] See references in Blau & Devaro, 2007 for details.
positions are to some degree justified. Competitive fields need competitive people, career breaks are bad for productivity, and it’s easy to give the rewards to those who ask. But can we be sure that women will succeed if we address the issue of childcare and teach girls from a young age to compete, negotiate and tolerate risk? To fix the leadership gap we must also investigate whether confident, competitive women who are dominant in their fields face discrimination based purely on gender, rather than discrimination based on the above traits commonly associated with being a female professional.

In spite of women’s possible aversion to competition, M. Daniele Paserman (2010) finds that male and female professional tennis players perform similarly in high pressure, high stakes environments. Paserman’s work suggests that if women are actually placed in these environments, their performance will be comparable to their male peers. Claudia Goldin shows that discrimination, rather than behavioral characteristics, may play a role in inhibiting women’s advancement by examining a change in the audition process used by symphony orchestras. With the introduction of a blind audition where the applicant plays behind a screen, the probability of a woman being hired for a position in a top symphony increased by 50% (Goldin & Rouse, 1997).

Even once gender considerations are eliminated from the hiring process, it is impossible to prevent gender from being observed during every day interactions in the workplace. Female musicians cannot play behind screens forever, nor should they need to. This study hopes to add to the existing literature on the leadership gap by revealing a hidden arena in the workplace where biases (both conscious and unconscious) may be affecting women’s chances for promotions to high-level positions once they are hired.
An individual’s prospects for promotion depend not only on his or her actual job performance, but also on reputation and visibility. Group discussions are utilized in many workplaces to improve the decision-making process, and are an important aspect of interaction between colleagues. An individual can build a good reputation and increase his or her visibility by performing well in group discussion. An employee should be able to increase the probability that manager will think of him or her when it comes time to give out promotions and raises by confidently proposing ideas with merit. The purpose of this study is to examine how men and women perform in these group discussion settings, and to see if bias is impacting the benefits that women should gain by performing well.

Experiments have been used previously to examine gender differences in behavior during group discussions. Katherine Baldiga Coffman finds that women may be less willing to contribute ideas during group discussions in stereotypically male fields, but that increasing contributions from women improves the performance of the group overall (Coffman, 2014). Similarly women have been shown to underperform due to time pressure and task stereotypes (Shurchkov, 2012). I will build on this evidence by using both observational data and an experimental approach to investigate whether gender differences in the consideration of ideas can be attributed to bias against women in group discussions, rather than to behavioral differences between men and women. Results using both approaches indicate that women are less likely to reap the benefits of proposing their best ideas, particularly women who are less dominant in the discussion.

The remainder of the paper is organized as follows. In Chapter 2, I present my observational data analysis and discuss the implications of the results. In Chapter 3, I describe in detail the design of the experimental approach. Chapter 4 examines the data
collected during the experiment and discusses these results in relation to the results from
the observational data method. Chapter 5 concludes, and suggests ideas for further
research.
Chapter 2: Analysis of Observational Data

The first part of my study is focused on a data set collected through observation of the Russian television program Что? Где? Когда? (in English, What? Where? When?). This show is a game show,\(^2\) which revolves around a highly stylized group discussion environment. Every episode features a team of five players who must attempt to correctly answer very difficult logical reasoning questions. Each question is posed to the entire team, and then the players are given one minute to discuss potential answers amongst themselves. After the discussion period, the team captain may either answer the question or designate a different player to do so. The player chosen answers on behalf of the entire team. If the team gets five answers correct before they miss five, they win the game.

2.1. Overview of Observational Data

This format allows us to examine what factors most affect a given player's probability of being chosen to answer the question on behalf of the team. Data was collected from 56 episodes of the 2011-2014 seasons of the TV show. 66 players participated during the time period covered by the data set, including 11 different captains. The data is set up as a panel: for each question in an episode, there is one observation for each player on the team, for a total of 3,816 observations. The dataset contains variables about the group discussion including which player in the group was the ‘dominant’ player (i.e. the player who most commanded the attention of the others), how confident the group as a whole seemed about whether or not they could answer correctly, and who was ultimately chosen to answer the question. The data also indicates the gender of each player.

\(^2\) Previous studies in behavioral economics have also used observational data from game shows to investigate concepts such as prospect theory (Post, Assem, J. van den Assem, Baltussen, & Thaler, 2012) and the relationship between physical attractiveness and cooperation (Darai & Grätz, 2012).
(including the captain), the confidence of the player chosen to answer when they are actually answering the question, whether any given answer is correct, and how many questions the team has answered correctly at that point in the episode.

Table 1: Observational Data Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Male Average</th>
<th>Female Average</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Questions Played</td>
<td>84.26</td>
<td>86.89</td>
<td>-2.63*</td>
</tr>
<tr>
<td>Total Questions Dominant</td>
<td>9.61</td>
<td>11.95</td>
<td>-2.34***</td>
</tr>
<tr>
<td>Percent of Questions Dominant</td>
<td>11%</td>
<td>14%</td>
<td>-0.03***</td>
</tr>
<tr>
<td>Total Questions Answered</td>
<td>16.09</td>
<td>10.67</td>
<td>5.42***</td>
</tr>
<tr>
<td>Percent of Questions answered</td>
<td>18%</td>
<td>13%</td>
<td>0.05***</td>
</tr>
<tr>
<td>Number of Answers Correct</td>
<td>9.3</td>
<td>6.21</td>
<td>3.09***</td>
</tr>
<tr>
<td>Percent of Answers Correct</td>
<td>57%</td>
<td>56%</td>
<td>0.01*</td>
</tr>
<tr>
<td>Confidence when Answering</td>
<td>3.49</td>
<td>3.63</td>
<td>-0.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Male answering</th>
<th>Female Answering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Group Confidence</td>
<td>3.32</td>
<td>3.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of players</td>
<td>51</td>
<td>15</td>
</tr>
<tr>
<td>Number of captains</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001

Several important factors are lacking. Demographic data other than gender are unavailable for the players, so I cannot control for unobservable factors such as age, education and occupation. The sample of players is not representative, as they overwhelmingly tend to be high-ranking professionals in fields such as law, business and
engineering. The process by which individuals are selected to be on the TV series is unclear, although it is certainly not a random sample. The composition of the teams and the selection of the captain are also not randomized within the sample. The data lacks a key variable, providing no indication of which players proposed the correct answer to a given question during the discussion period. Another concern is the subjectivity inherent in rating concepts such as confidence on a numerical scale. The rating was performed by a singular female rater, and every effort was made to maintain a consistent system of rating confidence and dominance, including a review that reexamined a substantial portion of the data. The results of the review and the original coding of the data were very similar, indicating little variation over time in collection procedures.

Despite these concerns (which provide the motivation for the experimental approach) the analysis of this observational data is worthwhile because the game show shares characteristics with real-world group discussion environments in the workplace. Random selection does not determine the occupants of boardroom seats, nor the selection of decision makers. The fields which possess the most evident leadership gaps are similar to those represented in the sample of players. The summary statistics (Table 1) show that there are far fewer female players and female captains than male players and male captains, indicating not only a leadership gap but lower representation of women on the show to begin with. However, on average, women tend to perform just as well as men in terms of answering questions correctly when chosen. Women are also on average more likely to be dominant and exhibit similar to greater confidence when answering compared to their male counterparts. Yet, on average women are less likely than men to be chosen to answer a question. T-tests reveal both the 3% difference in the average probability of being
dominant (which favors women) and the 6% difference in the probability of being chosen to answer (which favors men) to be statistically significant at the 1% level.

2.2. Regression Model for Observational Data

Players in this game should be able to increase their likelihood of being chosen to answer a question by similar methods that employees might use to increase their chances of being promoted: improving his or her visibility amongst colleagues, and building a reputation for excellent work. Therefore dominant, confident players with good track records for answering correctly should be chosen most often to answer. The data show that on average women match or exceed their male counterparts on each of these measures, which raises the possibility of gender bias in the captains’ choice patterns. I explore this possibility further with regression analysis, using both linear and logit models. With each, I estimate five specifications represented by forms of Equations 1 and 2.

\[
Chosen_{it} = \beta_1 Female_{it} + \beta_2 Dominant_{it} + \beta_3 Male\ Captain_{it} + Z_{it}\gamma + \theta_t + \epsilon_{it} \quad (1)
\]

\[
Chosen_{it} = [Gender \times Interaction]'_{it}\beta + Z_{it}\gamma + \theta_t + \epsilon_{it} \quad (2)
\]

where:
Gender \epsilon \{Female, Male\}
Interaction \epsilon \{Dominant, Captain Gender, Group Confidence\}
    Dominant \epsilon \{Not Dominant, Dominant\}
    Captain Gender \epsilon \{Male, Female\}
    Group Confidence \epsilon \{Low, High\}
i=player and t=question
\[Z_{it}\] is a vector of controls containing a player’s average confidence, cumulative percent of answers correct, and the stakes associated with the question
\[\theta_t\] represents fixed effects for the question timing and year
The dependent variable is an indicator variable reflecting whether a player was chosen by the team captain to answer that particular question. In Specifications 2-4 I include a vector of controls intended to measure individual player and question characteristics (represented by $Z_i$). I use the average of a player’s confidence when answering as a proxy for that player’s overall confidence, and the cumulative percent of questions a player has answered correctly up until that point as a proxy for reputation effects (i.e. how skillful a player appears to his or her teammates). I also include a control for whether or not a question is a high stakes question, meaning that if the player were to answer incorrectly, the team would automatically lose the game. Also included in $Z_i$ are fixed effects for year and how late in the episode a question was played, to control for any time-related variation in captains’ choice patterns. In all specifications, standard errors are clustered at the episode level.

In Specifications 3-5 (represented by Equation 2), I generate specific variables to examine the interactions between gender, dominance in the discussion, gender of the team captain and the confidence of the group during that particular discussion. This allows a more in-depth analysis of which situations are most likely to be affected by gender bias. For example, if we observe a negative coefficient on the $Female$ indicator variable in Specifications 1 and 2 (represented by Equation 1), we can then use Specification 4 see if the probability of being chosen to answer is lower for both women who are dominant in the discussion and women who are not. Conclusions from these specifications may be useful in determining how to eradicate gender bias from group discussions.
2.3. Discussion of Regression Results

The linear OLS and logit models give very similar probability estimates in all specifications. Both models are commonly used to estimate regressions when the outcome variable is dichotomous. The choice between the two depends on the shape of the data. If the probabilities estimated are extremely close to 0 or 1, the linear model is not appropriate, but otherwise the two models tend to give very similar results, as in this case. Both models estimate identical signs and very similar magnitudes for all coefficients, and comparable values of $r^2$ in all specifications.

Result 1: Women are approximately 5% less likely to be chosen to answer a question than their male counterparts (Tables 2 and 3)

Result 1 is present in Specification 1 and remains consistent (and statistically significant at the 1% level) with the addition of the vector of controls in Specification 2. Specification 2 also indicates a strong positive effect of being the dominant player on the probability of being chosen to answer (about 25% with OLS and 19% with the logit model). This effect is broken down in Specification 4, where I examine dominant women, dominant men, and non-dominant women relative to the omitted condition of non-dominant men.

Result 2: Non-dominant women are 6-8% less likely to be chosen to answer than non-dominant men. (Tables 2 and 3)

The gender bias in Result 1 appears to be driven specifically by a bias against women who are less dominant. Being the dominant player essentially equalizes the probabilistic advantage in being chosen to answer for women and men at about 25%. An F-test on the coefficients for dominant and non-dominant women in Specification 4 indicates that the difference is statistically significant at the 1% level (Tables 4 and 5)
Table 2: OLS Estimation Results of Probability of Being Chosen to Answer

<table>
<thead>
<tr>
<th></th>
<th>(I)</th>
<th>(II)</th>
<th>(III)</th>
<th>(IV)</th>
<th>(V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-0.047***</td>
<td>-0.053***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.019)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant</td>
<td>0.255***</td>
<td>0.254***</td>
<td>0.256***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.028)</td>
<td>(0.029)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Captain</td>
<td>0.003</td>
<td>0.004</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.020)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female and Female Captain</td>
<td>0.034</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female and Male Captain</td>
<td>-0.071***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male and Female Captain</td>
<td>-0.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Dominant Female</td>
<td></td>
<td>-0.064***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.019)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant Male</td>
<td></td>
<td>0.235***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.030)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant Female</td>
<td></td>
<td>0.250***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.055)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female in a Low Confidence Group</td>
<td></td>
<td>-0.053**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.025)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female in a High Confidence Group</td>
<td></td>
<td>-0.043*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.022)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male in a High Confidence Group</td>
<td></td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls for average confidence, cumulative percent of answers correct, and stakes</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Fixed Effects for Year and Question Timing</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>3,816</td>
<td>3,382</td>
<td>3,382</td>
<td>3,382</td>
<td>3,382</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.003</td>
<td>0.052</td>
<td>0.054</td>
<td>0.053</td>
<td>0.052</td>
</tr>
<tr>
<td>Omitted Condition in Interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001
Standard errors clustered on episode
Table 3: Logit Estimation Results of Probability of Being Chosen to Answer

<table>
<thead>
<tr>
<th></th>
<th>(I)</th>
<th>(II)</th>
<th>(III)</th>
<th>(IV)</th>
<th>(V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-0.050**</td>
<td>-0.056***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.021)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant</td>
<td>0.194***</td>
<td>0.193***</td>
<td>0.195***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Captain</td>
<td>0.003</td>
<td>0.004</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female and Female Captain</td>
<td></td>
<td></td>
<td>0.030</td>
<td></td>
<td>-0.079***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.045)</td>
<td></td>
<td>(0.025)</td>
</tr>
<tr>
<td>Female and Male Captain</td>
<td></td>
<td></td>
<td>-0.077***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male and Female Captain</td>
<td></td>
<td></td>
<td>-0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Dominant Female</td>
<td></td>
<td></td>
<td>-0.079***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant Male</td>
<td></td>
<td></td>
<td>0.173***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant Female</td>
<td></td>
<td></td>
<td>0.181***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.033)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female in a Low Confidence Group</td>
<td></td>
<td></td>
<td></td>
<td>-0.055**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>Female in a High Confidence Group</td>
<td></td>
<td></td>
<td></td>
<td>-0.046*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>Male in a High Confidence Group</td>
<td></td>
<td></td>
<td></td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.009)</td>
<td></td>
</tr>
<tr>
<td>Controls for average confidence, cumulative percent of answers correct, and stakes</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Fixed Effects for Year and Question Timing</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>3,816</td>
<td>3,382</td>
<td>3,382</td>
<td>3,382</td>
<td>3,382</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.003</td>
<td>0.047</td>
<td>0.049</td>
<td>0.049</td>
<td>0.047</td>
</tr>
<tr>
<td>Omitted Condition in Interaction</td>
<td>Male and Male Captain</td>
<td>Non-Dominant Male</td>
<td>Male in a Low Confidence Group</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001

Standard errors clustered on episode
**Table 4: F-tests on Interaction Coefficients from OLS Estimates (H₀ : β₁ = β₂)**

<table>
<thead>
<tr>
<th></th>
<th>β₁</th>
<th>β₂</th>
<th>F-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female and Male Captain</td>
<td>Female and Female Captain</td>
<td>3.524</td>
<td>0.066*</td>
<td></td>
</tr>
<tr>
<td>Non-Dominant Female</td>
<td>Dominant Female</td>
<td>37.320</td>
<td>0.000***</td>
<td></td>
</tr>
<tr>
<td>Female in a Low Confidence Group</td>
<td>Female in a High Confidence Group</td>
<td>0.144</td>
<td>0.706</td>
<td></td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001

**Table 5: Chi² tests on Interaction Coefficients from Logit Estimates (H₀ : β₁ = β₂)**

<table>
<thead>
<tr>
<th></th>
<th>β₁</th>
<th>β₂</th>
<th>Chi statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female and Male Captain</td>
<td>Female and Female Captain</td>
<td>4.877</td>
<td>0.027**</td>
<td></td>
</tr>
<tr>
<td>Non-Dominant Female</td>
<td>Dominant Female</td>
<td>57.450</td>
<td>0.000***</td>
<td></td>
</tr>
<tr>
<td>Female in a Low Confidence Group</td>
<td>Female in a High Confidence Group</td>
<td>0.094</td>
<td>0.759</td>
<td></td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001

In Specification 3, I look more closely at how the gender of the captain may be affecting choice patterns. This aspect of the analysis is somewhat limited, due to the extremely low number of female captains included in the data. Additionally, one of the two female captains included was on a team with only other men. Therefore when we look separately at women on teams with female captains, men on teams with female captains and women on teams with male captains, relative to men on teams with male captains, it is unsurprising that the coefficients for the first two categories are statistically insignificant. The statistically significant negative effect of about seven percentage points for women on teams with male captains is a reconfirmation of Result 1, since this is by far the most prevalent situation represented in the data.³

Lastly, specification 5 takes into account the confidence of the group during the discussion of the question before the captain chooses the team representative. This specification was included to check for evidence of the phenomenon known colloquially as

³ Due to this imbalance, results from Specification 3 should be interpreted with caution.
the ‘glass cliff’. In this scenario women are put forward for leadership positions in times of crisis, and then are blamed for negative outcomes, reaffirming stereotypes about women in positions of power. If this effect is present in the data, we should see a higher probability of women being chosen to answer when the group is less confident. Women in low confidence groups are in fact about 1 percentage point less likely to be chosen to answer than women in high confidence groups, and are 4-5% less likely to be chosen than men regardless of whether group confidence is high or low. Therefore we can conclude that no evidence of the glass cliff is present in this particular data set.

The most powerful message from these results is the importance of dominance in equalizing the probability of being chosen for men and women. Dominance in any given discussion has a far larger effect than personal confidence or good results in the past (the coefficients of which were insignificant in all specifications). Men can stay relatively quiet and still expect their ideas to be given attention while women enjoy no such luxury. While it is discouraging that women must be the dominant person in the group to eradicate the disadvantage of being female, this analysis does suggest a path forward for women in the meeting room. If women can manage to capture and retain the attention of a boardroom full of their colleagues, they can enjoy the same consideration as their male counterparts.
Chapter 3: Experimental Approach

This laboratory experiment enables me to gauge the impact of a variety of factors that are unobservable in the data gathered from the game show. The experimental design preserves the original elements of “What? Where? When?” that are critical to the investigation of my research question. Specifically, teams of three randomly selected players and one randomly chosen captain discuss the best answer to a question, and one subject is ultimately picked by the captain to answer the question on behalf of the group. The experiment allows me to identify the mechanisms which influence the captain’s decision in greater detail.

First, I am able to measure and control for baseline performance of all subjects in the task (see section 3.1 below). Second, the post-experiment questionnaire allows me to observe a rich set of individual demographic characteristics. Third, I am able to separate the effects of gender per se from the effects of traits that are perceived to be related to gender (such as confidence) by randomly assigning subjects to different treatments that (a) reveal or do not reveal the gender of the group members, and (b) vary the degree to which subjects are able to communicate their confidence in their idea or display dominance of any kind (see section 3.3). The following sections provide detailed explanations of the tasks and treatments used in the experiment.

3.1. The Task and Baseline Ability

The experiment remains centered around a task that a group of four subjects must complete in a specified period of time, followed by a player randomly designated as the captain delegating another player to submit an answer on behalf of the group. I chose the
experimental task such that the best solution can be determined by logical reasoning, but there is still room for disagreement among subjects during the discussion period. To mirror real-world discussion, in which multiple ideas could be applicable but some are better than others, I elected to use questions from *Family Feud.* At the beginning of the experiment, each participant answered four questions individually before any group work took place. For each question, the subject had one minute to submit an answer. Subjects answered these questions solely using their own judgment, and the cumulative score from their answers was used to establish a measure of baseline ability for this particular task. As an example, consider a representative *Family Feud* question and the corresponding table with answers below.

“Name an article of clothing that you can’t wash in the washing machine,”

<table>
<thead>
<tr>
<th>Answers</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoe</td>
<td>29</td>
</tr>
<tr>
<td>Bra</td>
<td>21</td>
</tr>
<tr>
<td>Hat</td>
<td>16</td>
</tr>
<tr>
<td>Coat</td>
<td>13</td>
</tr>
<tr>
<td>Sweater</td>
<td>9</td>
</tr>
<tr>
<td>Suit</td>
<td>7</td>
</tr>
<tr>
<td>Gown</td>
<td>3</td>
</tr>
</tbody>
</table>

The number of points is equal to the number of people who gave that answer on a 100 person survey. The goal for subjects in this experiment was to provide the answer that will earn the most points. Subjects were informed of the relationship between the score and the answers, so that they understood that to receive high scores they must select answers that were the most popular on the survey, not necessarily those which they felt were the most correct or the most inventive. Only answers that received two or more

---

4 Questions were selected from the database at http://familyfeudfriends.arjdesigns.com/
survey responses count for points. If the answer submitted did not appear in the table of answers, the subject received zero points.

3.2. **Group Discussion and Selection of the “Representative”**

The procedure for the group discussion portion of the experiment varied depending on treatment, but several key aspects remained constant across treatments. Each participant first had 15 seconds to read the given question and submit a self-reported confidence level elicited in the following manner:

"On a scale of 0-10, please indicate how confident you feel about your ability to submit a high-scoring answer to this specific question."

Next, the team of four randomly selected members “discussed” the question for two minutes. At the end of the discussion period, each player submitted his or her choice of another group member to potentially serve as the group’s chosen “representative” – the individual who would actually submit his or her answer to the question.\(^5\) Once each individual participant’s choice of representative had been submitted, one of the four group members was randomly selected as “captain,” and his or her choice of representative was implemented. Thus, each player behaved as if he or she was the captain, knowing his or her choice could be implemented. This procedure, known as “the strategy method” provided more key data points regarding the selection of the representative for every question played.

In addition to selecting a preferred representative, each subject answered two follow-up questions:

\(^5\) In order to abstract away from potentially gendered effects of social and risk preferences on the decision to personally answer the question, I restrict the choice of “representative” to be any team member other than the respondent.
(1) “What answer do you expect the chosen representative to submit?”

(2) “On a scale of 1-10, please indicate how confident you feel that your chosen representative will submit the highest scoring answer.”

The first question aimed to tease out a potential gender difference in the perception of whether the representative is likely to go with his or her own answer or to change the answer to somebody else’s submission. The second question gauged the respondent’s overall confidence in his or her chosen representative.

Each session of the experiment contained four rounds of group discussion questions, and team members were randomly re-matched before each question. This allowed for natural variation in the gender composition of each team, and prevented players from establishing reputations within teams that might affect their choices of whom to designate as the team’s representative.

Teams were not able to communicate with each other outside of the discussion period. Each player was assigned a unique ID number at the beginning of the experiment. These ID numbers, and in some treatments, the subjects’ first name or nickname, was displayed in list form on the screen of each player in the group, so that players could track who was contributing which ideas during the discussion period.

---

6 Because this question comes after the submission has already been made, I avoid the answer biasing the choice of representative in the given round. Note however that some learning may occur over rounds, as participants observe actual submissions by their representatives.

7 In this experiment, it is crucial for players to recall the identifiers of other players, in order to be able to make an informed decision for the “representative.”
### 3.3. Treatments

The design includes six experimental treatments which vary the degree of interaction among group members, and change the way in which confidence and dominance can be expressed. The treatments also vary players’ awareness of the genders of their teammates. The various permutations of treatments are presented in Table 6.

**Table 6: Treatment Structure of Experimental Design**

<table>
<thead>
<tr>
<th></th>
<th>GENDER KNOWN</th>
<th>GENDER NOT KNOWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Confidence Signal (Control)</td>
<td>Treatment A: All subjects submit an answer and see all other players’ submissions before selecting a representative. Gender revealed by subject name and roll call.</td>
<td>Treatment B: All subjects submit an answer and see all other players’ submissions before selecting a representative. Gender not known (players identified by ID #).</td>
</tr>
<tr>
<td>Self-Reported Confidence Signal</td>
<td>Treatment C: same as A, but subjects also view each others’ self-reported confidence before selecting a representative. Gender revealed by subject name and roll call.</td>
<td>Treatment C: same as A, but subjects also view each others’ self-reported confidence before selecting a representative. Gender not known (players identified by ID #).</td>
</tr>
<tr>
<td>Self-Reported Confidence Signal and Non-Face-to-Face Communication</td>
<td>Treatment E: same as C, but subjects also participate in written communication via chat room. Subjects view transcript of chat in addition to players’ self-reported confidence and submitted answers. Gender revealed by subject name and roll call.</td>
<td>Treatment E: same as C, but subjects also participate in written communication via chat room. Subjects view transcript of chat in addition to players’ self-reported confidence and submitted answers. Gender not known (players identified by ID #).</td>
</tr>
</tbody>
</table>

Changing how subjects communicate and express confidence and dominance was intended to help tease out the root cause of any gender gap in the probability of a subject being chosen to answer, such as we saw in the observational data analysis. There are two possible explanations for any observed discrimination against female players. It may be taste-based, with captains simply preferring men, in which case we would expect to see no gap in treatments where gender is not revealed and women being disproportionately less
likely to be picked in treatments with gender revealed across all levels of communication. However, the ways in which women express confidence and dominance could be influencing outcomes rather than the effect being directly due to gender. In this case we would expect to see that women are chosen equally in treatments at some levels of communication, but are less likely to be chosen at other levels of communication, regardless of whether or not gender is revealed.

The first treatment dimension is the ability of subjects to observe the gender of others. In all treatment cells under “gender known” in Table 6, subjects were implicitly aware of the gender of their teammates. In treatments A, C, and E, gender was revealed in two ways. First, players participated in a vocal roll-call at the beginning of each discussion period. Second, subjects viewed the preferred first name or nickname of the others in their group. In all treatment cells under “gender not known” in Table 6 subjects were only identified by an ID number.

The second treatment dimension is the extent to which subjects can freely communicate during discussion. The most basic treatments (A and B) feature the most controlled manner of interaction. All subjects had the same amount of time (one minute) to choose their preferred answer from a list of options. All players saw the submissions by the

---

8 In the treatments where gender is supposed to be observable, it is important that subjects recall this part of the others’ identity. Therefore, I believe it is not enough to inform the subjects of the gender of others only at the beginning of the discussion period via roll call as in Bordalo et al. (2016). It is also inappropriate and unrealistic to explicitly state the gender of the other participants, because by making the gender extremely salient, we would be risking revealing the purpose of the study directly and therefore limiting the possibility of finding any bias. Thus, in the known gender treatments, I choose to subtly hint at the subject’s gender with their preferred name. Note that subjects will not be precluded from reporting a “fake name” in case they prefer to remain anonymous. Finally, I acknowledge that gender-neutral or ambiguous names will necessarily introduce measurement error into my estimates. Therefore, I will attempt to mitigate this concern by also conducting the roll call of subjects at the beginning of each discussion period.
others, and from that information made the decision about whom they would like to 
nominate to answer on behalf of the group. Treatments C and D are identical to Treatments 
A and B, with the addition of eliciting a numeric measure of players’ confidence in their 
ability to answer the question. This measure was then displayed to the other subjects in the 
group and will provided additional information that may have informed these subjects’ 
choice of representative (see section 3.2).

In Treatments E and F, the discussion was free-form. In both treatments, subjects 
typed messages to the other member of their group in a chat room. In Treatment E, gender 
was revealed by the same method as in Treatments A and C. Every contribution submitted 
to the chat room was accompanied by the subject’s name in Treatment E or the subject’s ID 
number in Treatment F. After the discussion, subjects received a transcript of the 
discussion with the corresponding ID numbers or names (depending on treatment) of the 
players, indicating who had made which contributions. This transcript, combined with the 
self-reported confidence levels and each player’s proposed answer was available to the 
subjects as they made the choice of representative for the group. Discussions in these 
treatments were reviewed by the experimenters to measure dominance.

The structure of these two treatments allows me to examine a different variety of 
dominance which is not based purely on the quantity of contributions, but also takes into 
account the timing and quality. The experiment design has players first participate in the 
chat, then individually submit their preferred answers, and then select a representative. In 
about 83% of groups observed when the experiment was conducted, at least 3 out of the 4 
players in a group came to a consensus about their preferred answer during the chat and 
all submitted it individually. In none of these cases did the chosen representative
subsequently nominated to answer on behalf of the group deviate from this agreed-upon answer. By examining the chat transcripts, it was possible to construct a variable indicating which player within a group first proposed the final answer submitted by the group's chosen representative. This allows me to examine the effect of proposing a popular answer quickly on whether or not an individual is nominated by his or her teammates to answer on behalf of the group. This variable represents the effect of dominance via being proactive, and presenting high quality ideas that are popular with the rest of the group, whereas having the greatest percentage of chat entries represents dominance by being vocal and prolific.

In addition to the six treatments listed in the table, two other treatments were considered: one where subjects would communicate over a headphone set and another where we would allow face-to-face communication. However, I decided that they did not ultimately fit into the framework of this experiment. Firstly, it is impossible to design a treatment that replicates either oral or face-to-face discussion but conceals gender, which would prevent me from establishing a clear comparison group. Secondly, both treatments introduce confounds that would be difficult, if not impossible, to fully control. In the oral discussion treatment, the tone of voice and pitch can affect how persuasive one can be in their arguments. Face-to-face discussion further introduces potential bias stemming from the ‘beauty premium’ which is difficult (though not impossible) to control for in a laboratory setting. These treatments would enhance the external validity of my experiment and are therefore potentially fruitful to pursue in a follow-up study, as they reflect real-world discussions more closely than a chat room. However, the concerns listed above have
caused me to reject them for this particular experiment and focus on the 2x3 design shown in Table 6.

### 3.4. Payoffs

Subjects accrued points for performance in each part of the experiment. Points in the individual task were equal to the cumulative score over four Family Feud questions, answered individually ($Points_i = \sum_{r=1}^{4} Score_{ir}$, where $r$ refers to question number or round). Points in the group discussion part of the experiment were equal to the total number points earned by the player based on his or her group performance ($Points_g = \sum_{r=1}^{4} Score_{Gr}$). Whether the player gets points based on individual performance or on the group round performance will be randomly determined.  

Overall, the experiment was incentivized by means of the following payoff function:

$$Payoff = \$10 \text{ show-up fee} + \$0.12 \times Points + \$0.50 \text{ for each time that player was the team's chosen representative}$$

where $Points$ refer to points from either the individual or the group part of the experiment.

Subjects were informed of the structure of the payoff function at the beginning of the experiment, and it is therefore designed to encourage quality participation from all players.

---

9 I am aware of the potential problems that due to the wealth effects associated with the accumulation of points over multiple rounds. A way to address this concern would be to randomly select only one of the 8 total rounds of questions to receive payment. However, due to the condition that subjects must receive at least $20 for participation in this experiment, I elected to base payments on the cumulative score from one of my two rounds of play. This is because of the potentially wide variation in points possible for a given question (which may be as low as zero).

10 The laboratory utilized for the experiment requires a minimum payment of $20, so any subjects who made less than $20 under the payoff function received an additional completion fee.
3.5. Procedures

The experiment was conducted at the Computer Lab for Experimental Research (CLER) at Harvard Business School (HBS) and took place in March 2017. After signing the informed consent form, participants were seated at individual computer terminals. Subjects received written, oral, and on-screen instructions programmed using the standard zTree software package (Fischbacher 2007). Participants were encouraged to ask questions in private if they did not understand these instructions, but communication between subjects was disallowed other than when instructed. Subjects only received the instructions relevant to the immediate part of the experiment (Part 1 or 2). At the end of the experiment, subjects were informed about their performance and payment and filled out a post-experiment questionnaire. Subjects were paid in cash and in private by the experimenters.

---

11 CLER recruits subjects via an online registration procedure. Subjects first register for the CLER subject pool. Then, they sign up for studies of their choosing. Most subjects are students at Harvard University (undergraduates and graduates), although students from other Boston-area universities, such as MIT and Boston University, also participate. At any point, a subject can remove him or herself from the study for any reason.
Chapter 4: Analysis of Experimental Data

The data collected in the lab for the purposes of this paper consists of two sessions each of Treatments E and F (See Table 6 for summary). Hereafter I will refer to E as the gender treatment (GT) and F as the no gender treatment (NGT). A smaller pilot session of the GT is also included in the summary statistics and some regression specifications. The pilot session lacked the information from the post-experiment survey, so the data was dropped in the specifications which control for demographics. During the pilot, one of the questions in the group portion of the experiment did not elicit substantive discussion from subjects. In later sessions, this question was swapped with a question in the individual portion, so the pilot data is not directly comparable to that from later sessions of the GT. Overall 92 subjects participated in the experiment (including the pilot), 44 in the GT and 48 in the NGT. The data is organized such that for each group discussion question, there is one observation for each player, resulting in 368 total observations.

4.1. Overview of Experimental Data

Summary statistics for the lab data reveal trends similar to those found in the observational data with the added dimension of comparing across treatments to see how knowledge of gender changes the mean outcomes. Table 7 presents summary statistics, both of the sample’s demographics and the experimental outcomes, broken down by treatment and gender.
### Table 7: Demographic Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Gender Treatment</th>
<th>No Gender Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of Subjects</td>
<td>44</td>
<td>48</td>
</tr>
<tr>
<td>Average Age</td>
<td>23.37</td>
<td>22.33</td>
</tr>
<tr>
<td>Average Undergraduate GPA</td>
<td>3.57</td>
<td>3.49</td>
</tr>
<tr>
<td>Proportion of Sample by Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.43</td>
<td>0.56</td>
</tr>
<tr>
<td>Male</td>
<td>0.57</td>
<td>0.44</td>
</tr>
<tr>
<td>Proportion of Sample By Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.31</td>
<td>0.40</td>
</tr>
<tr>
<td>Asian</td>
<td>0.29</td>
<td>0.33</td>
</tr>
<tr>
<td>Black</td>
<td>0.20</td>
<td>0.08</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>Other</td>
<td>0.11</td>
<td>0.10</td>
</tr>
</tbody>
</table>

### Table 8: Lab Data Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Gender Treatment</th>
<th>No Gender Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Average Score in Individual Task</td>
<td>100.76</td>
<td>90.32</td>
</tr>
<tr>
<td>Average Confidence</td>
<td>7.51</td>
<td>7.29</td>
</tr>
<tr>
<td>Average Percent of Chat Entries</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td>Average Score of Proposed Answer</td>
<td>9.96</td>
<td>10.72</td>
</tr>
<tr>
<td>Probability of Dominating a Chat</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td>Probability of Suggesting Final Answer</td>
<td>0.25</td>
<td>0.21</td>
</tr>
<tr>
<td>Probability of Suggesting Better Answer</td>
<td>0.25</td>
<td>0.22</td>
</tr>
<tr>
<td>Probability of Being Chosen as Representative</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>Average Number of Votes per Round</td>
<td>0.97</td>
<td>1.01</td>
</tr>
<tr>
<td>Total Number of Subjects</td>
<td>25</td>
<td>19</td>
</tr>
</tbody>
</table>

*p<0.10, **p<0.05, ***p<0.001
We see in Table 8 that in neither treatment are men more likely to dominate the chat portion of the experiment, where dominance is defined as contributing the highest percentage of chat entries within a group in a given round. Similarly, in neither treatment are men significantly more likely achieve ‘quality’ dominance by being the first to contribute the group’s final answer. We have no reason to expect revealing gender to affect a subject’s score in the individual portion of the experiment as the roll call and the submission of names to indicate gender takes place after the individual task is complete. It is possible that the roll call works as a prime activating the stereotype threat, which might explain the gender difference in self-assessed confidence (Gneezy et al., 2003). However, in that case we would observe the difference in the GT rather than the NGT.\textsuperscript{12}

Just as in the observational data, the summary statistics for the experimental approach do not indicate that women are on average less confident, less dominant or less capable than their male counterparts. Indeed, we observe that when gender is concealed, on average women have a 14% greater chance of being selected as the chosen representative by at least one of their teammates than men. Yet men and women are chosen approximately equally when gender is revealed. Since on average we observe no other significant gender differences by treatment, we can attribute this change solely to revealing the genders of the players, thus suggesting some form of gender-based discrimination.\textsuperscript{13}

\textsuperscript{12} Therefore I attribute the differences in individual task score and self-assessed confidence between men and women to Type 1 error due to the small size of my sample.

\textsuperscript{13} This conclusion comes from examining the data indicating whether or not a player was chosen in a given round by at least one other player. We see a similar pattern in the total number of votes a player receives, however this is not statistically significant and the magnitude of the difference is much smaller. This reflects a pattern in the underlying data that women are much less likely to receive 1 vote, but about equally likely to receive 2-3 votes.
4.2. Regression Model for Experimental Data

Since the experiment is structured slightly differently than the original TV show from which the observational data was collected, I use a similar but slightly modified linear probability regression model to examine the data more thoroughly. The four specifications are represented by the following equations:

\[
Chosen_{it} = [Gender \times Treatment]^{'}_{it} \beta + \beta_t Chat Dominance_{it} +
\beta_2 First to Propose Final Answer_{it} + \beta_3 Confidence_{it} + Z_{it} \gamma + \theta_t + \epsilon_{it} \tag{3}
\]

\[
Chosen_{it} = [Gender \times Interaction \times Treatment]^{'}_{it} \beta + Z_{it} \gamma + \theta_t + \epsilon_{it} \tag{4}
\]

where:
Gender ∈ {Female, Male}
Treatment ∈ {GT, NGT}
Interaction ∈ {Chat Dominance, Final Answer Proposal}
    Chat Dominance ∈ {Not Dominant, Dominant}
    Final Answer Proposal ∈ {First, Not First}
i=player and t=question
\(Z_{it}\) is a vector of controls containing a player’s race, age, undergraduate GPA, prior familiarity with the task, total score in the individual task.
\(\theta_t\) represents fixed effects for the question timing

In all, the dependent variable is a indicator variable reflecting whether in any given round, the individual was selected by one or more others in his or her group as their preferred representative for the group. The key independent variables in Specifications 1 and 2 (represented by Equation 3) are a set of interaction variables that indicate the gender of the player, and whether or not the player is in a treatment where gender is known. In the absence of discrimination, we ought to see comparable coefficients for women in the GT and women in the NGT after controlling for other factors. In the presence of gender bias, the coefficient for women in the GT will be lower than in the NGT.

In the Specification 1, I also include controls for a players’ individual self-reported confidence levels in that round (coded from 1-10, 10 being the highest), their total scores in
the individual portion of the experiment as a proxy for aptitude for the task, whether or not they were dominant in the chat, and whether they were first to suggest the group’s final answer. In Specification 2, I use the continuous value of a player's percentage of the group’s chat entries instead of dominance in the chat. I also add a vector of demographic controls including undergraduate GPA, Race and age, as well as a control for how familiar the subject is with *Family Feud*, which could have a positive effect on performance.

Specifications 3 and 4 (represented by Equation 4) utilize three-way interactions to separately examine the two different types of dominance identified in section 3.3. First I compare women who are not dominant in the chat with dominant women in each treatment, to see how the effect of revealing gender might impact these types of women differently. I then follow the same method for women who are not first to propose the group’s answer, compared to those who are. This breakdown parallels my analysis of the combined effect of gender and dominance in the observational data, but separates out ‘quality’ and ‘quantity’ dominance instead of conflating them in a single variable.
4.3. Discussion of Experimental Data Regression Results

Table 9: OLS Estimates of the Probability Being Chosen as Representative

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male in GT</td>
<td>0.068</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Female in GT</td>
<td>0.090</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.088)</td>
</tr>
<tr>
<td>Female in NGT</td>
<td>0.170**</td>
<td>0.158**</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>Confidence</td>
<td>0.030**</td>
<td>0.047***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Submitted Final Answer First</td>
<td>0.242***</td>
<td>0.188***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Dominant in Chat</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td></td>
</tr>
<tr>
<td>Percentage of Chat Entries</td>
<td></td>
<td>0.310</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.287)</td>
</tr>
<tr>
<td>Controls for age, Race, GPA, prior</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>familiarity with task and total score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in individual task</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effects for question timing</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>336</td>
<td>276</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.091</td>
<td>0.158</td>
</tr>
<tr>
<td>Omitted Condition in Interaction</td>
<td>Male in</td>
<td>Male in</td>
</tr>
<tr>
<td></td>
<td>the NGT</td>
<td>the NGT</td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001

Robust Standard Errors
Result 3: Women have a 17% greater probability of being chosen than their male counterparts when gender is not known, but this advantage almost disappears when gender is known. (Table 9)

Specifications 1 and 2 (Table 9) both show a statistically significant higher probability of being chosen for women in the NGT when compared to men in the NGT. However in the GT, this probabilistic advantage drops from 15-17% more likely to be chosen down to a statistically insignificant value as low as 2% in Specification 2. An F-test on the difference of these coefficients yields a p-value of 0.112 (Table 11). For a slightly larger sample, it is not unreasonable to imagine this value becoming significant. Also of note is the statistically significant coefficient on confidence, indicating a 3-5% increase in the probability of being chosen associated with an increase of one level in self-reported confidence. Neither the coefficient of dominance in the chat, nor that of the percentage of chat entries is statistically significant, although both are positive. In contrast, the effect of being the first player to propose the group’s final answer is both large (24% in Specification 1 and 19% in Specification 2) and is significant at the 1% level. Thus it appears that ‘quality’ rather than ‘quantity’ dominance is foremost in determining who is chosen to answer.
Table 10: OLS Estimates of the Probability Being Chosen as Representative Continued

<table>
<thead>
<tr>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-Dominant Male in GT</strong></td>
<td>0.091 (0.106)</td>
</tr>
<tr>
<td>Dominant Male in GT</td>
<td>0.170 (0.113)</td>
</tr>
<tr>
<td>Dominant Male in NGT</td>
<td>0.048 (0.115)</td>
</tr>
<tr>
<td>Dominant Female in GT</td>
<td>-0.047 (0.150)</td>
</tr>
<tr>
<td>Non-Dominant in GT</td>
<td>0.073 (0.105)</td>
</tr>
<tr>
<td>Dominant Female in NGT</td>
<td>0.183* (0.105)</td>
</tr>
<tr>
<td>Non-Dominant Female in NGT</td>
<td>0.153 (0.097)</td>
</tr>
<tr>
<td>Confidence</td>
<td>0.050*** (0.016)</td>
</tr>
<tr>
<td>Total Score in Individual Task</td>
<td>-0.000 (0.001)</td>
</tr>
<tr>
<td>Submitted Final Answer First</td>
<td>0.189*** (0.053)</td>
</tr>
<tr>
<td>Dominant in Chat</td>
<td></td>
</tr>
<tr>
<td>Controls for Race, GPA, Age, Round and prior familiarity with task</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>276</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.159</td>
</tr>
<tr>
<td>Omitted Condition in Interaction</td>
<td></td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001
Robust standard errors
Table 11: F-tests on Interaction Coefficients (H₀ : β₁ = β₂)

<table>
<thead>
<tr>
<th></th>
<th>β₁</th>
<th>β₂</th>
<th>F-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female in GT</td>
<td>Female in NGT</td>
<td>2.412</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td>Non-Dominant Female in GT</td>
<td>Non-Dominant Female in NGT</td>
<td>2.384</td>
<td>0.124</td>
<td></td>
</tr>
<tr>
<td>Non-First Proposer Female in GT</td>
<td>Non-First Proposer Female in NGT</td>
<td>4.157**</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>First Proposer Female in GT</td>
<td>First Proposer Female in NGT</td>
<td>3.184*</td>
<td>0.076</td>
<td></td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.001

Result 4: ‘Quantity’ dominant women appear to be penalized for this type of dominance when gender is known. (Table 10)

The 3-way interactions in Specifications 3 and 4 also point towards a pattern of gender bias. Both women who are dominant in the chat and those who are not enjoy an approximately 15-18% advantage over non-dominant men in the NGT (although this value is only significant for the dominant women). Dominant women in the GT are actually less likely to be chosen than non-dominant men in NGT. Although F-test on the coefficients for dominant women in the GT and dominant women in the NGT shows the difference of 24 percentage points not to be statistically significant, the p-value is again fairly low at 0.124 (Table 11).

Result 5: No gap in probability exists for ‘quality dominant’ men and women. Gender bias is experienced primarily by non-quality dominant women. (Table 10)

The results from the 3-way interaction of being the first to propose the answer, gender, and treatment are more concrete, which is consistent with the larger and more
significant coefficient on this ‘quality’ dominance measure in Specifications 1 and 2.

Relative to men who do not propose the final answer first when in the NGT, women in the NGT who do propose the final answer first are 26% more likely to be chosen as the representative. Similarly, women who are not first proposers have 17% greater probability of being chosen in the NGT. However, relative to these same men in the NGT, women who do propose the final answer first in the GT are still more likely to be chosen, but women who do not are less likely to be chosen (although this value is not statistically significant).

An F-test indicates the difference between the coefficients for non-first proposer women in the NGT and those in the GT to be statistically significant (Table 11). This pattern resembles the results from the observational data which indicated that gender bias is experienced primarily by women who are not dominant.
Figure 1 plots the coefficients for the indicated categories of women in Specifications 2-4 relative to the respective omitted categories, with bars indicating the standard error for each coefficient. This plot indicates a consistent relationship between gender being revealed, and a decreased probability of women being chosen as the representative. The trend is strongest for ‘quantity’ dominant women and ‘quality’ non-dominant women, which is in accord with research showing that women tend to be penalized for verbal assertiveness, but less verbal dominance is viewed more favorably (Williams, 2016). Being dominant in the chat is comparable to being assertive (or aggressive) whereas contributing a good idea early is likely not perceived this way.

---

14 In specification 1, men with gender not known, in specification 2 non-dominant men with gender not known, and in specification 3 non-first proposer men with gender not known.
Chapter 5: Conclusions and Further Research

Both the results from the observational data and the experimental data point to a similar conclusion: women who are comparable to their male counterparts in ability, dominance and confidence experience bias in group discussion forums. This trend in the workplace could result in adverse outcomes for women, simply because their contributions are not being appropriately valued and reported to those in the position to make decisions about promotions and raises. This could also harm firms in general, if good ideas are being discounted or ignored simply because of the gender of those who propose them.

In this study, due to time and budget constraints, I was not able to identify the source of the discrimination as strictly statistical or taste-based. Further research will include conducting sessions of all of the treatments described in Table 6. Additionally, two components of the experimental design will be changed for the purpose of getting better measures of each subject’s preferences for the group’s representative. Instead of simply choosing one person as representative, subjects will be asked to rank all members of their group, which will provide a much more detailed measure of the group’s collective opinion of an individual. We will also ask subjects for their best guesses about the answers to the questions before they discuss with their teammates, and then allow them to update this guess afterwards. This will allow us to observe their best idea prior to the chat, which can be used as a proxy for how they would have answered individually.

My results suggest that in order to avoid a bad outcome relative to a reality in which gender is not observable, women must be ‘quality’ dominant. Unlike men, women do not seem to benefit simply from making their voices heard often and loudly. Additionally, women cannot afford to not contribute, whereas men are able be less inventive and still
enjoy a higher probability of having their ideas considered. This points to a potentially useful strategy for women whose ideas are being ignored during group discussions: say it early and say it exclusively. To have an idea considered at its full value, women should choose their best idea and propose it early, and then not present too many others, which puts them at risk of being viewed unfavorably as ‘too’ assertive and aggressive. This in no way eliminates the source of the bias, but until we as a society can change the inherent biases towards women in the workplace, it may prove a useful mechanism to circumvent the worst effects of gender-based discrimination in group discussions.
Appendix 1: Experiment Instructions for all Treatments Detailed in Chapter 3

Part 1: Individual Task (All Treatments)

**DECISION-MAKING EXPERIMENT INSTRUCTIONS**

Note that payments may be adjusted slightly to assure that we meet minimal payment guidelines at CLER depending on results of pilot sessions

*Welcome and general instructions*

(To be printed and distributed to each station and read aloud by the experimenter at the beginning of the session. Instructions will also appear on screen.)

Thank you for coming today. At this time, please turn off and/or put away any electronic devices/phones/other things you may have brought with you. From this point forward, we ask you to please focus on the tasks of the experiment. You are not permitted to browse the internet at any time during this experiment. You may not talk to other participants during the experiment. Anyone caught violating these rules will be dismissed without pay. If you have questions at any point, you should raise your hand and one of us will come over to assist you.

Whenever you are asked to enter text by typing on the computer keyboard, please be mindful of misspellings and typographical errors.

This experiment has 2 parts, followed by a post-experiment questionnaire. In each part, you have the opportunity to earn points. One part of this experiment will be randomly chosen to determine your payment. At the end of the session, you will be told which part has been randomly chosen to determine your payment. This payment will be added to your $10 show-up fee. You may also qualify for a completion fee at the end of the experiment.

You will be paid in cash and dismissed at the conclusion of the experiment, when all subjects in the laboratory have completed the experiment. This will take no longer than 60 minutes.
We will begin by measuring your individual performance on the task that you will later perform in groups of four. This task consists of answering several open-ended questions. Each question is a survey question that may have multiple correct answers.

For example:

“Name a word a judge might yell out during a tennis match”

<table>
<thead>
<tr>
<th>Answers</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fault</td>
<td>25</td>
</tr>
<tr>
<td>Foul</td>
<td>17</td>
</tr>
<tr>
<td>Love</td>
<td>14</td>
</tr>
<tr>
<td>Out</td>
<td>10</td>
</tr>
<tr>
<td>Order</td>
<td>6</td>
</tr>
<tr>
<td>Net</td>
<td>4</td>
</tr>
<tr>
<td>Point</td>
<td>3</td>
</tr>
</tbody>
</table>

- Each answer is worth a different number of points, equal to the number of people who gave that particular answer in a 100 person survey. The goal is to submit the highest scoring (i.e. most popular) answer.

- You will be given four questions to answer. For each you will have 30 seconds after the question is displayed on your screen to submit your answer. You may only submit one answer.

- After you submit your answer, the score breakdown will be displayed on your screen.

For example, if you answered “love”, you would see the following chart and earn 14 points:

<table>
<thead>
<tr>
<th>Answers</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fault</td>
<td>25</td>
</tr>
<tr>
<td>Foul</td>
<td>17</td>
</tr>
<tr>
<td>Love</td>
<td>14</td>
</tr>
<tr>
<td>Out</td>
<td>10</td>
</tr>
<tr>
<td>Order</td>
<td>6</td>
</tr>
<tr>
<td>Net</td>
<td>4</td>
</tr>
<tr>
<td>Point</td>
<td>3</td>
</tr>
</tbody>
</table>

- Only answers that have received two or more survey responses count for points. If your answer does not appear in the table of answers, you will receive zero points.
• If you misspell your answer (for example, if you typed “lofe” or any other spelling), you will receive zero points. You may capitalize the first letter, but none of the other letters. Please, be careful in typing your submission!

• Your score in this first task will be the sum of the number of points you accrue on each question, and may affect your final payoff at the end of the experiment.

If there are no questions we will begin the individual portion of the experiment now.

**Part 2: Group Task (Treatment A)**

In this part of the experiment, you will participate in 4 rounds of decision-making as a member of a group of 4 participants. In each round, individual participants will be re-matched to form new groups.

Before you begin your group interaction, we will ask you to submit your first name on the computer. You may use your actual first name, a nickname, or any other preferred name. If there is more than one subject in your group with the same name, we will randomly assign a unique numerical value to accompany the name. For example, if there are two individuals named Alex in your group, we will distinguish between them as Alex1 and Alex2. To make sure you have completed this step, we will also conduct a verbal roll call of all group members.

In each round, a question of the form identical to that used in Part 1 of the experiment will appear on each player’s computer screen. Each player will have 15 seconds to read the question.

Next, all players will have 30 seconds to submit their best guess of the highest scoring answer to the survey question.

After all players have submitted their answers, you will be shown a table of player names and the answers they have proposed, and you will be asked the following question:

Please select one player to act as the chosen representative for the group. If selected, this chosen representative will submit one final answer on behalf of the entire group. You may not select yourself.

After specifying your preferred chosen representative, you will answer another question:

What answer do you expect your chosen representative to submit on behalf of the group if they are chosen to fulfill this role?
Once all the players have answered the above question, the **group leader** will be randomly selected. Each player has an equal chance of being designated the group leader.

The group leader’s chosen representative will then submit one final answer to the question on behalf of the entire group. The chosen representative may submit an answer of his or her choosing.

After the chosen representative has submitted an answer, the group’s score breakdown will be displayed on each group member’s screen.

- In each round, all members of the group receive the points equal to the score of the answer submitted by the chosen representative.
- Recall that only answers that have received two or more survey responses count to obtain points. If the chosen representative’s answer does not appear in the table of answers, you will receive zero points.
- Being selected as the group’s chosen representative earns additional points.

The experiment will end with a post-experiment questionnaire. Upon completion of the questionnaire, you will receive your cash payment.

**Part 2: Group Task (Treatment B)**

In this part of the experiment, you will participate in 4 rounds of decision-making as a member of a group of 4 participants. In each round, individual participants will be re-matched to form new groups.

In each round, a question of the form identical to that used in Part 1 of the experiment will appear on each player’s computer screen. Each player will have 15 seconds to read the question.

Next, all players will have 30 seconds to submit their best guess of the highest scoring answer to the survey question.

After all players have submitted their answers, you will be shown a table of player ID numbers and the answers they have proposed, and you will be asked the following question:

Please select one player to act as the **chosen representative** for the group. If selected, this chosen representative will submit one final answer on behalf of the entire group. You **may not** select yourself.

After specifying your preferred chosen representative, you will answer another question:
What answer do you expect your chosen representative to submit on behalf of the group if they are chosen to fulfill this role?

Once all the players have answered the above question, the **group leader** will be randomly selected. Each player has an equal chance of being designated the group leader.

The group leader’s chosen representative will then submit one final answer to the question on behalf of the entire group. The chosen representative may submit an answer of his or her choosing.

After the chosen representative has submitted an answer, the group’s score breakdown will be displayed on each group member’s screen.

- In each round, all members of the group receive the points equal to the score of the answer submitted by the chosen representative.
- Recall that only answers that have received two or more survey responses count obtain points. If the chosen representative’s answer does not appear in the table of answers, you will receive zero points.
- Being selected as the group’s chosen representative earns additional points.

The experiment will end with a post-experiment questionnaire. Upon completion of the questionnaire, you will receive your cash payment.

**Part 2: Group Task (Treatment C)**

In this part of the experiment, you will participate in 4 rounds of decision-making as a member of a group of 4 participants. In each round, individual participants will be re-matched to form new groups.

Before you begin your group interaction, we will ask you to submit your first name on the computer. You may use your actual first name, a nickname, or any other preferred name. If there is more than one subject in your group with the same name, we will randomly assign a unique numerical value to accompany the name. For example, if there are two individuals named Alex in your group, we will distinguish between them as Alex1 and Alex2. To make sure you have completed this step, we will also conduct a verbal roll call of all group members.

In each round, a question of the form identical to that used in Part 1 of the experiment will appear on each player’s computer screen. Each player will have 15 seconds to read the question and answer the following:

   On a scale of 1-10, please indicate how confident you feel about your ability to submit a high-scoring answer to this specific question?
After all players have submitted their answers, you will be shown a table of player names, the answers they have proposed, and their self-reported confidence in that answer. You will be asked the following questions:

1. Please select one player to act as the **chosen representative** for the group. If selected, this chosen representative will submit one final answer on behalf of the entire group. You **may not** select yourself.
2. On a scale of 1-10, please indicate how confident you feel that your chosen representative will submit the highest scoring answer.

After specifying your preferred chosen representative, you will answer another question:

What answer do you **expect** your chosen representative to submit on behalf of the group if they are chosen to fulfill this role?

Once all the players have answered the above question, the **group leader** will be randomly selected. Each player has an equal chance of being designated the group leader.

The group leader's chosen representative will then submit one final answer to the question on behalf of the entire group. The chosen representative may submit an answer of his or her choosing.

After the chosen representative has submitted an answer, the group's score breakdown will be displayed on each group member's screen.

- In each round, all members of the group receive the points equal to the score of the answer submitted by the chosen representative.
- Recall that only answers that have received two or more survey responses count obtain points. If the chosen representative's answer does not appear in the table of answers, you will receive zero points.
- Being selected as the group’s chosen representative earns additional points.

The experiment will end with a post-experiment questionnaire. Upon completion of the questionnaire, you will receive your cash payment.

**Part 2: Group Task (Treatment D)**

In this part of the experiment, you will participate in 4 rounds of decision-making as a member of a group of 4 participants. In each round, individual participants will be re-matched to form new groups.
In each round, a question of the form identical to that used in Part 1 of the experiment will appear on each player’s computer screen. Each player will have 15 seconds to read the question and answer the following:

On a scale of 1-10, please indicate how confident you feel about your ability to submit a high-scoring answer to this specific question?

After all players have submitted their answers, you will be shown a table of player ID numbers, the answers they have proposed, and their self-reported confidence in that answer. You will be asked the following questions:

1. Please select one player to act as the **chosen representative** for the group. If selected, this chosen representative will submit one final answer on behalf of the entire group. You **may not** select yourself.
2. On a scale of 1-10, please indicate how confident you feel that your chosen representative will submit the highest scoring answer.

After specifying your preferred chosen representative, you will answer another question:

What answer do you **expect** your chosen representative to submit on behalf of the group if they are chosen to fulfill this role?

Once all the players have answered the above question, the **group leader** will be randomly selected. Each player has an equal chance of being designated the group leader.

The group leader’s chosen representative will then submit one final answer to the question on behalf of the entire group. The chosen representative may submit an answer of his or her choosing.

After the chosen representative has submitted an answer, the group’s score breakdown will be displayed on each group member’s screen.

- In each round, all members of the group receive the points equal to the score of the answer submitted by the chosen representative.
- Recall that only answers that have received two or more survey responses count obtain points. If the chosen representative’s answer does not appear in the table of answers, you will receive zero points.
- Being selected as the group’s chosen representative earns additional points.

The experiment will end with a post-experiment questionnaire. Upon completion of the questionnaire, you will receive your cash payment.
**Part 2: Group Task (Treatment E)**

In this part of the experiment, you will participate in 4 rounds of decision-making as a member of a group of 4 participants. In each round, individual participants will be re-matched to form new groups.

Before you begin your group interaction, we will ask you to submit your first name on the computer. You may use your actual first name, a nickname, or any other preferred name. If there is more than one subject in your group with the same name, we will randomly assign a unique numerical value to accompany the name. For example, if there are two individuals named Alex in your group, we will distinguish between them as Alex1 and Alex2. To make sure you have completed this step, we will also conduct a verbal roll call of all group members.

In each round, a question of the form identical to that used in Part 1 of the experiment will appear on each player's computer screen. Each player will have 15 seconds to read the question and answer the following:

> On a scale of 1-10, please indicate how confident you feel about your ability to submit a high-scoring answer to this specific question?

After all players have submitted their answers, you will be able to communicate with the other members of your group to discuss potential answers. The discussion will take place via a chat room, and the discussion period will last 1 minute.

Once the discussion period has concluded, you will be able to view a transcript of the discussion, with players’ names indicating their responses, as well as a table of player names, their proposed answer and their self-reported confidence in their answer. You will then be asked the following questions:

1. Please select one player to act as the **chosen representative** for the group. If selected, this chosen representative will submit one final answer on behalf of the entire group. **You may not** select yourself.
2. On a scale of 1-10, please indicate how confident you feel that your chosen representative will submit the highest scoring answer.

After specifying your preferred chosen representative, you will answer another question:

> What answer do you expect your chosen representative to submit on behalf of the group if they are chosen to fulfill this role?

Once all the players have answered the above question, the **group leader** will be randomly selected. Each player has an equal chance of being designated the group leader.
The group leader’s chosen representative will then submit one final answer to the question on behalf of the entire group. The chosen representative may submit an answer of his or her choosing.

After the chosen representative has submitted an answer, the group’s score breakdown will be displayed on each group member’s screen.

- In each round, all members of the group receive the points equal to the score of the answer submitted by the chosen representative.
- Recall that only answers that have received two or more survey responses count to obtain points. If the chosen representative’s answer does not appear in the table of answers, you will receive zero points.
- Being selected as the group’s chosen representative earns additional points.

The experiment will end with a post-experiment questionnaire. Upon completion of the questionnaire, you will receive your cash payment.

**Part 2: Group Task (Treatment F)**

In this part of the experiment, you will participate in 4 rounds of decision-making as a member of a group of 4 participants. In each round, individual participants will be re-matched to form new groups.

In each round, a question of the form identical to that used in Part 1 of the experiment will appear on each player’s computer screen. Each player will have 15 seconds to read the question and answer the following:

> On a scale of 1-10, please indicate how confident you feel about your ability to submit a high-scoring answer to this specific question?

After all players have submitted their answers, you will be able to communicate with the other members of your group to discuss potential answers. The discussion will take place via a chat room, and the discussion period will last 1 minute.

Once the discussion period has concluded, you will be able to view a transcript of the discussion, with players’ ID numbers indicating their responses, as well as a table of player ID numbers, their proposed answer and their self-reported confidence in their answer. You will then be asked the following questions:

1. Please select one player to act as the **chosen representative** for the group. If selected, this chosen representative will submit one final answer on behalf of the entire group. **You may not** select yourself.
2. On a scale of 1-10, please indicate how confident you feel that your chosen representative will submit the highest scoring answer.
After specifying your preferred chosen representative, you will answer another question:

What answer do you expect your chosen representative to submit on behalf of the group if they are chosen to fulfill this role?

Once all the players have answered the above question, the group leader will be randomly selected. Each player has an equal chance of being designated the group leader.

The group leader’s chosen representative will then submit one final answer to the question on behalf of the entire group. The chosen representative may submit an answer of his or her choosing.

After the chosen representative has submitted an answer, the group’s score breakdown will be displayed on each group member’s screen.

• In each round, all members of the group receive the points equal to the score of the answer submitted by the chosen representative.
• Recall that only answers that have received two or more survey responses count obtain points. If the chosen representative’s answer does not appear in the table of answers, you will receive zero points.
• Being selected as the group’s chosen representative earns additional points.

The experiment will end with a post-experiment questionnaire. Upon completion of the questionnaire, you will receive your cash payment.

**Final Payoffs (All Treatments)**

At the end of the experiment, one of the two parts of the experiment (individual task or group task) will be randomly selected for payment based on the cumulative points earned in that part. In addition, you will have a chance to earn extra points for every time you were chosen as the group’s representative to answer a question in Part 2.

In summary, your total payoff from the experiment will be calculated as follows:

$10 show-up fee  
+ $0.12 for every point you accumulated over the 4 rounds in Part 1 OR in Part 2  
+ $0.50 for each time you were chosen as the group’s representative to answer a question in Part 2  
+ completion fee (if any)
For example, a player who accumulated 90 points in Part 1 (individual play) and 100 points in Part 2 (group task), with Part 1 randomly selected for payment, and who was chosen as the group’s representative twice would receive the following pay-off (without a completion fee):

\[
$10.00 + ($0.12 \times 90) + ($0.50 \times 2) = $21.80
\]

\[
\text{Part 3: Post Experiment Questionnaire (All Treatments)}
\]

[Note that all questionnaires will be computerized.]

Thank you for your participation in this experiment. You will now be asked a short series of questions regarding your demographic information and your strategy in the experiment. We ask you to answer all questions as truthfully as possible. Your answers will help us tremendously! Feel free to leave any of the answers blank if you prefer not to answer.

Please remember that all information collected during the experiment will be kept private and that your decisions will remain anonymous to other participants and the experimenters.

After completing this brief survey, you will be privately paid your experimental earnings and will be free to go.

Please do not hesitate to speak to the experimenter should you have any questions regarding the experiment or your experience as a participant.

1. Gender:  □ Male  □ Female

2. Age:  

3. Marital status:
   □ Now married
   □ Widowed
   □ Divorced
   □ Separated
   □ Never married

4. Which of the following best describe(s) your race (check all that apply)?
   □ Asian
   □ Black
   □ Arabic or North African
   □ Hispanic
☐ White
☐ Other (please specify) __________________

5. Country of birth: __________________

6. Country of citizenship: __________________

7. Primary language: ______________

8. Please rate your English language proficiency:

☐ Elementary  ☐ Intermediate  ☐ Advanced  ☐ Fluent  ☐ Native or bilingual

9. Are you currently a student? ☐ Yes ☐ No

1) If yes, what is your student status? ☐ Bachelors student ☐ Graduate student

2) If yes, what is (are) your field(s) of study? __________________________

10. What is (was) your undergraduate grade, on a 4.0 scale? __________________________

11. What is your family’s total household income?

☐ Under $20,000  ☐ $20,000 to $45,000  ☐ $45,000 to $65,000  ☐ $65,000 to $90,000  ☐ $90,000 to $125,000  ☐ $125,000 to $200,000  ☐ $200,000 or more
12. From the list of experiences below, please select all that apply.

- I do not have any past leadership experience.
- I have been a leader of a group of others as part of my educational activities at school or university.
- I have been a leader of a group of others as part of my extracurricular activities (such as sports, student organizations, etc.).
- I have been a leader of a group of others as part of my work or internship experiences.

Please, give specific examples of your leadership experiences below (if any):

____________________

13. Are you at all familiar with the television show *Family Feud*?

- No
- Somewhat familiar
- Very familiar

14. When engaging in group discussions, what factors motivate you to consider the proposals of others? Please mark all that apply.

- The fundamental content of the proposal
- The reputation of the individual who is making the proposal
- The confidence and clarity with which the proposal is presented
- The frequency of proposals from that individual
- Lack of confidence in my own proposals

15. Did you recognize anyone in your group as a personal friend, colleague, or classmate?

- Yes
- No

If yes, did that knowledge change your decisions in the experiment?

- Yes
- No

If yes, please explain how so: __________________________

16. Did you use your real name in the experiment? (only for treatments with name identifier).

- Yes
- No
References


