Journal of Economic Behavior and Organization

(Un)Willing to lead? Men, Women and the Leadership Gap

--Manuscript Draft--

Manuscript Number:	JEBO-D-20-01486
Article Type:	VSI: Men and women behaviour
Keywords:	Experiments; Coordination, Gender; Leadership
Corresponding Author:	Yaxiong Li, Ph.D. Xi'an Jiaotong-Liverpool University Auckland City, NEW ZEALAND
First Author:	Yaxiong Li
Order of Authors:	Yaxiong Li
	Ananish Chaudhuri
	Erwann Sbai
Abstract:	We explore the causes behind the gender gap in leadership; as one moves up the organizational hierarchy, one encounters fewer women. We use the weak-link game paradigm to simulate intra-organization coordination problems, where participants can volunteer for leadership roles. The leaders' job is to resolve potential coordination failures. We look at whether: (1) there are systematic gender differences in the willingness to lead and (2) followers are less likely to follow female leaders. We find that, compared to men, fewer women volunteer to lead, particularly when the leader's gender is revealed to the followers. But, by and large, male and female leaders choose similar messages and/or actions in this game, and controlling for those, groups achieve similar levels of coordination success regardless of the leader's gender. We do not find evidence of resistance against female leadership, even though anticipation of such backlash may lie behind the female reluctance to lead.

Declaration of interests

To: Journal of Economic Behavior & Organization

Date: 26 November 2020

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Sincerely,

Yaxiong (Sherry) Li

BS318 International Business School

Suzhou 8 Chongwen Road

Suzhou Dushu lake Science and Education Innovation District

Suzhou Industrial Park

Suzhou

P.R. China

215123

Email: sherry0520@gmail.com

Yaxiong.li@xjtlu.edu.cn

Tel: +86 (0)512 8188 9052

(Un)Willing to lead? Men, Women and the Leadership Gap

Abstract

We explore the causes behind the gender gap in leadership; as one moves up the organizational hierarchy, one encounters fewer women. We use the weak-link game paradigm to simulate intra-organization coordination problems, where participants can volunteer for leadership roles. The leaders' job is to resolve potential coordination failures. We look at whether: (1) there are systematic gender differences in the willingness to lead and (2) followers are less likely to follow female leaders. We find that, compared to men, fewer women volunteer to lead, particularly when the leader's gender is revealed to the followers. But, by and large, male and female leaders choose similar messages and/or actions in this game, and controlling for those, groups achieve similar levels of coordination success regardless of the leader's gender. We do not find evidence of resistance against female leadership, even though anticipation of such backlash may lie behind the female reluctance to lead.

JEL Classification Codes: C91, D91, J16

Keywords: Experiments, Coordination, Gender, Leadership

1. Introduction

The "gender gap" in the workplace is a well-documented phenomenon. However, it is important to recognize that this gap, in reality, refers to two phenomena, which are certainly related but still, not exactly the same. The first is the gender gap in wages: that mean/median earnings for men is higher than that for women. The second refers to the gender gap in leadership roles: There are fewer women as we move up organizational hierarchies.

In this paper, we focus on the second phenomenon; the gender gap in leadership and explore two issues. First, we look at whether there are systematic gender differences in the willingness to lead. Second, we explore whether there are differences in the perception of female leaders compared to male leaders, in the sense of followers' willingness to follow male leaders more than female leaders. In particular, we intend to see whether messages from male and female leaders – even where the content of the message is identical – have a different impact on the actions of the worker. In order to study this issue, we rely on the minimum effort coordination game paradigm, which serves as a good vehicle for simulating intra-organizational coordination problems. We use a modified version of the Brandts and Cooper (2006, 2007) corporate turnaround game. We explain the details of the game below in the section on experimental design.

We present results from two sets of studies. In one, both leaders and followers face the same payoffs while in another the leaders experience a payoff matrix that is different from that for the followers. Within these studies we manipulate two factors: in one treatment, the followers get to learn the leaders' gender while in another they do not. We also manipulate the nature of messages that leaders can send to their followers. In one treatment, the message sent is pre-determined and written by the experimenter. The leader only gets to choose how frequently to send this message. This ensures that, if and when sent, the message coming from the leader is identical across genders. In a second treatment, leaders are allowed to write free-form messages. This generates a 2x2 protocol consisting of four treatments: (1) Leader's gender not revealed; pre-set message; (2) Leader's gender revealed; pre-set message; (3) Leader's gender not revealed; free-form message and (4) Leader's gender revealed; free-form message.

Overall, we find that fewer women volunteer to the leadership role compared to men. This is true whether the payoff matrix for the leaders is the same as that of the followers or not. But the decision to lead or not does depend crucially on whether the leader's gender is revealed to group members or not. When gender information is revealed, there is a significant gap in the rate at which men and women volunteer to lead with men being much more willing; when gender is not revealed, this gap shrinks considerably. However, when it comes to leader effectiveness, we do not find evidence of significant differences. By and large, messages and actions taken by leaders are similar, even if not identical, and this, in turn, implies that groups led by women enjoy similar levels of coordination success as those led by men.

If women are more reluctant to avail of leadership roles, then this may partially explain the existing gender gap in leadership. To an extent, this reluctance to volunteer for leadership may be predicated on the assumption that followers are less likely to follow female leaders. We do not find evidence of any systematic resistance to female leadership. This seems to offer a learning opportunity that oft-held presumptions may not be necessarily correct and provides support for more pro-active equity-based practices in the workplace.[1] We proceed as follows. In Section 2, we provide a brief review of the relevant background literature. In Section 3 we outline our experimental design and procedures. We present our results in section 4 and make concluding remarks in Section 5.

2. The gender gap in wages and leadership roles

Despite the progress in female educational attainment and increasing parity in paid hours of work, and occupational choice, women are still earning less than men in same or similar jobs (Goldin, Katz and Kuziemko, 2006; Bertrand, Goldin and Katz, 2010; Blau and Kahn, 2000; 2006, Drolet, 2001, Goldin, 2014, Weichselbaumer and Winter-Ebmer, 2005). This is true not only for women in the general labour force, but also for those who have graduated from MBA and other professional programs, presumably with the aim of pursuing ambitious managerial/professional careers. (Babcock and Laschever, 2003; Bertrand,

[1] See Erkal, Gangadharan and Xiao (2019) for an example of one such policy initiative based on changing the organizational default from opting-in to leadership roles to opting-out of leadership roles. Bohnet (2016) and Eckel et al. (2020) provide other examples of changes in institutional rules may achieve greater gender parity.

Goldin and Katz, 2010; Carter and Silva, 2010). Wood, Corcoran and Courant (1993), in a study of law school graduates find that even after controlling for the "motherhood penalty" and other factors such as school performance and work-history, one-third to one-quarter of the wage gap is left unexplained. Jena, Olenski and Blumenthal (2016) confirm a significant gender gap in wages among more than 10,000 physicians at 24 public medical schools in United States after controlling for factors such as age, experience, specialty, faculty rank, measures of research productivity and clinical revenue. Moreover, the gender gap in wage at the top of the wage distribution has declined much more slowly than at the middle and bottom (Blau and Kahn, 2017).[2]

We refrain from elaborating any futher on the gender gap in wages since our focus in this paper is on the gender gap in leadership. Women currently hold 5% of CEO positions in S&P 500 companies. Among the CEOs of the largest publicly listed companies in the European Union, only 5.5% are women. Women constitute only 9.4% of the 540 C-level executives among Canada's 100 largest publicly traded companies. (Catalyst, 2018). Sandberg (2013) comments that at the time of writing: Of the 195 independent countries in the world, only 17 were led by women. Women held just about 20% percent of seats in parliaments globally and about 14% of executive officer positions, 17% of board seats, and constituted 18% of elected congressional officials in the US.

Of course, if more men than women are occupying higher paid jobs, this will also show up as a higher average wage for men. Some of the factors that contribute to the gender gap in wages almost certainly also create the gender gap in leadership roles. Early research in the gender wage gap attributed this phenomenon mainly to differences in human capital and either taste-based or statistical discrimination. (E.g., Altonji and Blank, 1999). But, as Bertrand (2011) points out, in recent years a large body of research – to a large extent experimental – suggests the possibility that observed gender gaps may arise due to

[2] The OECD defines the gender wage gap as the difference between median earnings for males and females relative to the median earnings for males. It should be noted that there are significant cross-country differences. Looking at data for 2015-2018, we find that the average gender wage gap for all OECD countries is 13.5%. At the most unequal end we have Korea (with a difference of 35%), followed by Japan (25%), Israel (22%) and then Canada, USA and Finland (around 18%). At the other extreme, we have Belgium, Greece, Costa Rica, Denmark and Italy, all hovering around 5% with Ireland, Norway and Sweden at about 6% and New Zealand at a little less than 8%.

important differences in psychological attributes and preferences between men and women. Such differences may include gender differences in risk preferences, in attitudes towards competition and negotiation and in other-regarding preferences. This, in turn, may also have implications for gender differences in occupational choice, work-place strategies and consequently wages and career advancement.

Experimental evidence suggests that women tend to shy away from participating in highly competitive environments (Gneezy, Niederle and Rustichini, 2003; Niederle and Vesterlund, 2007); are less likely to engage in negotiations for pay and promotions (Babcock and Laschever 2003; Exley, Niederle and Vesterlund 2020) and are more risk-averse compared to their male counterparts (Croson and Gneezy, 2009; Charness and Gneezy, 2012, Coffman, 2014). There is also evidence suggesting that as the environment becomes more competitive, the performance and participation of men increase relative to that of women. (Andersen, Ertac, Gneezy, List and Maximiano, 2013; Vandegrift & Yavas, 2009, Gneezy, Leonard and List, 2009). In addition, recent studies have found that women are more willing to take on tasks that are less likely to benefit their performance evaluations or advancement prospects whereas men tend to focus more on high-promotability tasks. (Babcock, Recalde, Weingart and Vesterlund, 2017). Finally, in agency relationships, compared to male employers, female employers routinely offer more generous contract terms to employees, which does not necessarily elicit higher effort from the latter. This has a negative impact on the earnings of female principals. (Chaudhuri, Cruickshank and Sbai, 2015).

Babcock and Laschever (2003) and Niederle and Vesterlund (2007) argue that greater female aversion towards competition may explain why one finds fewer women occupying positions of power. Sandberg (2013) suggests that in order to be successful in the work-place women need to adopt more assertive negotiating tactics, i.e., they need to "*lean in*" more, a trait usually associated with males. Bohnet (2016) not only provides a comprehensive overview of what we currently know about the gender gap, but also makes numerous policy suggestions as to how we can go about creating a more equal work-place.[3]

[3] Such include, for instance, the introduction of "blind" auditions for orchestras (Goldin and Rouse, 2000) or the use of structured interview questions with the same questions being asked in the same sequence of all job candidates as well as evaluating job candidates contemporaneously rather than sequentially. (Bohnet, van Geen and Bazerman, 2016).

The experimental literature looking at gender differences is large. We refer the interested reader to Eckel and Grossman (2008) and Croson and Gneezy (2009) for reviews primarily with regards to differences in risk attitudes and other-regarding preferences. Babcock and Laschever (2003) and Bohnet (2016) provide comprehensive discussions of the topic from a labour economics perspective. For a selection of findings from the psychological point of view, see Walters, Stuhlmacher, and Meyer (1998) or Sax (2005).

This line of inquiry overlaps with research in leadership; specifically whether men and women tend to adopt different styles when it comes to dealing with employees. Rosener (1990), building on concepts introduced by Burns (1978), argues that men typically tend to be "*transactional*" leaders and see job performance as a series of transactions with subordinates involving rewards for services rendered and punishments for inadequate performance. Women on the other hand are seen as being more "*transformational*", relying less on explicit rewards and punishments and more on a democratic and participative style. Li, Sbai and Chaudhuri (2020) explore this issue further in the context of a principal agent relationship by looking at the nature of contract choices by male and female leaders. They define a "transformational" relationship as one where the employer-employee relationship is based on mutual trust and reciprocity. A "transactional" relationship, on the other hand, relies on contracts that specify explicit rewards and punishments for effort provided (or not). Li et al. (2020) do not find significant differences in the rates at which male and female leaders choose one over the other contract. In fact, there is a preponderant tendency for leaders of both genders to choose the trust-based transformational contracts.

Eagly and Johnson (1990) undertake a meta-analysis of 162 studies on leadership and find little difference between male and female leadership styles. They find some support for the view that women adopt a more democratic style while men tend to adopt a more authoritarian style.[4] Eagly, Karau, and Makhijani (1995) undertake a further meta-analysis which extends the analysis of leadership styles to the issue of leadership effectiveness. They report that men and women are equally effective as leaders except

[4] The authors suggest that these differences may arise in part from the fact that women, being outnumbered by men as leaders, face greater resistance from employees and feel the need to seek greater employee input.

that men tend to be more effective in occupations that are typically defined in more masculine terms such as the military while women are more effective in occupations defined in primarily feminine terms such as nursing. See Moran (1992) for a succinct overview of much of this work.

However, some other studies report that men are perceived as more effective leaders when gender role expectations spill over into leadership roles. Indeed, men are found to exert more influence than women in mixed groups and are more resistant to female leadership. (Gangadharan, Jain, Maitra and Vecci, 2016). Followers are likely to ignore and/or dismiss actions by female leaders because they do not view women as legitimate leaders. (Ridgeway, 2001). MacNell, Driscoll and Hunt (2015) examine students' evaluations of instructors in an online course. They find that regardless of the actual gender of the instructor and even after controlling for teaching quality, students give higher ratings to perceived male instructors than perceived female instructors. Such biased perceptions may lead to undervaluation of a woman's effectiveness as a leader.

There are three prior studies that are similar in spirit to ours. In the first of these, Grossman, Eckel, Komai and Zhan (2019) examine followers' perception of a leader's gender using experiments involving a coordination game with groups of five participants, interacting repeatedly for 20 rounds. Prior to the 11^{th} round, an exogenously appointed leader provides advice to the group members on how to coordinate on the payoff maximising outcome. At the end of the session, the followers choose a bonus payment for the leader. The authors find that for such exogenously appointed leaders, women are assessed less positively and rewarded less generously than men even if the former are equally effective in resolving intra-organizational coordination problems.

Reuben and Timko (2018) examine gender differences in leadership using a minimum effort coordination game. Leaders are either democratically elected (the *election* treatment) or randomly selected (the *random* treatment). They remain leaders for three periods and then the process is repeated. At the beginning of the first of the three periods the leader can write a non-binding message to their group members. Reuben and Timko (2018) find no gender differences in the effectiveness of the leaders in the random treatment. In the election treatment, male leaders enjoy greater benefits in the initial interactions,

but this gender difference dissipates over time. Overall, groups perform better in the election treatment than in the random treatment, and groups with female leaders perform as well as groups with male leaders. However, the authors note that unsuccessful male leaders are re-elected more often than unsuccessful female leaders.[5]

Grossman, Komai and Jensen (2015) undertake a study that is closest in spirit to ours. Groups of three players have to choose whether to invest or not in a project. In effect, this sets up a coordination game with strategic complementarities resulting in multiple payoff-ranked equilibria. Subjects are presented with three scenarios and the game parameters are such that in two of the three situations social efficiency dictates all players choosing to invest resulting in the payoff dominant outcome; in the third scenario, all three should choose not to invest leading to the "secure" outcome. The leader receives information about which of the three scenarios is applicable and then decides whether to invest in the project or not. The two followers get to see the leader's decision and then decide simultaneously as to whether to invest or not.

Grossman et al. look at four different treatments: one treatment, where groups consist of both genders and the leader's gender is revealed; a second, with mixed gender groups but with the leader's gender hidden; a third, with all male members and a fourth, with all female members. The authors find that in both single gender treatments and in the treatment where the leader's gender is not revealed, female leaders are more likely to invest than male leaders in those situations where investment is commensurate with the social optimum. However, when the leader's gender is revealed, female leaders are less likely to invest than male leaders in a situation where followers' refusal to follow can reduce the leader's payoff. The followers' decisions on investment are not affected by the leader's gender. Below, we report some results that are similar to Grossman et al. (2015). First, female leaders show a hesitation to lead in mixed gender environments with gender information revealed in circumstances where followers' refusal to follow

[5] Gallup has been tracking workers' preferences for their superior's gender since 1953 when 66 percent of American adults said they would prefer a male boss and only 5 percent said they would prefer a female boss. Although this gap has narrowed as of 2013, it still remains. According to 2013 Pew Research Centre survey, around three-quarters of men and women who are currently working or ever worked have no preferences regarding their boss's gender, but among those who do have a preference, both genders are more likely to prefer a male boss.

can reduce the leader's payoff. Second, in our study, like theirs, the follower's decision is not affected by the leader's gender.

However, there are significant differences between the two approaches. As Grossman et al. point out, they reduce the leadership question down to its basic essentials by ignoring any issues of leadership style. Their leader is an average player who is selected at random and derives his/her legitimacy from the fact that he/she possesses better information about the state of the world. The random leader selection also implies that Grossman et al. are unable to comment on the issue of willingness to lead and whether there are gender differences in the same; a key focus in our paper. It is also the case that in our study the leader has significant experience in the coordination problem prior to volunteering to lead. Finally, we adopt a richer experimental design that allows us to explore issues involving leadership style and messaging that Grossman et al. strip away.

3. Experimental Design and Procedure

3.1. Experimental Design

We rely on a modified version of the Brandts and Cooper (2006, 2007) corporate turnaround game which, in turn, is derived from the minimum effort (weak-link) game of Van Huyck, Battalio and Beil (1990). We provide the instructions in Appendix A. Each session consists of 20 rounds with the first ten rounds being identical. Participants are randomly assigned to groups of five at the start of the session, and group composition remains unchanged for the entire session. In each of the first ten rounds, each participant simultaneously chooses an effort level ("hours of work") {0, 10, 20, 30, or 40}, where earnings depend on the participant's effort choice and the minimum effort level chosen in the group in that round. Earning for each player is determined by the underlying equation:

$$\pi_i = 400 - 5(E_i) + B\left(\underset{i\in\{1,2,3,4,5\}}{\text{Min}}(E_i)\right) \qquad (1)$$

where π_i indicates earning for player i, E_i is the effort level player i chooses, B is a constant set equal to 10 and *Min* (E_i) is the minimum effort level chosen in the group for that round. The earnings are depicted

in Table 1. The values are denoted in experimental dollars with 400 experimental dollars equal to 1 New Zealand dollar.

<Table 1 about here>

In this payoff matrix, any common effort level (hours) chosen by the five players constitutes a Nash equilibrium with everyone choosing 40 hours being the payoff dominant outcome while everyone choosing zero hours is the secure outcome. Each player faces the dilemma that higher effort level (and therefore, higher disutility of effort) is required in order to achieve higher earnings, but a higher effort level also entails more risk in the form of reduced payoffs, if even one member of the group deviates and chooses a lower effort level. Earlier studies suggest that over time play typically approaches the secure Nash equilibrium where all players choose the lowest effort level and that players find it difficult to coordinate to the payoff dominant equilibrium for any length of time. (Van Huyck et al., 1990; Knez and Camerer, 1994, 2000, Chaudhuri, Schotter and Sopher, 2009).

Prior to the start of round 11, participants are given an opportunity to volunteer to be the group leader. If more than one person wishes to be a leader, then the leader is chosen randomly from the group of contenders.[6] We conduct **two separate** experiments. In **Experiment 1**, following the selection of the leader, the payoff matrix from the leader changes. The leader's payoff is now determined by the following equation:

$$\pi_L = 100 + [(60 - 4B) \times \min_{i \in \{1,2,3,4,5\}} (E_i)] - \mathrm{c} \qquad (2)$$

If the leader's choice of hours exceeds that of the workers, then there is an additional cost to the leader of the form:

$$c = 2\left(E_L - \mathrm{Min}(E_{(-i)})\right) \qquad (3)$$

where E_L is the leader's effort level and $\mathrm{Min}(E_{(-i)})$ refers to the effort level chosen by other group members excluding the leader. We think of this as a psychological cost of feeling "ripped-off" when the leader

[6] A leader is chosen randomly among group members if no one volunteers. But it never happened that there were no volunteers.

chooses an effort level that is higher than that chosen by one or more followers. This psychological cost is zero if and when the leader's effort is the same as the group minimum. Table 2 shows the payoff for leaders.

<Table 2 about here>

In **Experiment 2**, the payoff matrices for both leaders and followers are the same for all 20 rounds and shown by the payoff matrix in Table 1. Other than the change in the leader's payoff matrix in Experiment 1, the two studies are conducted in the exact same way. Clearly, the payoff matrix in Table 2 (for Experiment 1) presents both gains and losses for the leader. We note a few points. From Table 1, it is clear that any team member can guarantee him or herself a payoff of 400 by choosing zero hours in the secure equilibrium. Now, looking at the leader's payoff matrix in Table 2, we find that if the minimum effort in the group ends up being zero or 10 hours, then the maximum the leader can earn is 300. It is only when the team manages to coordinate to a minimum effort level of 20 hours or more that the leader makes more than 400, which any team member can make in the secure equilibrium. If the team minimum is 20 hours, then the leader makes 500. The leader makes more if the team manages to coordinate to a higher minimum.

If we find greater female reluctance to lead in Experiment 1 where the leader's payoff changes, then this may partially be attributable to differences in risk attitudes. To control for this, we run Experiment 2, where the payoff matrix remains unchanged for both leaders and followers.[7] We also control for risk preferences using the Holt-Laury lottery choice mechanism (Holt and Laury, 2002). In what follows, we will periodically remind readers of the difference between Experiments 1 and 2; an easy mnemonic is "*1 for change, 2 for none*"; meaning that in Experiment 1, the leader's payoff changes following leader selection and prior to Round 11 while in Experiment 2, both leaders and followers continue to face the same payoff matrix following leader selection.

[7] Experiment 2 only partially controls for risks; some risks remain. The fact that we do not get different results for experiment 2, may be due to the fact that the gender difference is not due to differences in risk aversion and/or because we have not sufficiently eliminated the risk. We thank a referee for pointing this out to us.

3.2. Treatments

Within each experiment, we vary the experimental design along two dimensions: whether the gender of the leader is revealed or not; and whether the leader sends a fixed message, that is provided by the experimenter or the leader is allowed to write free-form messages. This generates a 2X2 protocol consisting of four treatments: (1) Leader's gender not revealed; pre-set message; (2) Leader's gender revealed; pre-set message; (3) Leader's gender not revealed; free-form message and (4) Leader's gender revealed; free-form message.

In our pre-set message treatments, participants are informed that if they choose to be a leader, then they will be provided with a message that they can send to the other members of their group. In each round, the leader moves first by choosing the number of hours. The leader can also choose to re-send the fixed message or not. Once the leader has chosen the number of hours, the information regarding the leader's choice of hours and the content of the message will be revealed to the group members. The fixed message that leaders could disseminate is:

> *You should choose to work 40 hours in each round. NOTICE, from the payoff matrix, that if every participant in a group follows the message then every participant will earn 600 experimental dollars. However, if even one of the participants does not follow the message and chooses a number different from 40, then each participant will make less money than if everyone chose 40.*[8]

In free-form message treatments, leaders are asked to type a message which they can send to their group members. Figure 1 shows a screenshot for a leader at the beginning of round 11 under Free-form message treatments. They can write a different message each round. In each round, the leader moves first by choosing the number of hours. The leader can also choose to re-send the message sent before or write a

[8] An astute reader may note that in Experiment 1, this is true of all the followers but not of the leader since the leader's payoff matrix changes and the leader gets to make more than 600 if everyone coordinates to 40 hours. This instruction is accurate for Experiment 2, where there is no change in payoffs for either followers or leaders. We decided to use this language in the interests of consistency. Otherwise, we would have had to use the word "followers" for Experiment 1 and "participants" for Experiment 2. This would have introduced a potential confound. In Experiment 1, we felt that given that the followers are getting this message *after* they have seen the leader's effort choice and the leader's message, it would be obvious in the context that here participant means the followers, who are receiving this message. In Experiment 2, this is correct that choosing 40 hours means all participants earn 600 units in each round.

different message. Once the leader has chosen the number of hours, the information regarding the leader's choice of hours and the content of the leader's message is revealed to the group members. The employees then choose their hours simultaneously. It is important to note that the person selected to be the leader **must** send a message prior to Round 11. From that point, the leader gets to choose whether to send any further messages or not. Not sending any more messages is an option as is the option to send a message prior to every round.

<Figure 1 about here>

In gender revealed treatments, the leader's gender is revealed to the group members in addition to the choice of hours and content of the message. Figures 2A and 2B are screenshots for team members in Free-form message treatment with gender not revealed (Figure 2A) and revealed (Figure 2B) at the beginning of round 11. The main difference is that in the former the followers do not know the leader's gender while in the latter this information is available on the screen. Each participant sees a screen that indicates if he or she has been selected as a leader. If not, the screen shows the ID number and gender of who has been selected as the leader. In the gender revealed treatments, the leader's gender information is made salient over multiple screen messages.[9] Experiment 1 results clearly demonstrate that any differences, if they arise at all, are more prominent in the gender revealed treatment. Therefore, in experiment 2, we only conduct the gender revealed treatments.

<Figures 2A and 2B about here>

[9] We chose to have the leader's gender being revealed as being male or female; in the sense that the followers get a message saying: "Your leader is female/male." One drawback to doing things this way is that this may create experimenter demand effects. There are different ways one can go here and all of them pose difficulties. One option used often is to ask participants to choose from a list of proto-typical male or female names, typically European ones such as Adam or Lisa. But the University of Auckland is one of the most international universities in the world. We were uncertain whether asking a large number of non-European students to explicitly choose an European pseudonym would be well received or not. On the other hand, if we used non-European names, it was not clear if the gender would automatically be salient to everyone. Other studies have used virtual avatars. It is our feeling that asking participants to adopt another name or a virtual avatar etc., at times, start to approach the line separating truth and deception. There is increasing scrutiny and concern about deception in experiments even if benign. We decided to avoid this and to go with the label male/female. This is partly because as Zizzo (2010) points out: experimenter demand effects become less of a concern where the demand effect explicitly relates to the variable, which is being manipulated and is of primary interest. Here, as in many other cases, one faces design trade-offs. We chose to go with this particular approach.

Prior to choosing whether to volunteer as a leader, each group member knows whether the leader's payoff will change and how (as in Experiment 1) or not (as in Experiment 2). Each person also knows whether they are in the pre-set message treatment (i.e., the message is provided to them by the experiments and all they need to do is to decide how often to send this message out to group members) or the free-form one. They also know that the leader has to send a message prior to Round 11 but can choose whether or not to do so for the subsequent rounds. All of the above is common knowledge.

3.3. Post-experiment tasks

Following the conclusion of the 20 rounds of the coordination game, participants are asked to take part in the Holt and Laury (2002) lottery choice game with choices being incentivized. Subjects are presented with 10 different binary lottery choices and must choose between either option A or option B for each paired gamble. (See Appendix B)

An individual's degree of risk aversion is determined by the point at which that person switches from option A to option B. A risk-neutral person would switch from option A to B at the fifth choice. An individual who switches to Option B earlier than the fifth choice (i.e. choices 1 through 4) is risk loving, while an individual who switches to option B at the sixth choice or later is risk averse. Participants are informed that this is a separate task, which will be pay out according to their choices. At the end of the session the computer randomly chooses one of the gambles (rows) and each participant is paid based on whether he/she chose Option A or Option B for that gamble. In our overall sample, we have 27% of all participants with inconsistent risk preferences, 42% of all participants are risk averse, 25% of all participants are risk neutral and 6% of participants are risk loving. In our regression analyses below, when we control for risk preferences, we will ignore the subjects with inconsistent preferences, leading to a loss in the number of observations in some cases.[10]

After completing the lottery task, participants completed a demographic survey (see Appendix C) collecting information regarding participant's gender, field of study, year in the undergraduate program,

[10] The existence of subjects who demonstrate inconsistent preferences in the Holt-Laury task is a well-documented drawback of this particular way of measuring risk preferences.

age, income, whether they were born in New Zealand, and their ethnicity. In some sessions, we also asked subjects to fill out a questionnaire regarding their attitudes toward female leadership.[11]

3.4. Experimental Procedure

A total 330 students took part in Experiment 1, and 85 students took part in Experiment 2. Tables 3 and 4 provide details for experiment 1 and 2, respectively. All sessions in this experiment were conducted in the DECIDE laboratory at The University of Auckland using Z-tree (Fischbacher, 2007). Participants were recruited via an email announcement and were students from undergraduate courses without any prior experience with the turnaround game.

<Table 3 about here>

<Table 4 about here>

Participants are directed to computer cubicles once they enter the lab. There are dividers between each cubicle so that participants are separated from one another and are unable to see any other participant's screen. They are also cautioned against communicating with others. Participants know that they are randomly assigned to a group of five and that the group composition will remain unchanged for the entire time. Each participant is assigned a subject ID number and never learn the actual identity of any of the others in his or her group. Participants know that the experiment consists of two parts of 10 rounds each. We read them the instructions for the first 10 rounds at the beginning. Participants then play the first 10 rounds of the game. After the conclusion of the first ten rounds, participants receive the instructions for the second part of the experiment. The instructions are read out loud to them.

[11] Subjects are asked to answer the following five questions. (1) Females do not possess good leadership quality. (2) Because leadership is viewed as a masculine trait, females will not be viewed as strong leaders. (3) Females are too emotional to lead effectively. (4) Females are capable of performing effectively in any leadership position. (5) Male leaders connect with the public better than female leaders. All questions are answered on a 1-5 scale from "Strongly Disagree" to "Strongly Agree" with appropriate scores reverse scored as relevant. Using various non-parametric tests, we do not find any significant differences between male and female responses regarding their view of female leaders. Therefore, in what follows, we do not elaborate on these survey responses.

Once the instructions for the second half of the experiments (after Round 10) have been provided, participants first see a screen, which asks them whether they wish to volunteer to be a leader or not. Once all participants have entered their decisions, one of them is randomly selected as the leader. At this point, all participants get to know who was selected as leader (via subject ID only), and in the gender revealed treatments, participants also learn whether this leader is male or female. Then, the leader enters his/her effort choice for Round 11 along with the message, either pre-set or free-form. The followers then get to see the leader's message and effort choice and enter their own effort choices. Everyone then gets to see the outcomes (the effort choices of the leader and of the followers) and earnings from that round. From Round 12 onwards, the leader has the choice of whether to send a message or not. Once all 20 rounds are complete, participants are asked to take part in the Holt-Laury lottery choice task and also fill out other questionnaires, including the one asking for demographic information. Figure 3 shows the timeline for both experiments 1 and 2.

<Figure 3 about here>

Each session lasts approximately 60 minutes. At the end of the session, participants are paid their earnings in cash plus a $5 show-up fee. Participants are told that their earnings are private information and they should keep this private. Once paid, they are free to leave. Average payoff for the leadership game is NZ $29 including the show-up fee. In addition, subjects would have earned another additional amount of approx. NZ $5, on average, depending on the outcome of the Holt-Laury lottery-choice game.

4. Results

We report our results in two sections. First, we examine whether men and women differ systematically in their willingness to lead. Second, we look at the impact of leaders' gender on the coordination level of group members in both pre-set message treatment and free-form message treatment.

4.1. Willingness to lead

As noted, upon the conclusion of the first 10 rounds and prior to beginning Round 11, participants are asked whether they wish to be a leader or not. We find that in both Experiment 1 (where the leader faces a different payoff matrix) and Experiment 2 (where the payoff matrix remains unchanged) a much larger proportion

of men volunteer to lead. Figure 4, Panels A and B show the breakdown. In Experiment 1, 76% men volunteer to lead as opposed to only 51% of women. There is a similar discrepancy in Experiment 2, with nearly 80% of men expressing a willingness to lead while only about 60% of women do so. This suggests that it is not the change in the payoff matrix that is driving these results.

<Figures 4A and 4B about here>

In Figure 5, Panels A and B, we provide a more detailed break-down of what happens when the leader's gender is revealed to followers as opposed to when it is not. We start by looking at Panel A of Figure 5, which shows the gender breakdown for the pre-set and free-form message treatments separately. This figure reveals some striking differences. It is clear that women are much more reluctant to lead, when the leader's gender is made known to the followers. If we look at the first and third pair of bars, which represent the situation when gender is not revealed, we find that any gender differences in the willingness to lead are not large. 76% of men volunteer as opposed to 63% women in the pre-set message gender not revealed treatment. The corresponding figures are 56% for men and 53% for women in the free-form message gender not revealed treatments. Using a sample proportions test, these differences are not statistically significant.

But if we now compare the second and fourth pairs of bars, for the gender revealed treatments, then the differences are dramatic. In the pre-set message gender revealed treatment, 73% of men express a willingness to lead, while only 44% of women do so. In the free-form message, gender revealed treatment, the differences are even more pronounced; 83% for men and 51% for women. Both of these differences are highly significant using a sample proportions test. ($|z|=2.84$, $p<0.01$, m=40, f=45 for the pre-set message gender revealed treatment and $|z|=5.33$, $p<0.01$, m=86, f=79 in the free-form message gender revealed treatments.)[12] The results of Experiment 2, where we only look at the gender revealed condition, are similar

[12] In order to increase the statistical power of our tests, we also carry out similar tests after pooling the data across the two message treatments in Experiment 1. Given that the decision regarding leadership is taken before sending any messages, we look at the gender-revealed and gender not revealed treatments in aggregate. This allows us to compare the decisions made by 39 males and 41 females in the two message treatments where gender is not revealed. There are no significant differences here. ($|z|=0.76$, $p>0.1$). But if we combine over the two message treatments, where

except that the proportions of men and women volunteering to lead in the pre-set message treatment are not significantly different (|z|=1.05, p>0.1, m=24, f=21) while the difference in the free-form message treatment is significant (|z|=1.86, p<0.1, m=17, f=23).

<Figure 5 Panels A and B about here>

In Table 5, we present the results of probit regressions for differences in the willingness to lead. The dependent variable is an individual's response before round 11 (=1 if willing to be leader, =0 otherwise). The first two columns (Models 1A and 1B) present results for Experiment 1, while the third and fourth columns (Models 2A and 2B) do so for Experiment 2. Given that the leadership choice occurs prior to sending messages, in this table we have combined the pre-set and free-form messages for both Experiments 1 and 2. In each case, the first specification (Models 1A and 2A) includes a female dummy (Female= 1 for female, = 0 for male), average effort in the first 10 rounds, average earnings in the first 10 rounds, and demographic controls. The second specification (Models 1B and 2B) control for risk preferences using the decisions made in the Holt-Laury (2002) lotter-choice task.[13] The female dummy is negative and significant at 1% level in Experiment 1, clearly suggesting that fewer women choose to be a leader across both message treatments. For Experiment 2, the female dummy is negative and significant at 5% when we do not control for risk preferences but loses significance once we do so.

<Table 5 about here>

We conclude this section by noting the following facts. First, regardless of whether leaders face a different payoff matrix or not, across both our experiments, a much larger proportion of men volunteer to lead. This is borne out by sample proportions tests and parametric probit regressions. This difference in the willingness to lead persists even after we control for risk preferences using the Holt-Laury lottery-choice task. The female reluctance to lead is much more pronounced when subjects know that the leaders' gender will be made known to the followers. When the leaders' gender is not revealed, the differences between the

gender is revealed, then we have 126 decisions by males and 124 by females. In this latter case, there is a significant gender difference in willingness to lead (|z|=5.37, p<0.01).

[13] As noted previously, we lose observations when we control for risk preferences due to excluding subjects who make inconsistent choices in this task.

proportion of men and women volunteering to lead are usually not significant. But when this information is made public, women exhibit greater reluctance to lead. A plausible conjecture regarding the greater female reluctance to lead is that female leaders anticipate greater resistance and backlash from followers as documented by Gangadharan, Jain, Maitra, and Vecci (2016). We now turn to exploring the extent to which this presumption of greater resistance to female leadership is borne out in the latter half of the game.

4.2 Perception and efficacy of leaders

In this part of the paper, our aim is to compare the performance of the groups led by male and female leaders, respectively. Are there systematic differences in the level of effort exerted, the degree of coordination success or earnings in male and female led groups? Therefore, here, we focus on the gender revealed treatments only. In Experiment 1 (where leader payoff changes), we have 17 groups with 11 male leaders and 6 female leaders in the pre-set message treatment, while we have 33 groups with 19 male and 14 female leaders in the free-form message treatment. In Experiment 2 (where leader payoff remains unchanged), we have 9 groups with 4 male leaders and 5 female leaders in the pre-set message treatment, and 8 groups with 2 male leaders and 6 female leaders in the free-form message treatment. This information is provided in Tables 3 and 4 above. Given that here the unit of observation is a group of five people, we will be dealing with a relatively small number of observations.

The dynamic nature of the process following the selection of a leader creates potential confounds since differences will arise endogenously. In an ideal setting, we want to hold everything other than the leader's gender constant. One way of doing this is to have an exogenously appointed leader and then look at follower responses. This is the approach adopted by Grossman et al. (2019). However, introducing an exogenous leader has an element of artificiality and may well raise questions about the authenticity of such a leader, who has no prior experience with the group or the task at hand. We felt that allowing endogenous emergence of leaders, who already have experience with the coordination problem, makes the setting more realistic. In any event, such endogenous appointment is the logical extension of the first part of our study, which allows participants to express a willingness to lead. This does create a trade-off between a degree of

realism in the design and the need for experimental control. As we explain below, we control for potential confounds as far as practicable and note relevant caveats along the way.

In the pre-set message treatment, we make sure that the message sent by the leaders is exactly the same. Given that this message is exhorting followers to choose the maximum possible effort level (forty hours), it stands to reason that leaders, when sending this message prior to Round 11, will also end up choosing the maximum possible effort level. Therefore, any differences that arise in the pre-set message treatment should be driven primarily by the leader's gender, since we expect no differences in the message sent by the leader and, few, if any, differences in the leader's effort choice in Round 11. Of course, over time, things may start to diverge; male and female leaders may choose different effort levels and also differ in the number of messages they send over the course of the second half of the game. However, we do know that behaviour in these weak-link games is extremely path dependent and so controlling for variations early on guarantees that the initial conditions are similar for the male and female led groups. If the leader's gender matters, then we would expect to see differences emerging in group performance even in this somewhat sanitized environment with tight control over the leader's actions and messages.[14]

On the other hand, in the free-form message treatment, we are allowing for much greater variation. Leaders here can differ along a number of dimensions: the content of their messages; their effort levels, especially if the message chosen asks followers to choose an effort level different from forty; how many times the message is sent, and how many words to use per message. But, if we find no differences in the pre-set message treatment, then we might be able to conclude that gender per se is not so important[15]. In

[14] In the pre-set message treatment, we had one problem. There is one (female) leader in this treatment who chose an effort level of 20 in spite of the fact that she was sending a message asking the followers to choose 40. To analyze the impact of a leader's gender on followers' effort level while holding effort level of 40 in round 11 and the content of the message constant, we exclude this one group where the leader did not choose 40 in round 11 for the rest of the analysis. This leaves us with 16 groups in pre-set message treatment with 11 male leaders and 5 female leaders in Experiment 1. Our result does not change if we include that group.

[15] Another explanation is that leaders are treated differently in the pre-set message treatment and the free-form message treatment. In the pre-set message treatment, both leaders and followers are aware that the message is fixed and provided by the experimenter. Leaders can only pass on the message provided to them by the experimenter and they only need to decide on whether to pass this information to their group members or not. On the other hand, in the free-form message treatment, leaders know that they are required to lead the team by using their own messages. Therefore, leaders in the free-form message treatment are more likely to be treated as an actual leader than those in the pre-set

that case, the free-form treatment allows us to focus on other sources of differences; for instance, whether there are systematic differences in message style/content or effort choices of the leaders. Further, as we explain below, we will control for message content, frequency and effort choice on the part of the leaders.

We start by showing the evolution of average effort levels in male and female led groups over the course of an entire session. Figures 6A and 6B are based on data from Experiment 1. They show what happens to effort levels over the course of the session, i.e., Round 1-20 for the pre-set and free-form message treatments, respectively. Figures 7A and 7B provide the same information for the pre-set and free-form message treatments respectively in Experiment 2. In each figure, the left panel shows the evolution of average effort levels over the first 10 rounds of the session, where there is no leader. The middle panel shows the average effort levels chosen by the leaders (male or female; at the expense of being labelled sexist, we have chosen to go with blue lines with diamonds for males and pink lines with squares for females) in Rounds 11-20. The right panel shows the average effort levels for male and female led groups (once again blue with diamonds for male led groups and pink with squares for female led groups) over the course of Rounds 11-20.

It is clear that there are no dramatic differences in average effort choices for either leader or followers. This is particularly true for Experiment 2 (with no change in leader payoff) where average leader and follower efforts are very similar. One can discern some differences in Experiment 1 (where leader payoff changes). There is a sharp downward spike in average leader effort in female led groups in the pre-set message treatment (middle panel of Figure 6A) but this does not seem to have had much of an adverse impact on average follower effort in those female led groups; if anything, average follower effort in female-led groups is marginally higher than that in male led groups (right panel of Figure 6A).

We can also see differences in the free-form message treatment of Experiment 1, where both average leader and follower effort appears to be lower in female led groups. We note two points before

message treatment. In the pre-set message treatment, the credibility of the message might be attributed to the experimenter and thus independent of the leader's gender.

proceeding to look into this in greater detail. First, when we look at the average effort level of leaders and/or groups, we are working with a small number of observations. Second, as noted previously, there is path-dependency in such weak-link games, implying that follower effort choices are affected crucially by leader effort choices. So, some of the dynamics here are being driven by the small number of leader observations.[16]

<Figures 6A and 6B about here>

<Figure 7A and 7B about here>

In order to explore any potential gender differences more rigorously, we next turn to regression analysis. In this part, we will look at two issues: first, are there differences in the effort levels of followers in male and female led groups? Second, we will explore whether there are differences in the earnings of male and female led groups. This will help establish the extent to which there are differences in the performance of male and female led groups and whether such differences, if any, result is differential earnings for these groups.

4.2.1 Follower effort levels in male and female led groups

From Figures 6 and 7, it seems clear that there are no dramatic differences in average leader /follower effort choices in Experiment 2 (where leader payoff does not change); differences,, if any, arise primarily in Experient 1 (where leader payoff is different). So, in this part of the paper, we will confine our attention to data from Experiment 1 alone. We present detailed analogous results for Experiment 2 in Appendix D.

In Table 6, we present results of random effects ordered probit regression to examine the impact of leader's gender on follower's effort choice for both treatments with standard errors clustered on groups.[17] The dependent variable is an individual's effort choice per round from rounds 11 to 20. The regressors

[16] We also look at whether, during the first set of 10 rounds, there are differences in the average effort levels of groups that went on to have a male or female leader. This is to ascertain that there were no systematic differences in the performance of these groups during the first part, which may have led to more (or fewer) female subjects volunteering to lead. So, effectively, we are asking whether these groups behaved differently prior to a male or female leader being appointed. We do not find evidence of any significant differences and therefore, refrain from elaborating any further on this aspect of the study.

[17] Given that there is a natural ordering of the effort choices ranging from zero to 40 in this game, an ordered probit regression is the appropriate way to model choices. Further, given that groups are fixed over time, it makes sense to cluster errors on groups rather than on individual subjects or at the session level.

include: Female = 1, if the follower is female and zero otherwise; Female Leader = 1, if the group leader is female and zero othertwise. The other variables are self-explanatory. For both treatments, the coefficient for round is negative and significant which is in line with what is apparent from Figure 6; that there is decay in effort choice over time. The coefficients for earnings in the previous round are postive and statistsically significant suggesting that the higher the earning followers receive in the previous round, the higher the likelihood of them choosing a higher effort in the following round, but the values of the coefficient are numerically small for both treatments. The coefficients for female leader are not significant for either treatment, which suggest that average follower efforts are not significantly different between male and female led groups. We present analogous results for Experiment 2 in Table A6 (and Table A6*) of Appendix D.

<Table 6 about here>

In Tables 7 and 8, we take a more disaggregated look at the issue of resistance to female leadership. In Table 7, we present the results from random effects ordered probit model for follower effort choices in the pre-set message treatment for Experiment 1. The dependent variable is the effort choice by follower *i* from group *g* in round *t*. The regressors include three dummies: male follower with female leader, female follower with male leader and female follower with female leader. The reference category is male follower with male leader. We also control for the group *g* leader's effort in round *t* (since the followers get to see the leader's effort choice prior to making their own choices), a dummy variable for whether the leader sent a message in round *t* or not and (given the path dependency in such games) the lagged minimum effort, i.e., minimum effort in round *t-1*.

We present results from three different specifications. The first one includes only the regressors without controlling for other demographics; the second one adds demographic controls, while the third and final specification adds both demographic controls as well controls for risk preference derived from the Holt-Laury lottery choice task. (As noted earlier, we lose observations when we control for risk preferences due to inconsistent choices made by a number of subjects, who get dropped.) Follower effort levels increase in leaders's effort and the lagged minimum effort, both of which facts make intuitive sense. The results do

not provide evidence of any backlash against female leadership. The main thing that stands out is that female followers exert lower effort than their male counter-parts and this is true both when the leader is female as well as when the leader is male. The relevant coefficients are significant at 5% in the second and third specifications. However, male followers do not behave differently regardless of whether the team leader is male or female. The results for Experiment 2 are similar and are reported in Table A7 of Appendix D.

<Table 7 about here>

In Table 8, we carry out a similar exercise for the free-form message treatment in Experiment 1 (where leader payoff changes). The regressors are similar to those in Table 7 except, here, we also control for the style of message sent by the leader. We classify messages into three different types: authoritarian, democratic and laissez-faire. We provide details on this classification exercise in Section 4.3 below. We provide four different specifications here: the first one includes only the main regressors of interest; the second one adds demographic controls, the third adds dummies for the message style (authoritarian/democratic/laissef-faire) and the fourth specification also controls for risk preference.

The results are similar and once again we do not see any evidence of resistance against female leadership. The only other fact that stands out (in the fourth specification) is that democratic messages from leaders seem to have a positive and significant impact on follower effort choice. We conclude this section by arguing that there is no systematic evidence of backlash against female leaders. Female followers choose lower effort levels in general but this is true for both male and female led groups. There are no differences in male follower effort levels regardless of whether the leader is male or female.[18]

<Table 8 about here>

4.2.2 Earnings in male and female groups

Figures 5 and 6 as well as the regression results presented in Tables 7 and 8 suggest that there are no systematic differences in the performance of male and female led groups in terms of average effort. This,

[18] The results for the same exercise for Experiment 2 are reported in Table A8 of Appendix D.

in turn, implies that we would not expect to see differences in the average earnings of male and female led groups as well. We present results for earnings from Experiment 1, in Table 9. As noted above, we expect differences, if any, to arise in Experiment 1 rather than Experiment 2, since the latter shows no difference in average effort levels at all. For the sake of completeness, we also provide the earnings regression for Experiment 2 in the Appendix D. The coefficient for the female leader dummy is not significant in any of the treatments, whether pre-set message or free-form; neither is it significant if we control for risk preference or not. We conclude that there is no evidence to suggest that male and female led groups differed in terms of their average earnings.[19]

<Table 9 about here>

4.3. Leadership style and message content in free-form message treatments

We rely on the Northouse (2012) methodology to identify the types of messages sent by the leaders. Bear in mind that the leaders have two choices for each round following Round 11; whether to send a message or not and if yes, then what type of message to send. We classify messages into four categories. "0" indicates that leaders did not send any message to their group members. "1" stands for an "authoritarian" message style, "2" for democratic and "3" for laissez faire.[20] Figure 8, Panels A and B show the distribution of message choices in Experiments 1 and 2. By and large there are no substantive differences here between the two genders; at most men showed a slight preference for authoritarian messages while women preferred democratic messages. But we fail to find any significant differences.

<Figure 8 about here>

[19] The results for the same exercise for Experiment 2 are reported in Table A9 of Appendix D.

[20] An "authoritarian" message is one where leaders give orders and/or demand compliance from group members. Examples include: "*I think I've made myself clear. All choose 40 and we make the most money!*", and "*All pick 40 hours for the remaining ten rounds for maximum pay*". Style 2 is defined as "democratic", where leaders exhort followers and make suggestions/requests. Following are examples: "*Hey guys, to make the most money let's all select 40 hours. That way you and I will both earn the most profits, we all get the optimal outcome ;)",* and *"Hi guys! If you choose more than 20 that would be great!".* We categorize Style 3 as laissez faire message style where messages have no substantive content. E.g., "*You know what to do. ;) I believe in you.*", and "*McDonalds tonight!*". We use three independent coders to undertake this coding. A subject is assigned to a particular category as long as two out of the three coders assigned the same score to this subject.

In Table 10, we show the relative frequencies of messages sent by leaders. While there are differences here, there do not seem to be any obvious patterns. It is also the case that these are repeated observations for the same leaders rather than independent observations and for Experiment 2, we have very few observations. If we look at the average number of words in the messages sent then women, on average, wrote 9.1 words (standard error of 3.3) while men wrote 9.6 words (standard error of 4.39). Once again there do not seem to be any significant differences in the style or frequency of messages sent by our leaders.

<Table 10 about here>

5. Conclusion

In this study, we examine whether there is a gender difference in the willingness to lead in a weak-link game designed to simulate frequently occurring intra-organization coordination problems. Overall, we find that fewer women volunteer to be the leader compared to men. This decision depends crucially on whether the leader's gender is revealed to group members or not. When the leader's gender is revealed to followers, fewer women choose to volunteer; when it is not, the willingness gap among men and women shrinks.

We do not find significant differences in the degree of coordination success (or failure) between male and female led groups. Neither do we find much difference between effort levels chosen or the content of messages sent by male and female leaders. While we do not find significant differences in followers' perception towards male and female leaders, we do find that women are less likely volunteer for leadership positions. As noted above, a plausible conjecture is that female subjects shied away from volunteering because they anticipated resistance from followers. But this resistance did not materialize in our study.

It is entirely possible that the lack of resistance toward female leaders is due to the fact that this study was carried out in New Zealand; the first country to endow women with the right to vote back in 1893. New Zealand also has a long history of strong female leaders including three female Prime Ministers in recent times and as noted above, the gender wage gap in New Zealand is small and lower than the OECD average. But, while this may explain why we do not find evidence of backlash against female leadership, it is harder to explain why female subjects are so much more reluctant to volunteer for leadership and why this reluctance is exacerbated in the gender revealed treatments. It seems difficult to explain this in any way

other than to appeal to internalized norms whereby women anticipate greater resistance based on their socialization and lived experiences. So, a potential lesson of this study is that the gender gap in leadership may arise due to a greater female reluctance to assume leadership roles but to the extent that these results may extend to other countries and cultures, it is possible that female leaders are over-estimating the degree of resistance to female leadership.

References

Altonji, J.G., and Blank, R.M. 1999. *Race and gender in the labor market.* Vol. 3, in *Handbook of Labor Economics*, by Orley Ashenfelter and David Card, 3143-3259. Amsterdam: Elsevier.

Andersen, S., Ertac, S., Gneezy, U., List, J.A., and Maximiano, S. 2013. "Gender, Competitiveness, and Socialization at a Young Age: Evidence From a Matrilineal and a Patriarchal Society." *Review of Economics and Statistics* 95 (4): 1438-1443.

Babcock, L., and Laschever, S. 2003. *Women don't ask.* Princeton, NJ: Princeton University Press.

Babcock, L., Pecalde, M.P., Vesterlund, L., and Weingart, L. 2017. "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability." *American Economic Review* 107(3): 714-47.

Bertrand, M. 2011. *New Perspectives on Gender.* Edited by Orley C Ashenfelter and David Card. Vol. 4b. Amsterdam: Elsevier.

Bertrand, M., Goldin, C., and Katz, L.F., 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2: 228-55.

Blau, F. D., and Kahn, L.M. 2000. "Gender Differences in Pay." *Journal of Economic Perspectives* 14: 75-99.

Blau, F. D., and Kahn, L.M. 2006. "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence." *Industrial and Labor Relations Review* 60: 45-66.

Blau, F. D., and Kahn, L.M. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3): 789-865.

Bohnet, I. 2016. *What works: gender equality by design.* Cambridge: Harward University Press.

Bohnet, I., van Green, A., and Bazerman, M. (2016). When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Management Science, 62*(5), 1225-1234.

Brandts, J., and Cooper, D.J. 2006. "A Change Would Do You Good ... An Experimental Study on How to Overcome Coordination Failure in Organizations." *American Economic Association* 96 (3): 669-693.

Brandts, J., and Cooper, D.J. 2007. "It's What You Say, Not What You Pay: An Experimental Study of Manager-Employee Relationships in Overcoming Coordination Failure." *Journal of the European Economic Association* 5 (6): 1223-1268.

Brown, A. 2014. *Perceptions about women bosses improve, but gap remains.* 7 August. http://www.pewresearch.org/fact-tank/2014/08/07/perceptions-about-women-leaders-improve-but-gap-remains/.

Burns, J M. 1978. *Leadership.* New York: Harper and Row.

Carter, N. M., and Silva, C. 2010. *Pipeline's Broken Promise.* 26 May. https://www.catalyst.org/wp-content/uploads/2019/01/Pipelines_Broken_Promise_Final_021710.pdf.

Catalyst. 2018. *Women in Management.* 30 July. http://www.catalyst.org/knowledge/women-management#footnote22_g8owtq6.

Charness, G., and Gneezy, U. 2012. "Strong Evidence for Gender Differences in Risk Taking." *Journal of Economic Behavior and Organization* 83 (1): 50-58.

Chaudhuri, A., Cruickshank, A. and Sbai, E. (2015). Gender differences in personnel management: experimental evidence. *Journal of Behavioral and Experimental Economics*, 58, October 2015, 20-32.

Chaudhuri, A., Schotter, A., and Sopher, B. (2009). Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Effort Games with Private, Almost Common and Common Knowledge of Advice. *Economic Journal*, *119*(534), 91-122.

Coffman, K. (2014). Gender Differences in Willingness to Guess. *Management Science, 60*(2), 434-448.

Croson, R., and Gneezy, U. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2): 448-474.

Drolet, M. 2001. *The Persistent Gap: New Evidence on the Canadian Gender.* Ottawa, Ontario: Statistics Canada.

Eagly, A. H., and Johnson, B.T. 1990. "Gender and Leadership Style: A Meta-Analysis." *Psychological Bulletin* 108 (2): 233-256.

Eagly, A. H., Karau, S.J., and Makhijani, M.G. 1995. "Gender and the Effectivness of Leaders: A Meta-Analysis." *Psychological Bulletin* 117 (1): 125-145.

Eckel, C. C., Gangadharan, L., Grossman, P.J. and Xue, N. (2020). The Gender Leadership Gap: Insights from Experiments in Chaudhuri, A. (Ed.) Research Agenda in Experimental Economics (Chapter 7), Cheltenham, UK: Edward Elgar (Forthcoming)

Eckel, C. C., and Grossman, P.J. 2008. "Differences in the Economic Decisions of Men and Women: Experimental Evidence." *Handbook of Experimental Economics Results* (Elsevier) 1: 509-519.

Erkal, N., Gangadharan, L., and Xiao, E. 2019. Leadership selection: Can changing the default break the glass ceiling? Monash University Working Paper.

Exley, C. L., Niederle, M., and Vesterlund, L. 2020. "Knowing When to Ask: The Cost of Leaning in." *Journal of Political Economy* 128 (3):816-854.

Fischbacher, U. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10 (2): 171-178.

Gangadharan, L., Tarun J., Maitra, P., and Vecci, J. 2016. “Social identity and governance: The behavioral response to female leaders.” *European Economic Review* 90: 302-325.

Gneezy, U., Leonard, K.L., and List, J.A. 2009. “Gender Differences in Competition: Evidence From a Matrilineal and Patriarchal Society.” *Econometrica* 77 (5): 1637-1664.

Gneezy, U., Niederle, M., and Rustichini, A. 2003. “Performance in Competitive Environments: Gender Differences.” *The Quarterly Journal of Economics* 118 (3): 1049-1074.

Goldin, C. 2014. “A Grand Gender Convergence: Its Last Chapter.” *American Economic Review* 104 (4): 1091-1119. https://doi.org/10.1257/aer.104.4.1091.

Goldin, C., and Rouse, C. 2000. “Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians.” *The American Economic Review* 715-741.

Goldin, C., Katz, L.F., and Kuziemko, I. 2006. “The Homecoming of American College Women: The Reversal of the College Gender Gap.” *Journal of Economic Perspectives* 20 (4): 133-56.

Grossman, P. J., Eckel, C., Komai, M., and Zhan, W. 2019. “It pays to be a man: Rewards for leaders n a coordination game.” *Journal of Economic Behavior and Organzation* 197-215.

Grossman, P. J., Komai, M., and Jensen, J.E. 2015. “Leadership and gender in groups: An experiment.” *Canadian Journal of Economics* 48: 368-388.

Holt, C. A., and Laury, S.K. 2002. “Risk Aversion and Incentive Effects.” *The American Economic Review* 92 (5): 1644-1655.

Jena, A. B., Olenski, A.R., and Blumenthal, D. M. 2016. “Sex differences in physician salary in U.S. public medical schools.” *JAMA Intern Medicine* 176 (9): 1294-1304. doi:10.1001/jamainternmed.2016.3284.

Knez, M., and Camerer, C. 1994. “Creating Expectational Assets in the Laboratory: Coordination in 'Weakest-Link' Games.” *Strategic Management Journal* 15: 101-119.

Knez, M., and Camerer, C. 2000. “Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination Games.” *Organizatinal Behavior and Human Decision Processes* 82 (2): 194-216.

Li, Y., Sbai, E., and Chaudhuri, A. (2020). An experimental study of gender differences in agency relationships. *Journal of Behavioral and Experimental Economics*. (Forthcoming)

MacNell, L., Adam D., and Hunt, A.N. 2015. “What's in a Name: Exposing Gender Bias in Student Ratings of Teaching.” *Innov High Educ* 40: 291-303.

Moran, B. B. 1992. “Gender Differences in Leadership.” *Library Trends* 40: 475-491.

Niederle, M., and Vesterlund, L. 2007. “Do Women Shy Away From Competition? Do Men Compete Too Much?” *The Quarterly Journal of Economics* 122 (3): 1067-1101.

Northouse , P. G. 2012. *Introduction to Leadership: Concepts and Practice.* Los Angeles: SAGE.

Reuben, E., and Timko, K. 2018. "On the effectiveness of elected male and female leaders and team coordination." *Journal of the Economic Science Association* 4 (2): 123-135.

Ridgeway, C. L. 2001. "Gender, Status, and Leadership." *Journal of Social Issues* 57 (4): 637-655.

Rosener , J. 1990. *Ways women lead.* Harvard Business Review.

Sandberg, S. 2013. *Lean in.* New York: Random House USA Inc.

Sax, L. 2005. *Why gender matters.* New York: Doubleday.

Van Huyck, J. B., Battalio, R.C. and Beil, R.O. 1990. "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *The American Economic Review* 80 (1): 234-248.

Vandegrift, D., and Yavas, A. 2009. "Men, women, and competition: An experimental test of behavior." *Journal of Economic Behavior and Organization* 72 (1): 554-570.

Walters, A. E., Stuhlmacher, A.F., and Meyer, L.L. 1998. "Gender and Negotiator Competitiveness: A Meta-analysis." *Organizational Behavior and Human Decision Processes* 76: 1-29.

Weichselbaumer, D., and Winter-Ebmer, R. 2005. "A Meta-Analysis of the International Gender Wage Gap." *Journal of Economic Surveys* 19: 479-511.

Wood, R. G., Corcoran, M.E., and Courant, O.N. 1993. "Pay Differences among the Highly Paid: The Male-Female Earnings Gap in Lawyer's Salaries." *Journal of Labor Economics* 11: 417-441.

Zizzo, D. J. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics* 13: 75-98.

Table 1: Payoff matrix for (i) first 10 rounds of Experiment 1 and (ii) all rounds of Experiment 2

		Minimum of hours chosen by other members of the group				
		0	10	20	30	40
Choice of hours by a particular player	0	400	400	400	400	400
	10	350	450	450	450	450
	20	300	400	500	500	500
	30	250	350	450	550	550
	40	200	300	400	500	600

Note: This payoff matrix shows the payoff to all players for the first 10 rounds of Experiment 1 (before leaders are selected) and for all 20 rounds of Experiment 2 (both before and after leader selection)

Table 2: Payoff matrix for leaders for rounds 11 through 20 in Experiment 1

		Minimum of hours chosen by other members of the group				
		0	10	20	30	40
Leader's choice of hours	0	100	100	100	100	100
	10	80	300	300	300	300
	20	60	280	500	500	500
	30	40	260	480	700	700
	40	20	240	460	680	900

Note: In Experiment 1, all players face the payoff matrix shown in Table 1 for rounds 1 through 10 before selection of leaders. For rounds 11 – 20, the followers continue to play with that same payoff matrix, while the payoff matrix for the leaders changes to the one shown in Table 2.

Table 3: Number of subjects in different treatments in Experiment 1

	Pre-set message; Gender not revealed	Pre-set message; Gender revealed	Free-form message; Gender not revealed	Free-form message; Gender revealed	Total
Male	21	40	18	86	165
Female	24	45	17	79	165
Total	45	85	35	165	330
Total 5-person Groups	9	17	7	33	66

Table 4: Number of subjects in different treatments in Experiment 2

	Pre-set message; Gender revealed	Free-form message; Gender revealed	Total
Male	24	17	41
Female	21	23	44
Total	45	40	85
Total Groups	9	8	17

Note: In Experiment 2, we only carry out the gender revealed treatments.

Table 5: Probit Model for leadership choice prior to round 11

	Experiment 1		**Experiment 2**	
Choice	**Model 1A**	**Model 1B**	**Model 2A**	**Model 2B**
Female	-0.261***	-0.231***	-0.223**	-0.120
	(0.046)	(0.055)	(0.089)	(0.103)
Average effort round 1-10	0.007***	0.011***	0.009	0.010
	(0.003)	(0.003)	(0.006)	(0.006)
Average earnings round 1-10	-0.001	-0.001**	-0.002***	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Risk Averse	-	0.018	-	0.032
	-	(0.059)	-	(0.105)
Constant	0.640**	0.657**	0.694***	0.764***
	(0.024)	(0.028)	(1.1045)	(0.049)
Demographic Control	Yes	Yes	Yes	Yes
pseudo R^2	0.1127	0.1073	0.157	0.178
Wald χ^2	48.76	33.29	16.35	11.54
Prob>χ^2	0.000	0.000	0.038	0.2407
Number of Observations	330	241	85	59

*Notes: Standard errors in parentheses; ***, ** and * represent significance at 1% , 5% and 10% respectively.*

Table 6: Random effects ordered probit model for follower's effort in Experiment 1 (errors clustered on the group)

Dependent variable: Choice of effort level by follower

Effort Choice	**Pre-set**	**Pre-set**	**Free-form**	**Free-form**
Round	-0.310***	-0.314***	-0.282***	-0.281***
	(0.061)	(0.061)	(0.060)	(0.060)
Female	-2.533*	-2.706**	-1.493	-1.472
	(1.316)	(1.275)	(1.603)	(1.545)
Female*Round	0.147**	0.152***	0.090	0.088
	(0.060)	(0.059)	(0.092)	(0.091)
Female Leader	-0.343	-0.440	-1.047	-1.016
	(0.733)	(0.772)	(0.720)	(0.759)
Lag Earning	0.005***	0.004***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Risk Averse	-0.797	-0.764	1.033***	0.953**
	(0.681)	(0.726)	(0.397)	(0.454)
Demographic Control	NO	YES	NO	YES
Number of observations	470	470	990	990
Wald χ^2	82.29	413.13	66.33	482.20
Prob > χ^2	0.000	0.000	0.000	0.000

*Notes: Standard errors in parentheses; ***, ** and * represent significance at 1% , 5% and 10% respectively.*

Table 7: Random Effects Ordered Probit model for effort level under pre-set message treatment in Experiment 1 (errors clustered on the group)

	Model 1	Model 2	Model 3
Round	-0.238***	-0.240***	-0.287***
	(0.041)	(0.042)	(0.042)
Female*Round	0.089*	0.092*	0.155***
	(0.048)	(0.049)	(0.054)
Male follower with female leader	0.298	0.264	-0.137
	(0.580)	(0.633)	(0.704)
Female follower with male leader	-1.423	-1.610	-2.658**
	(1.051)	(1.092)	(1.091)
Female follower with female leader	-1.802*	-1.951*	-2.864**
	(1.051)	(1.183)	(1.161)
Leaders effort	0.057***	0.057***	0.063***
	(0.020)	(0.020)	(0.024)
Message shown	-0.097	-0.099	-0.139
	(0.156)	(0.154)	(0.251)
Lag minimum effort	0.036***	0.035***	0.040***
	(0.008)	(0.007)	(0.007)
Risk averse	-	-	-0.470
	-	-	(0.492)
Constant 1	-3.619***	-3.800***	-4.546***
	(1.010)	(1.071)	(0.999)
Constant 2	-3.562***	-3.742***	-4.464***
	(0.993)	(1.060)	(0.970)
Constant 3	-2.736**	-2.913**	-4.006***
	(1.146)	(1.226)	(1.039)
Constant 4	-1.463	-1.639	-2.719**
	(1.101)	(1.188)	(1.166)
Log pseudolikelihood	-416.522	-415.097	-250.541
Demographic Controls	No	YES	YES
Wald χ^2 Prob > χ^2	124.83 0.000	5313.37 0.000	2018.53 0.000
Number of observations	640	640	470
Wald test for equality of coefficients			
Male follower with female leader = Female follower with male leader	χ^2=2.85 p>χ^2=0.09	χ^2=3.33 p>χ^2=0.07	χ^2=4.81 p> χ^2=0.03
Male follower with female leader = Female follower with female leader	χ^2=5.01 p>χ^2=0.03	χ^2=5.17 p>χ^2=0.02	χ^2=5.94 p>χ^2=0.01
Female follower with male leader = Female follower with female leader	χ^2=0.47 p>χ^2=0.49	χ^2=0.30 p>χ^2=0.58	χ^2=0.09 p>χ^2=0.76

Table 8: Random Effects Ordered Probit model for effort level under Free-form message treatment in Experiment 1 (errors clustered on the group)

Follower effort	Model 1	Model 2	Model 3	Model 4
Round	-0.169***	-0.165***	-0.165***	-0.181***
	(0.046)	(0.045)	(0.045)	(0.036)
Female*Round	0.058	0.053	0.055	0.025
	(0.061)	(0.061)	(0.061)	(0.081)
Male follower with female leader	-0.485	-0.480	-0.638	-0.696
	(0.469)	(0.465)	(0.485)	(0.453)
Female follower with male leader	-0.915	-0.703	-0.835	-0.736
	(1.040)	(1.017)	(1.027)	(1.360)
Female follower with female leader	-1.219	-0.981	-1.180	-0.920
	(0.992)	(0.963)	(1.006)	(1.235)
Leaders Effort	0.032***	0.033***	0.032***	0.037***
	(0.009)	(0.009)	(0.009)	(0.009)
Authoritarian Message Style	-	-	0.576**	0.536
	-	-	(0.261)	(0.346)
Democratic Message Style	-	-	0.526**	0.576**
	-	-	(0.223)	(0.243)
Laissez-Faire Message Style	-	-	0.694	0.745
	-	-	(0.453)	(0.467)
Lag minimum effort	0.036***	0.036***	0.036***	0.041***
	(0.008)	(0.009)	(0.008)	(0.010)
Risk averse	-	-	-	0.645**
	-	-	-	(0.306)
Constant 1	-3.642***	-3.726***	-3.570***	-3.365***
	(0.831)	(0.917)	(0.927)	(1.071)
Constant 2	-3.422***	-3.505***	-3.346***	-3.104***
	(0.842)	(0.934)	(0.946)	(1.087)
Constant 3	-2.857***	-2.941***	-2.774***	-2.413**
	(0.913)	(0.988)	(0.995)	(1.199)
Constant 4	-2.421***	-2.508***	-2.337**	-1.949
	(0.892)	(0.966)	(0.971)	(1.188)
Log pseudolikelihood	-551.775	-546.307	-542.730	-397.170
Demographic Controls	NO	YES	YES	YES
Wald χ^2	94.78	645.244	726.722	1181.628
Prob > χ^2	0.000	0.000	0.000	0.000
Number of observations	1320	1320	1320	990
Wald test for equality of coefficients				
Male follower with female leader = Female follower with male leader	χ^2=0.13 p=0.71	χ^2=0.04 p=0.84	χ^2=0.03 p=0.86	χ^2=0.00 p=0.98

Male follower with female leader = Female follower with female leader	χ^2=0.47 p=0.49	χ^2=0.26 p=0.61	χ^2=0.32 p=0.57	χ^2=0.03 p=0.86
Female follower with male leader = Female follower with female leader	χ^2=0.44 p=0.51	χ^2=0.41 p=0.52	χ^2=0.68 p=0.41	χ^2=0.13 p=0.72
Authoritarian Message Style = Democratic Message Style	---	---	χ^2=0.04 p=0.85	χ^2=0.02 p=0.90
Authoritarian Message Style = Laissez-Faire Message Style	---	---	χ^2=0.06 p=0.81	χ^2=0.11 p=0.74
Democratic Message Style = Laissez-Faire Message Style	---	---	χ^2=0.14 p=0.71	χ^2=0.09 P=0.77

Table 9: Random Effects Model for follower's round earnings in Experiment 1 (errors clustered on the group)

Earnings	Pre-set	Pre-set	Free-form	Free-form
Round	-11.136***	-8.767***	-8.441***	-8.184***
	(3.421)	(2.628)	(2.607)	(2.821)
Female	-61.546*	-2.644	-22.629	-29.332
	(33.223)	(36.178)	(32.783)	(37.469)
Female Leader	-12.469	-16.237	-9.852	-11.777
	(13.366)	(15.259)	(14.252)	(14.189)
Female*Round	3.995*	0.087	1.627	2.123
	(2.385)	(2.506)	(2.097)	(2.407)
Leader's Effort	1.205	1.195	2.154***	1.539**
	(0.934)	(0.975)	(0.710)	(0.727)
Lag Minimum Effort	3.768***	3.685***	3.230***	3.484***
	(0.427)	(0.509)	(0.623)	(0.561)
Risk Averse	-	0.405	-	7.590
	-	(8.303)	-	(6.804)
Constant	551.393***	521.925***	497.061***	504.101***
	(46.791)	(34.761)	(47.028)	(51.153)
Wald χ^2	321.314	707.255	251.730	511.960
Prob > χ^2	0.000	0.000	0.000	0.000
Number of observations	640	470	1320	990

Table 10: Frequency of messages sent by leaders

		Pre-set message treatment	Free-form message treatment
	Female leader	0.73 (f=10)	0.95 (f=12)
Experiment 1	Male leader	0.69 (m=16)	0.79 (m=13)
Sample proportions test		*z=0.55* *p=0.58*	*z=-3.68* *p<0.01*
Experiment 2	Female leader	0.98 (f=5)	0.78 (f=6)
	Male leader	0.78 (m=4)	1 (m=2)

Notes: Leaders have to send a message in Round 11 but from that point onward, leaders can choose whether to send a message or not. If the leader chooses to send a message in all 10 rounds, then this is denoted as 100%. This is normalized to 1 in the table. We average over all leaders in a particular treatment. In the case of Experiment 1, within the pre-set and free-form message treatments, we combine the gender revealed and gender not revealed treatments. E.g. 0.73 implies that in that particular treatment, female leaders sent an average of 7.3 messages out of the maximum of 10 messages possible. Given the very small number of observations in Experiment 2, we have dispensed with formal statistical tests.

Figure 1: Screenshot for Leader under Free-form message treatment

Figure 2A: Screenshot for followers in Free-form message <u>gender not revealed</u> treatment

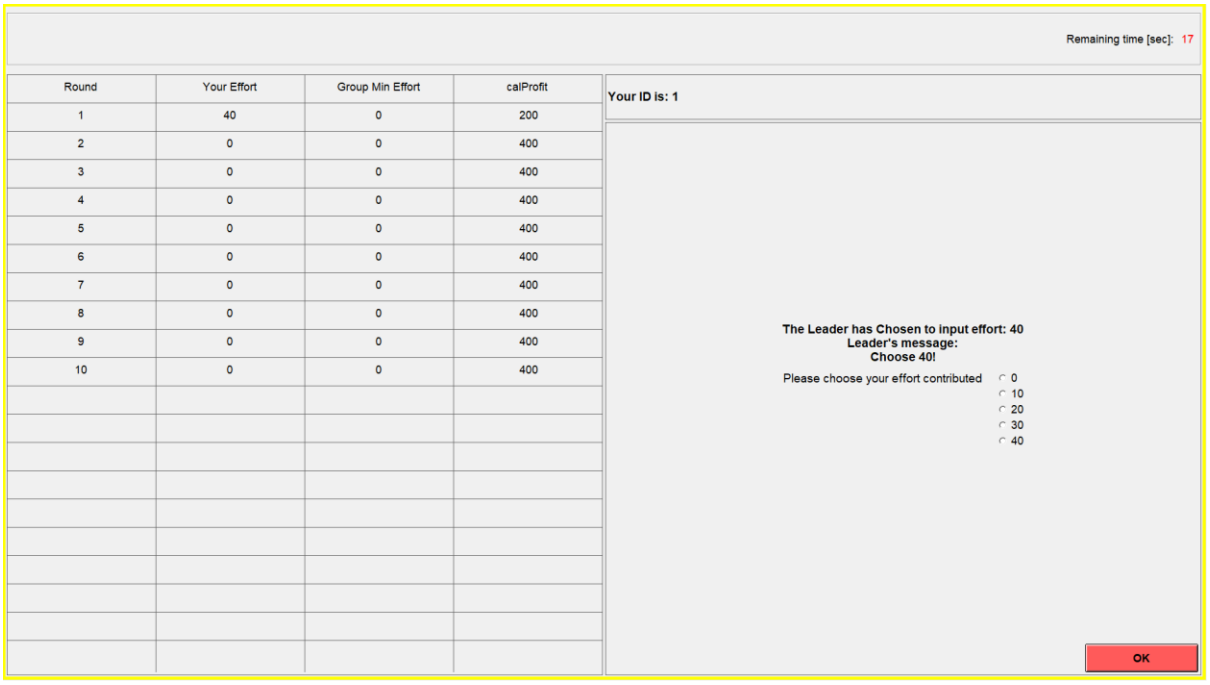


Figure 2B: Screenshot for team members in Free-form message <u>gender revealed</u> Treatment

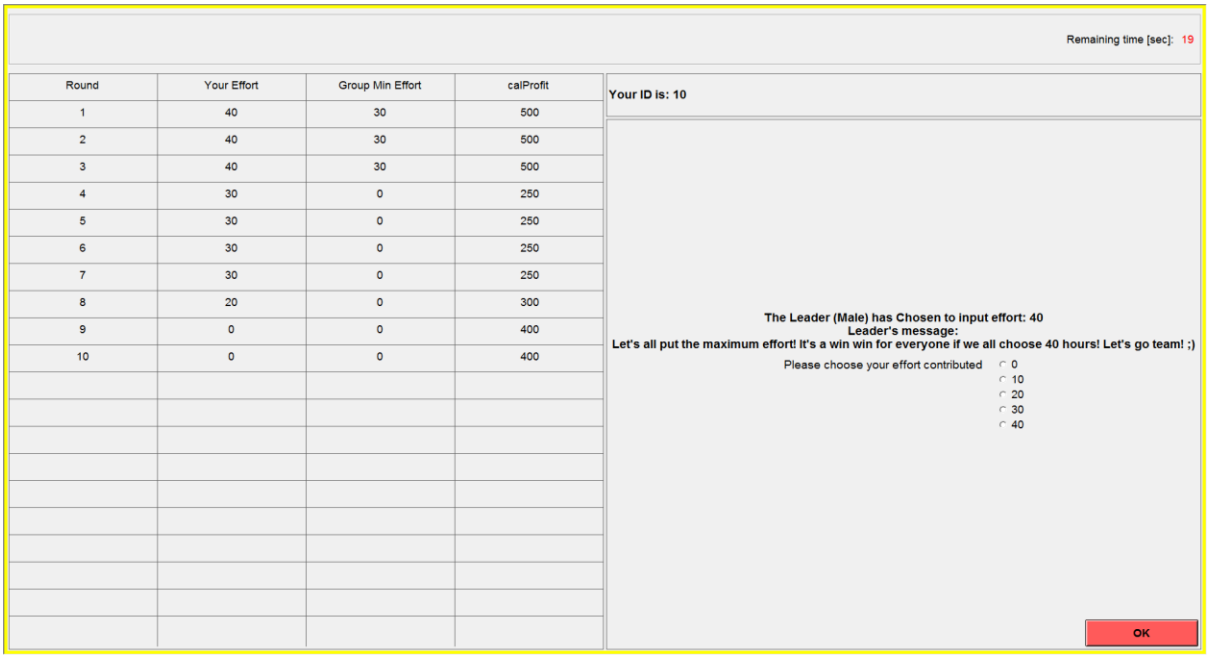
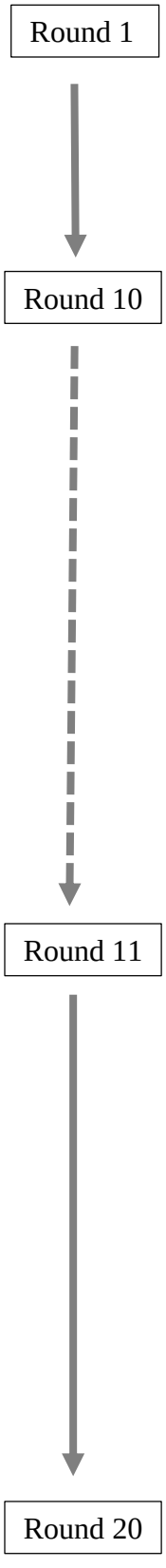


Figure 3: Timeline of Experiments 1 and 2



Prior to Round 1: Instructions for Round 1-10; on-screen and read out loud.

Rounds 1 to 10: Participants play corporate turnaround game with same payoff matrix for everyone as shown in Table 1.

(1) Instructions for Round 11-20; on-screen and read out loud.

In *Experiment 1, leader's payoff changes; In Experiment 2, leader's payoff unchanged*; follower payoffs unchanged in both experiments; this is common knowledge.

(2) Participants asked if they wish to volunteer as leader or not.

Once all decisions entered, one participant randomly selected as leader.

(3) Participants learn outcome: leaders get message stating he/she has been chosen as leader.

In gender not revealed treatment, followers get message: You are not the leader. The leader is Subject ID X.

In gender revealed treatment, follower gets message: You are not the leader. The leader is Subject ID X (Male/Female)

Round 11 starts:

(1) Leader enters effort choice and sends message; Depending on treatment message is either pre-set or free-form. If pre-set, then message available on-screen. Leader only has to click button next to message. If free-form, then leader has to write message. In Round 11, leader has to either click button or write message as the case may be.

(2) Followers get to see leader's effort choice, the leader's message and in the gender revealed treatment, the leader's gender: male/female.

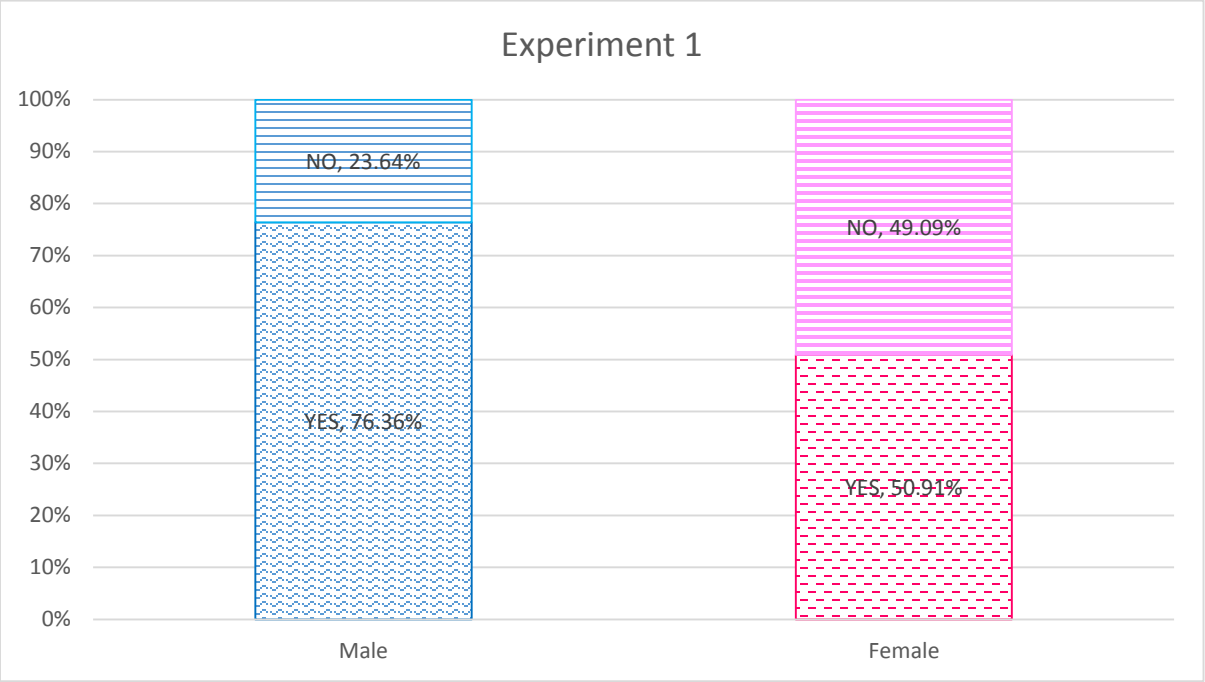
(3) Followers enter their effort choice.

(4) Everyone gets to see outcome of Round 11.

Rounds 12 through 20 proceed in the same way. In Experiment 1, payoff matrix for leader different from that of followers as shown in Table 2. In Experiment 2 payoff matrix same for both leaders and followers as shown in Table 1.

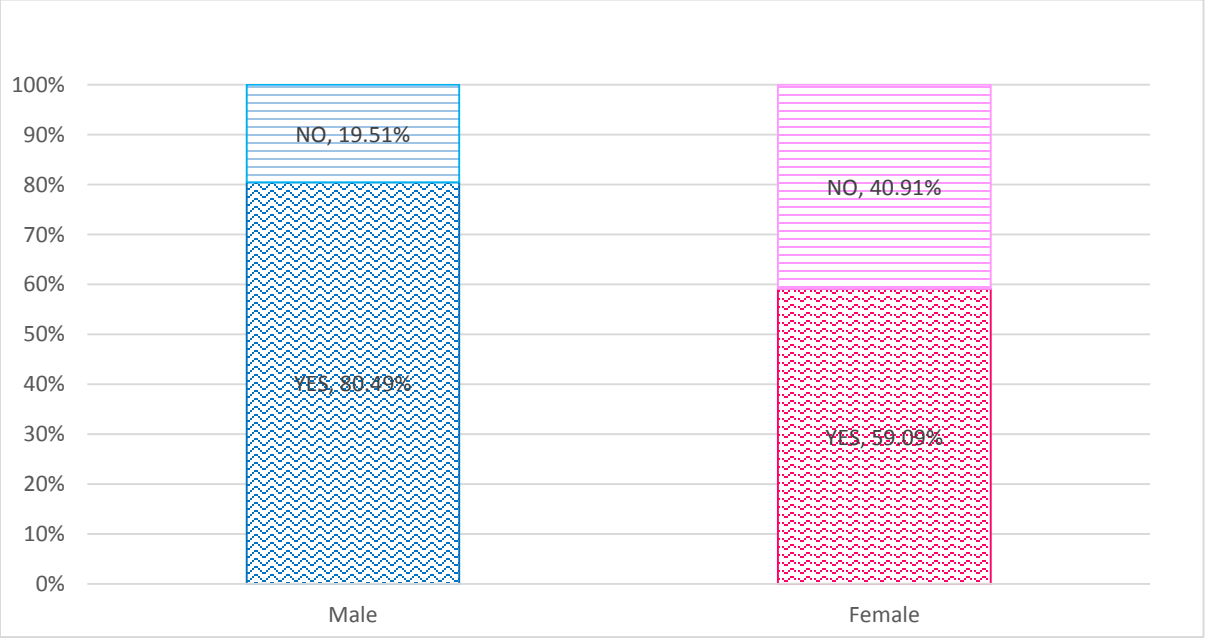
Leader free to send or not send message in any round.

Figure 4A: Difference in the willingness to lead in Experiment 1



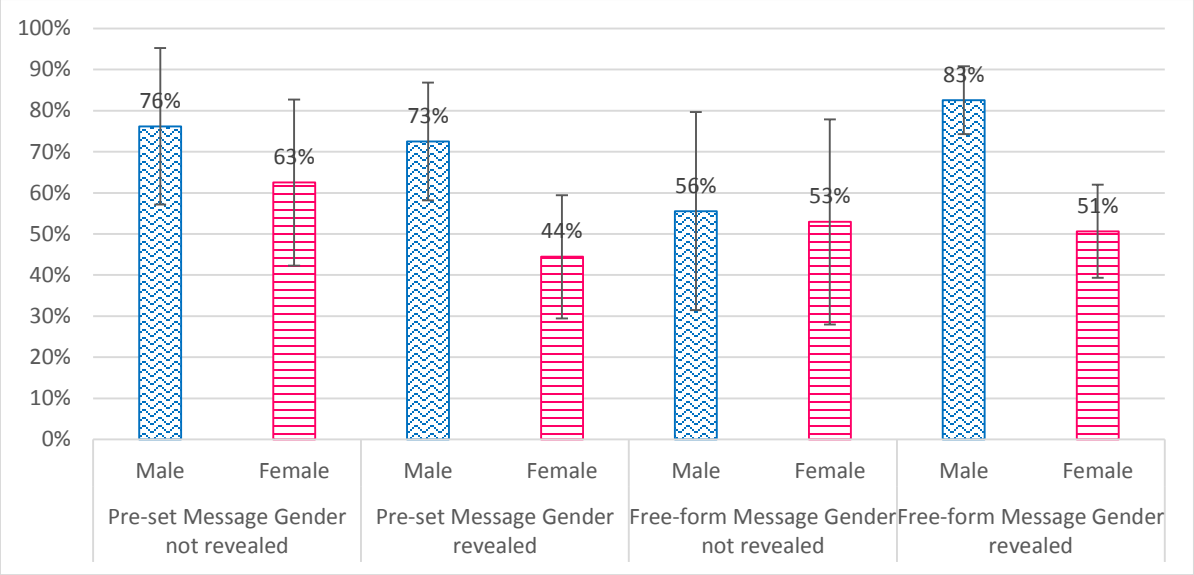
Note: Data for 165 males and 165 females from Experiment 1; simple count of how many men and women volunteered to lead prior to Round 11

Figure 4B: Difference in the willingness to lead in Experiment 2



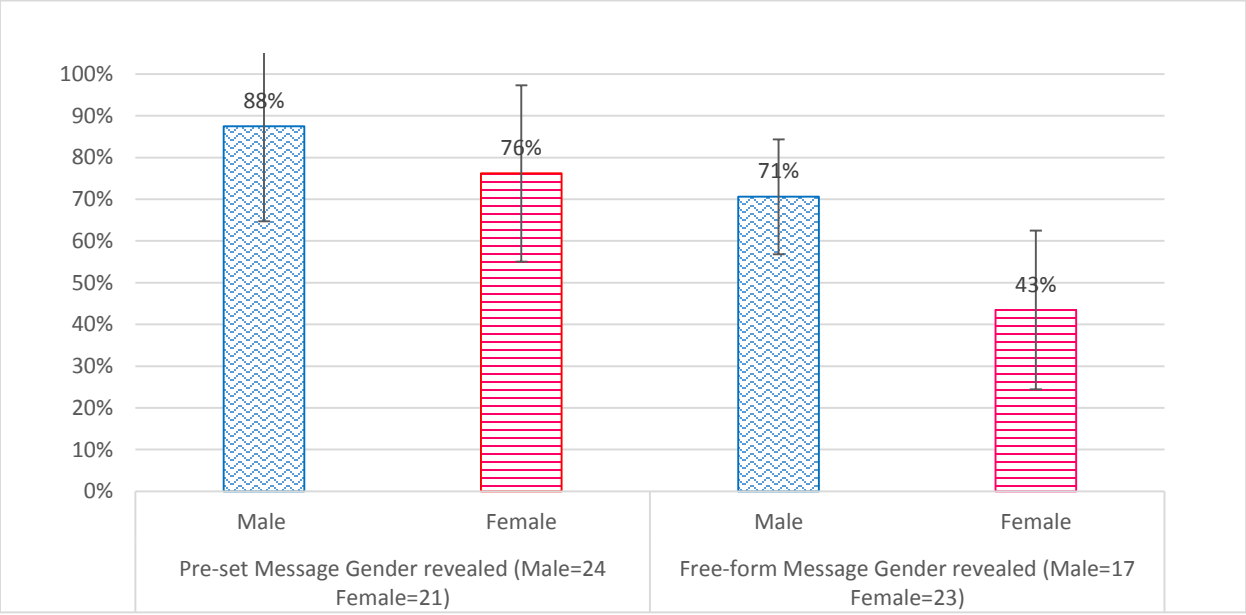
Note: Data for 41 males and 44 females from Experiment 2; simple count of how many men and women volunteered to lead prior to Round 11

Figure 5A: Difference in willingness to lead broken up by treatments (gender revealed vs. not revealed; pre-set message vs. free-form) in Experiment 1



Note: Pre-set message gender not revealed: 21 males and 24 females; pre-set message gender revealed 40 males and 45 females; free-form message gender not revealed: 18 males and 17 females; pre-set message gender revealed: 87 males and 79 females.

Figure 5B: Difference in willingness to lead broken up by treatments (pre-set message vs. free-form) in Experiment 2



Notes: In Experiment 2, there is no gender not revealed treatment since we undertake only gender revealed treatments.

Figure 6A: Average effort choice in Rounds 11-20 in pre-set message treatment (Experiment 1)

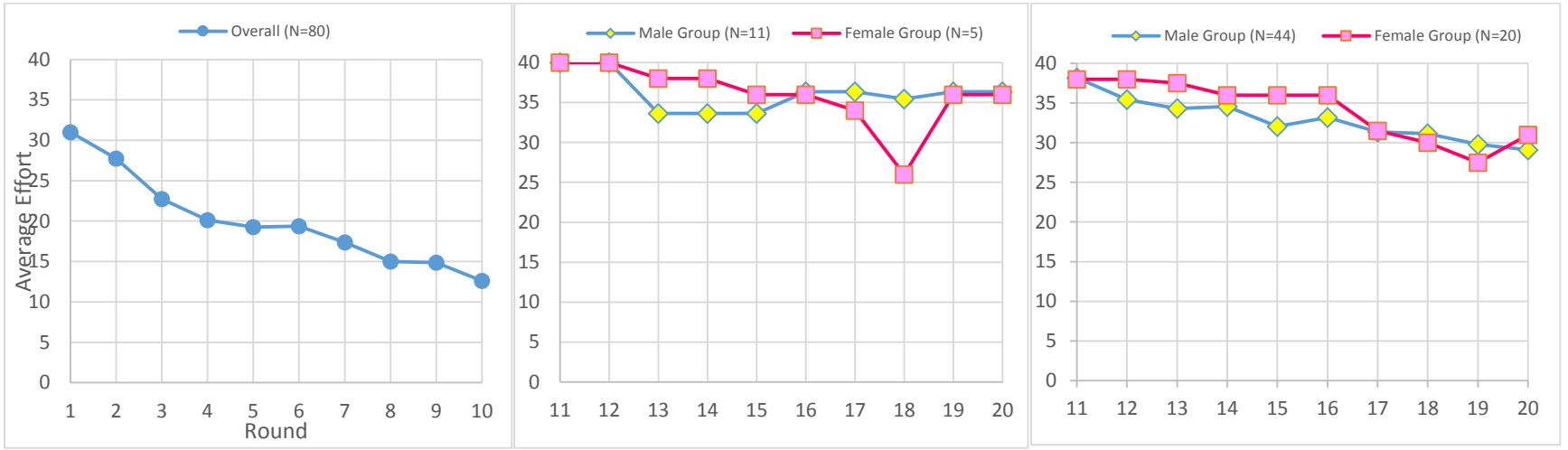


Figure 6B: Average effort choice Rounds 11-20 in free-form message treatment (Experiment 1)

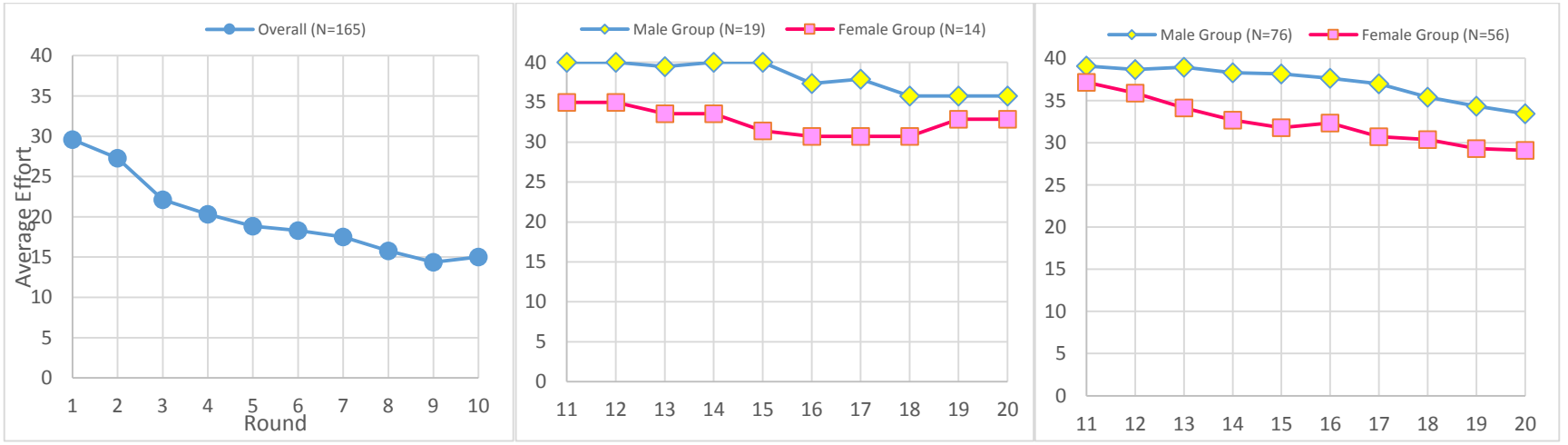


Figure 7A: Average effort choice in Rounds 11-20 in pre-set message treatment (Experiment 2)

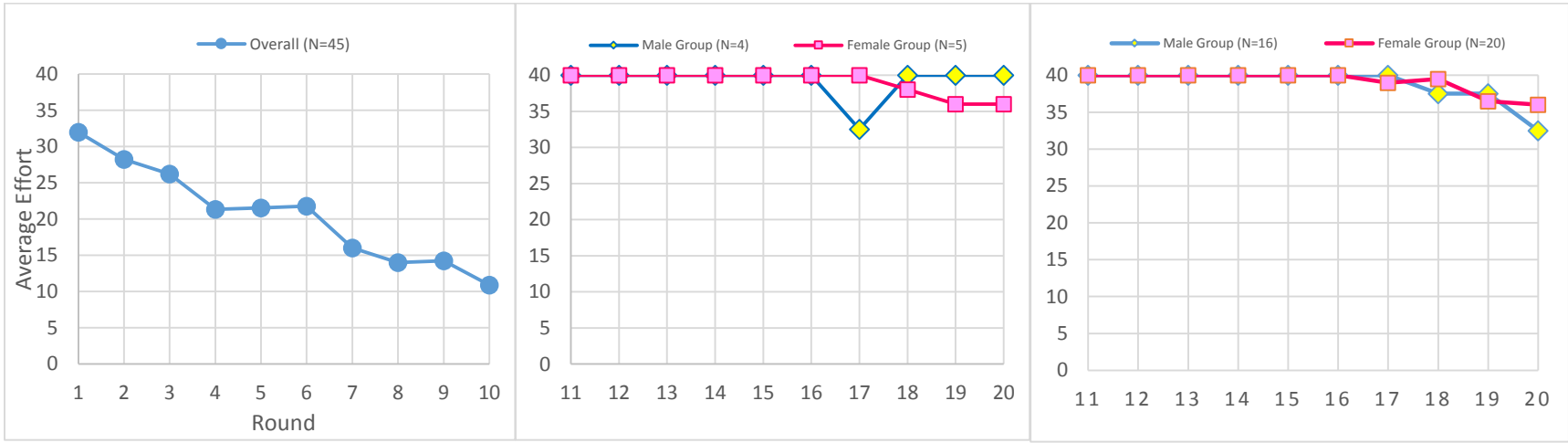


Figure 7B: Average effort choice in Rounds 11-20 in free-form message treatment (Experiment 2)

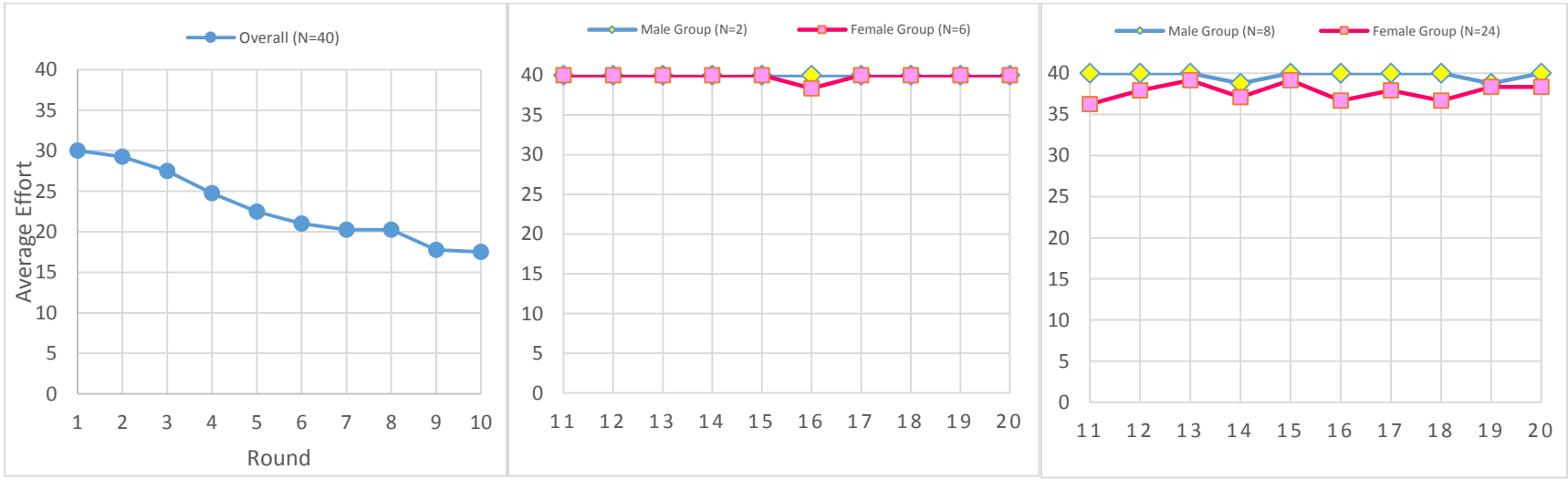


Figure 8A: Distribution of messages in Experiment 1

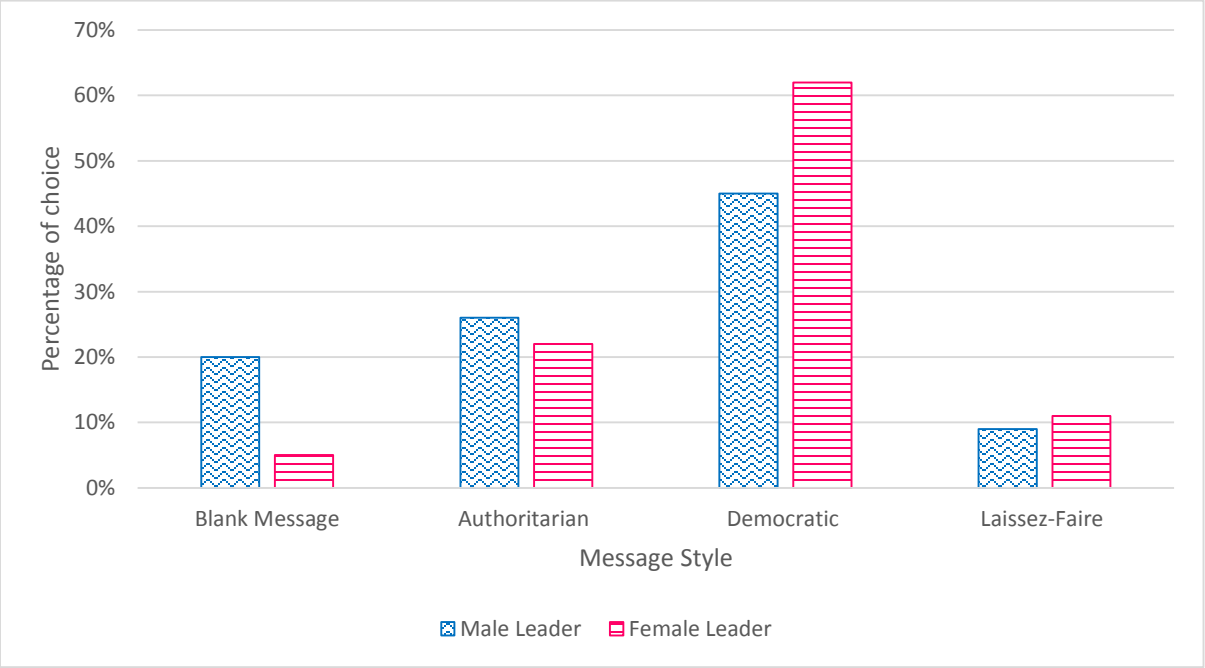
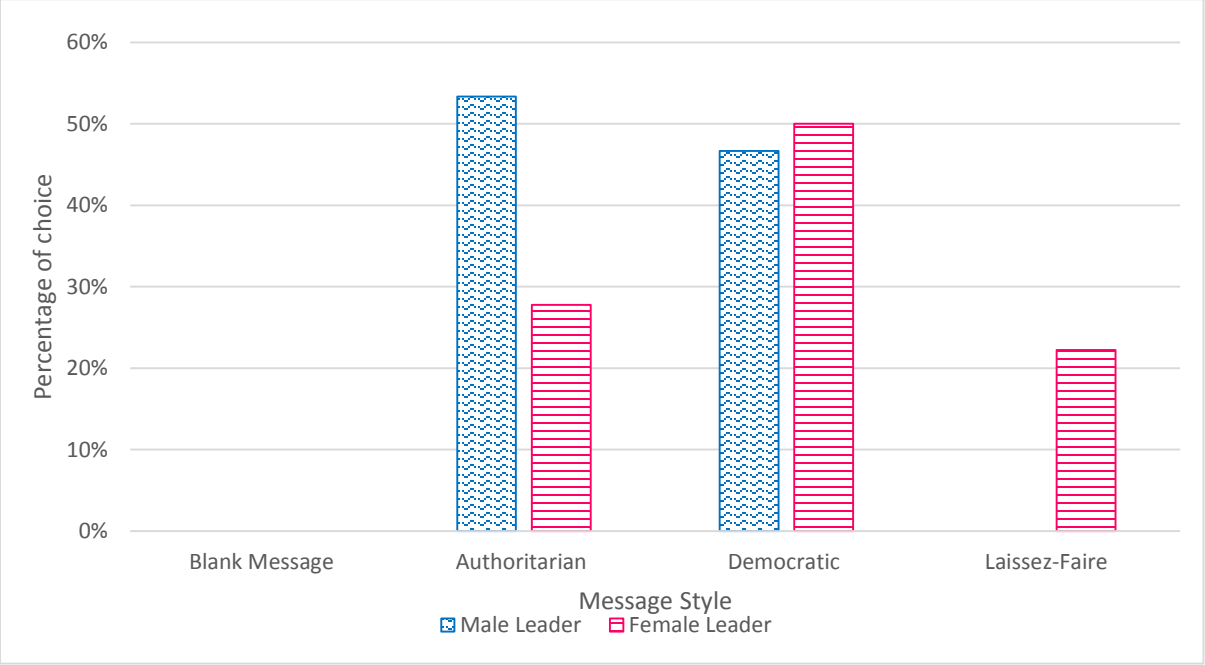


Figure 8B: Distribution of messages in Experiment 2



Un)Willing to lead? Men, Women and the Leadership Gap

Yaxiong (Sherry) Li[1,2,*], Erwann Sbai[1], Ananish Chaudhuri[1,3]

1. University of Auckland; 2. Xi'an Jiaotong-Liverpool University; 3. CESifo Munich
*; Corresponding author: sherry.li@auckland.ac.nz

e.sbai@auckland.ac.nz; a.chaudhuri@auckland.ac.nz

Abstract

We explore the causes behind the gender gap in leadership; as one moves up the organizational hierarchy, one encounters fewer women. We use the weak-link game paradigm to simulate intra-organization coordination problems, where participants can volunteer for leadership roles. The leaders' job is to resolve potential coordination failures. We look at whether: (1) there are systematic gender differences in the willingness to lead and (2) followers are less likely to follow female leaders. We find that, compared to men, fewer women volunteer to lead, particularly when the leader's gender is revealed to the followers. But, by and large, male and female leaders choose similar messages and/or actions in this game, and controlling for those, groups achieve similar levels of coordination success regardless of the leader's gender. We do not find evidence of resistance against female leadership, even though anticipation of such backlash may lie behind the female reluctance to lead.

JEL Classification Codes: C91, D91, J16

Keywords: Experiments, Coordination, Gender, Leadership

Acknowledgements:

Funds for this study were provided by the University of Auckland Business School through a Faculty Research Development Fund grant. We thank William Zhi for programming help. The paper has benefited from extensive discussion with Bram Cadsby and Fei Song during the initial stages. We are grateful to Ryan Greenaway-McGrevy, Phillip Grossman, John Hillas, Steffen Lippert and Martin Kocher for comments. We gratefully acknowledge feedback received from seminar participants at QUT, University of Guelph, Indian Institute of Management Kolkata, University of Auckland Leadership Institute, University of Auckland Department of Economics as well as those from participants at numerous conferences including the ESA Annual Meeting in Berlin (2018), ESA North American Meetings (Tucson, 2016 and Richmond 2017), Australia New Zealand Workshop in Experimental Economics (Melbourne 2017 and Auckland 2018) and the Women in Economics Network (WEN) Meeting at RMIT in 2019. All errors are our responsibility.

Appendix A: Instructions

General instructions

Welcome. The University of Auckland has provided funding in order to conduct this research. The instructions are simple. If you follow them closely and make appropriate decisions, you may make an appreciable amount of money. For this experiment all earnings are denoted in experimental dollars. At the end of the session you will be paid your earnings in cash at the rate of 400 experimental dollars = NZ $1. This money is in addition to the $5 show-up fee that you get.

The experiment will be conducted using computers. Please do not talk at any point during the experiment. If you have any questions then please raise your hand and one of us will come to you to answer it.

Specific instructions

This experiment consists two parts. In each part, there will be 10 rounds. We will give you the instructions to the first set of ten rounds now. We will give you the instructions to the next set of ten rounds at the end of the first ten rounds. At the completion of the 20 rounds of the game, we will ask you fill out two questionnaires.

Instructions for Part 1: first 10 rounds

One way to think about this experiment is that you are an employee of an organization. You will be part of a group consisting of 5 employees, that is you and four other employees. However, you will not learn the identity of the other people in your group in any round. The composition of these groups will remain unchanged for the entire time. This means that you will be interacting with the same four other people during the whole experiment.

In each round every employee will decide how many hours to work. The number of hours you may choose to work are {0, 10, 20, 30, or 40}. Your earnings in each round will depend on the number of hours that you pick and the smallest number of hours chosen by any participant in your group, including your own choice of hours.

Table 1 below tells you the potential payoffs you may receive. The earnings in each round may be found by looking across from the value you choose on the left hand side of the table and down from the smallest value chosen by any participant from the top of the table. For example, if you choose to work 20 hours while the minimum number of hours chosen in your group is 10 (that is, there is at least one person who has chosen to work for 10 hours and no one has chosen 0 hours), then you will earn 400 experimental dollars for that round. If you choose to work 30 hours and the smallest value chosen is 20 then you will earn 450 experimental dollars for that round. If you choose to work 40 hours and the smallest value chosen is 0, then you will earn 200 experimental dollars for that round.

Table 1:

Your choice of hours	Minimum of hours chosen by other members of the group				
	0	10	20	30	40
0	400	400	400	400	400
10	350	450	450	450	450
20	300	400	500	500	500
30	250	350	450	550	550
40	200	300	400	500	600

Notice: If someone chooses a particular value X as the number of hours and the others in the group all choose to work for at least X hours or more, then the smallest number of work hours chosen in the group is X. For example, if the number of hours chosen by the five members of a group are 0, 10, 20, 30 and 40, then the minimum number of hours is 0. Alternatively, if the number of hours chosen by the five members of a group are 10, 20, 20, 30 and 30, then the minimum number of hours is 10.

As noted above, the experiment will consist of 20 rounds. After we have finished reading the instructions you will proceed to play the first 10 rounds of this game.

We will stop at the end of the 10^{th} round. At that point we will give you further instructions about how the next ten rounds (Rounds 11 – 20) will work.

Please do NOT continue on to the 11^{th} round of this game until asked by the experimenter to do so.

We will pay you your earnings from the experiment at the end of the session. You are free to go once you have been paid. Your earnings are private information and we encourage you to keep this information private. If at any point you have any questions or problems, please raise your hand and ask for assistance.

Instructions for free-form message leader's gender revealed treatment in Experiment 1

Instructions for Part 2: Rounds 11 through 20

Before round 11 starts, you have a chance to volunteer to be a leader for your group. You will be asked if you wish to be a leader and you will have to respond by choosing "Yes" or "No." If you choose to be a leader by saying "Yes" then you will be asked to type a message. This message will be shown to the other 4 group members. If more than one member of your group volunteer to be the leader then the leader will be chosen at random by the computer. If no one has volunteered to be the leader then the leader will also be chosen at random by the computer.

There are still five people in each group, except now four of them are employees and one of them is the leader. More importantly, the Leader is now paid differently. The leader's payoff now depends on the hours of work chosen by the leader and also the minimum number of hours chosen by someone in the group, *including the leader*. This is shown in the payoff table below – Table 2. The larger the minimum number of hours chosen in the group, the higher is the leader's payoff.

The payoff table for the employees remains unchanged and is shown in Table 1.

Table 2: The leader's payoff is summarized in the following table:

Leader's Payoff	
Minimum hours chosen by employee	Earnings
0	100
10	300
20	500
30	700
40	900

But if the leader's choice of hours exceeds that of the workers, then there is an additional cost to the leader of the form Cost = 2(Leader's choice of hours – Minimum hours chosen by someone in the group).

For instance, if the leader chooses 20 hours and the minimum chosen is also 20 hours, that is, no one chooses to work for less than 20 hours then the difference between the leader's choice and the minimum choice is zero. In this case, the leader earns 500 – 2(20-20) = 500. On the other hand, if the leader chooses 40 hours while the minimum hours chosen in the group is 0, then the leaders' payoff is 100– 2(40-0) = 100 – 80 = 20.

For your convenience, Table 3 incorporates the information in Table 2 plus any additional cost incurred when leader's choice is larger than the minimum number of hours chosen in the group.

Table 3:

		Minimum of hours chosen by other members of the group				
		0	10	20	30	40
Leader's choice of hours	0	100	100	100	100	100
	10	80	300	300	300	300
	20	60	280	500	500	500
	30	40	260	480	700	700
	40	20	240	460	680	900

Once the leader has been chosen and the message written by the leader has been shared with the four employees, then the game proceeds as follows. In each round the leader moves first by choosing the number of hours. The leader can also choose to re-send the message sent before or write a different message. Once the leader has chosen the number of hours, the information regarding the leader's choice of hours, the leader's gender and the content of the leader's message will be revealed to the group members. The employees will then choose their hours simultaneously, i.e., each employee will make a choice without knowing the choice of the other employees.

Your earnings will be denoted in experimental dollars as before.

Your $5 show up fee and earnings for the first 10 rounds are unaffected.

Are there any questions? If at any point you have any questions or problems, please raise your hand and ask for assistance.

We will tell you when we are ready to proceed.

Instructions for free-form message leader's gender revealed treatment in Experiment 2

Instructions for Part 2: Rounds 11 through 20

Before round 11 starts, you have a chance to volunteer to be a leader for your group. You will be asked if you wish to be a leader and you will have to respond by choosing "Yes" or "No." If you are chosen as the leader, then the experimenter will provide you with a message that you can send to the other members of the group. This message will be shown to the other 4 group members. If more than one member of your group volunteer to be the leader then the leader will be chosen at random by the computer.

There are still five people in each group, except now four of them are employees and one of them is the leader. The payoff table for both leaders and employees remain unchanged and as shown in Table 1.

Once the leader has been chosen and leader's message has been shared with the four employees, then the game proceeds as follows. In each round the leader moves first by choosing the number of hours. The leader can also choose to re-send the message sent before or not. Once the leader has chosen the number of hours, the information regarding the leader's choice of hours, the leader's gender and the content of the leader's message will be revealed to the group members. The employees will then choose their hours simultaneously, i.e., each employee will make a choice without knowing the choice of the other employees.

Your earnings will be denoted in experimental dollars as before.

Your $5 show up fee and earnings for the first 10 rounds are unaffected.

Are there any questions? If at any point you have any questions or problems, please raise your hand and ask for assistance.

We will tell you when we are ready to proceed.

Appendix B: Holt and Laury Lottery Task

This experiment consists of ten paired lottery-choice, you are required to choose between option A and option B for each of these ten pairs. The following table shows these options and their probabilities and payoffs:

Lottery number	Option A	Option B	YOUR CHOICE
1	1/10 of $6.00, 9/10 of $4.80	1/10 of $11.55, 9/10 of $0.30	
2	2/10 of $6.00, 8/10 of $4.80	2/10 of $11.55, 8/10 of $0.30	
3	3/10 of $6.00, 7/10 of $4.80	3/10 of $11.55, 7/10 of $0.30	
4	4/10 of $6.00, 6/10 of $4.80	4/10 of $11.55, 6/10 of $0.30	
5	5/10 of $6.00, 5/10 of $4.80	5/10 of $11.55, 5/10 of $0.30	
6	6/10 of $6.00, 4/10 of $4.80	6/10 of $11.55, 4/10 of $0.30	
7	7/10 of $6.00, 3/10 of $4.80	7/10 of $11.55, 3/10 of $0.30	
8	8/10 of $6.00, 2/10 of $4.80	8/10 of $11.55, 2/10 of $0.30	
9	9/10 of $6.00, 1/10 of $4.80	9/10 of $11.55, 1/10 of $0.30	
10	10/10 of $6.00, 0/10 of $4.80	10/10 of $11.55, 0/10 of $0.30	

Please choose either Option A or Option B for those 10 lottery pairs. The computer will randomly select a lottery number after everyone has made their decisions.

Appendix C: Demographic Questionnaire

Please answer ALL of the questions on this survey as accurately as you can.

1. What is your Age? _____

2. What is your Gender? _____

3. What is your field of study at the University? _____

4. Are you an Undergraduate Student (which year) or a Postgraduate Student? Circle one.

   UG (Year _____)

   PG

5. What is your after-tax weekly income from ALL SOURCES (including salary, allowances & scholarships)? Circle one of the options below.

   Less than $250

   Between $250 and $750

   Between $750 and $1,250

   More than $1,250

6. Were you born in New Zealand? YES NO

7. Which ethnic group do you belong to? Circle one of the options below.

   New Zealand European

   Maori

   Samoan

   Cook Island Maori

   Niuean

   Chinese

   Indian Subcontinent (including Pakistan and Bangladesh)

   Other (Please specify) _____

Appendix D: Additional Results

Table A6: Random effects ordered probit model for follower's effort in Experiment 2 (cluster on the group); ***counter-part to Table 6 that reports the results for Experiment in the paper.***

Dependent variable: Choice of effort level by follower

Effort Choice	**Pre-set**	**Pre-set**	**Free-form**	**Free-form**
Round	-0.594***	-1.013***	-0.030	0.012
	(0.107)	(0.156)	(0.058)	(0.072)
Female	-4.694**	-12.999***	-0.320	2.109*
	(1.974)	(2.891)	(1.728)	(1.230)
Female*Round	0.555***	1.011***	0.051	0.034
	(0.130)	(0.156)	(0.100)	(0.106)
Female Leader	-0.259	0.279	-1.087	-8.236***
	(0.446)	(0.296)	(0.722)	(0.911)
Lag Earning	0.001	-0.003*	0.008***	0.007**
	(0.002)	(0.002)	(0.001)	(0.003)
Risk averse	-1.105*	-1.208**	-0.026	2.625***
	(0.589)	(0.531)	(0.461)	(0.328)
Demographic Control	NO	YES	NO	YES
Number of observations	270	270	220	220
Wald χ^2	2232.72	-	157.51	-
Prob > χ^2	0.000	-	0.000	-

*Notes: Standard errors in parentheses; ***, ** and * represent significance at 1%, 5% and 10% respectively.*

Notes: In experiment 2, the majority of followers choose effort of 40, therefore the random effects ordered probit model is not a good fit.

Table A6*: Random effects model for follower's effort in Experiment 2 (cluster on the group)

Dependent variable: Choice of effort level by follower

Effort Choice	**Pre-set**	**Pre-set**	**Free-form**	**Free-form**
Round	-1.046***	-1.048***	-0.045	-0.017
	(0.250)	(0.257)	(0.040)	(0.060)
Female	-12.066***	-11.965***	1.089	1.220
	(3.311)	(3.595)	(4.022)	(3.825)
Female*Round	0.874***	0.883***	0.007	0.009
	(0.224)	(0.231)	(0.186)	(0.186)
Female Leader	-0.010	-0.128	-1.840	-1.442
	(0.644)	(0.508)	(1.403)	(0.917)
Lag Earning	0.022	0.021***	0.020***	0.018***
	(0.005)	(0.005)	(0.007)	(0.007)
Risk averse	-1.527	-2.270*	1.120	2.222
	(1.240)	(1.234)	(1.340)	(2.405)
Demographic Control	NO	YES	NO	YES
Number of observations	270	270	220	220
Wald χ^2	126.70	-	1414.69	-
Prob > χ^2	0.000	-	0.000	-

*Notes: Standard errors in parentheses; ***, ** and * represent significance at 1%, 5% and 10% respectively.*

Table A7: Random Effects model for effort level under Pre-set message treatment in Experiment 2 (cluster on the group); ***counter-part to Table 7 that reports results for pre-set message treatment in Experiment 1 in the paper.***

Follower's Effort Choice	(1)	(2)	(3)	(4)
Round	-0.707***	-0.865**	-0.795***	-0.800***
	(0.269)	(0.350)	(0.283)	(0.293)
Female*Round	0.538***	0.674***	0.627***	0.639***
	(0.199)	(0.262)	(0.217)	(0.228)
Male follower with female leader	0.527	0.678	0.500	0.565
	(0.558)	(0.589)	(0.381)	(0.632)
Female follower with male leader	-6.913***	-8.817**	-8.345***	-8.021**
	(2.676)	(3.652)	(3.165)	(3.455)
Female follower with female leader	-7.021**	-9.120**	-8.581***	-8.312**
	(2.730)	(3.842)	(3.257)	(4.097)
Message shown	-	-	1.331	1.412
	-	-	(1.855)	(1.960)
Leaders Effort	0.335**	0.343***	0.365***	0.352***
	(0.140)	(0.121)	(0.103)	(0.112)
Lag minimum effort	0.097**	0.123**	0.117**	0.113**
	(0.048)	(0.062)	(0.054)	(0.052)
Risk averse	-	-	-1.065	**-1.618*** **
	-	-	(0.772)	**(0.628)**
Constant	32.620***	34.099***	31.462***	31.645***
	(7.619)	(7.978)	(6.120)	(7.090)
Demographic Controls	NO	NO	NO	YES
Wald χ^2	411.699	6088.546	-	-
Prob > χ^2	0.000	0.000	-	-
Number of observations	360	270	270	270
Wald test for equality of coefficients				
Male follower with female leader = Female follower with male leader	Wald χ^2=6.82 Prob > χ^2=0.01	Wald χ^2=5.90 Prob > χ^2=0.02	Wald χ^2=7.13 Prob > χ^2=0.01	Wald χ^2=6.80 Prob > χ^2=0.01
Male follower with female leader = Female follower with female leader	Wald χ^2=6.76 Prob > χ^2=0.01	Wald χ^2=5.77 Prob > χ^2=0.02	Wald χ^2=7.25 Prob > χ^2=0.01	Wald χ^2=5.20 Prob > χ^2=0.02
Female follower with male leader = Female follower with female leader	Wald χ^2=0.66 Prob > χ^2=0.42	Wald χ^2=0.33 Prob > χ^2=0.57	Wald χ^2=0.10 Prob > χ^2=0.75	Wald χ^2=0.09 Prob > χ^2=0.77

*Notes: Standard errors in parentheses; ***, ** and * represent significance at 1% , 5% and 10% respectively.*

Table A8: Random Effects model for effort level under Free-form message treatment in Experiment 2; *counter-part to Table 8 that reports results for free-form message treatment from Experiment 1 in the paper.*

Follower's Effort Choice	(1)	(2)	(3)	(4)	(3)	(4)
Round	-0.170	-0.187	-0.046	-0.170**	-0.187*	-0.046
	(0.120)	(0.123)	(0.123)	(0.081)	(0.105)	(0.080)
Female*Round	-0.224	-0.244	-0.029	-0.224	-0.244	-0.029
	(0.158)	(0.157)	(0.165)	(0.310)	(0.320)	(0.170)
Male follower with female leader	-1.163	-1.568	-1.817	-1.163	-1.568	-1.817
	(0.708)	(0.975)	(1.214)	(1.035)	(1.622)	(1.463)
Female follower with male leader	3.810	4.123	1.418	3.810	4.123	1.418
	(2.686)	(2.672)	(3.219)	(4.851)	(5.005)	(3.407)
Female follower with female leader	2.554	2.602	-0.187	2.554	2.602	-0.187
	(2.539)	(2.569)	(2.722)	(4.518)	(4.915)	(3.695)
Leaders Effort	1.454***	1.389***	1.266***	1.454***	1.389***	1.266***
	(0.208)	(0.206)	(0.253)	(0.184)	(0.200)	(0.098)
Lag minimum effort	0.221***	0.211***	0.122***	0.221***	0.211***	0.122**
	(0.023)	(0.024)	(0.026)	(0.083)	(0.080)	(0.060)
Authoritarian Message Style	-	-1.296	-0.720	-	-1.296	-0.720
	-	(0.945)	(0.961)	-	(1.375)	(1.084)
Democratic Message Style	-	-1.617**	-1.139	-	-1.617	-1.139
	-	(0.708)	(0.843)	-	(1.583)	(1.615)
Laissez-Faire Message Style	-	-0.204	-0.660	-	-0.204	-0.660
	-	(0.955)	(0.989)	-	(0.918)	(1.063)
Risk averse	-	-	0.930	-	-	0.930
	-	-	(0.799)	-	-	(1.167)
Constant	-23.892***	-19.370**	-14.574	-23.892***	-19.370**	-14.574***
	(8.491)	(8.556)	(10.433)	(5.565)	(7.980)	(3.323)
Demographic Controls	NO	NO	NO	YES	YES	YES
Wald χ^2	178.528	179.173	71.621	-	-	-
Prob > χ^2	0.000	0.000	0.000	-	-	-
Number of observations	320	320	220	320	320	220

Notes: There were very few choices low effort choices in this treatment in Experiment 2. As a result, we could not run a random effects ordered probit regression. Therefore, we have chosen to report results from a random effects regression instead.

Table A9: Random Effects Model for follower's round earnings in Experiment 2 Pre-set Message treatment (clustered on the group); ***counterpart to Table 9 that reports earnings from Experiment 1 the paper.***

Earnings	**Pre-set**	**Pre-set**	**Free-form**	**Free-form**
Round	-14.057***	-13.771***	-2.895	-1.991
	(5.377)	(4.926)	(2.077)	(2.151)
Female	-101.771**	-129.919**	-24.028	-4.512
	(50.985)	(56.826)	(24.984)	(16.012)
Female Leader	10.786	5.074	-7.176	-11.577
	(9.750)	(7.312)	(11.919)	(12.459)
Female*Round	7.601**	9.610**	1.023	0.346
	(3.659)	(4.025)	(1.214)	(0.868)
Leader's Effort	9.481***	9.231***	3.440	2.747
	(1.302)	(1.351)	(2.880)	(1.824)
Lag Minimum Effort	2.018**	1.927**	2.507**	1.643*
	(0.976)	(0.917)	(1.128)	(0.915)
Risk Averse	-	-1.476	-	-2.298
	-	(6.555)	-	(2.766)
Constant	336.368***	348.933***	409.076***	456.839***
	(94.971)	(95.688)	(108.984)	(75.124)
Wald χ^2	2911.998	10893.538	901.805	2359.411
Prob > χ^2	0.000	0.000	0.000	0.000
Number of observations	360	270	320	220

*Notes: Standard errors in parentheses; ***, ** and * represent significance at 1%, 5% and 10% respectively.*