Gender, Communication Styles, and Leader Effectiveness

Timko, Krisztina

16 February 2017

Gender, Communication Styles, and Leader Effectiveness

Krisztina Timko[1]
University of Helsinki, e-mail: krisztina.timko(at)gmail.com

February 2017

Abstract

We study gender differences in the behavior, communication, and effectiveness of randomly selected leaders in a laboratory experiment using the turnaround game. Leaders can send non-binding pre-play text messages to try to convince followers to coordinate on the Pareto-efficient equilibrium. The treatment variations consist of the gender of the leader, and whether the communication is one-way (only leaders send messages) or two-way (first followers send messages to their leader, and subsequently the leader sends messages to the group). We find that male leaders communicate more assertively. Communication with the followers induces female leaders to express significantly more often that they are part of the group, rather than standing above the group. Despite the different paths in communication, both men and women are equally likely to request the highest effort contribution. Men and women are equally effective leaders.

JEL Codes: M14; M54; J16; C92
Keywords: gender differences; leadership; leader effectiveness; coordination

1 Introduction

Gender equality is a major concern for some politicians, international organizations, and the general public. In many parts of the world girls and women have limited access to education, health care, and the labor market, and they have limited rights to individual decision-making and representation in political and economic decision-making processes. Gender gaps in labor force participation, compensation, and career development are still a concern in Western societies as well. Globally, about three quarters of working-age men participate in the labor force, compared to half of working-age women (ILO 2013; OECD 2016). The global gender pay gap is around 15-20%, and the annual pay for women reached only today the amount men were earning ten years ago (OECD 2014; Eurostat 2014; World Economic Forum 2015). While gender parity is reached in university attendance, the gender gap in taking skilled roles is 25%, and the gender gap in leadership roles is huge with 72% (World Economic Forum 2015). One of the global sustainable development goals of the United Nations is the economic and political empowerment of women and to close significant gender gaps in terms of labor market, compensation, and access to and control over decision-making, as well as to support women in participating more fully in political processes, spheres of public life, and in taking on more leadership roles (United Nations 2015).

Our paper contributes to the scientific insights about the possible explanations why women are underrepresented in leadership roles, and what kind of changes to expect if women fill empowering

[1]I would like to thank my doctoral thesis supervisor, Ernesto Reuben for his valuable and helpful guidance.

positions (Goldin 2002; Akerlof and Kranton 2005; Eagly and Johannesen-Schmidt 2007; Wang et al. 2011; Dezső and Ross 2012). We study the interaction of followers and leaders in a laboratory experiment using the turnaround game, a variation of the weak-link (or minimum effort) coordination game (Brandts and Cooper 2007). In particular, we study whether there are any differences in the way male and female leaders communicate to their followers, and whether this affects their effectiveness.[2]

The basic weak-link game by Van Huyck et al. (1990) models a coordination problem in an organizational setting, where a group of players engages in a joint economic activity with complementarities. Players choose their individual effort contribution simultaneously, and the lowest performing player, the "weak link", determines the performance of the organization. Classical examples are "the assembly line that moves no faster than the slowest line worker", collaborative work that is incomplete until the final contribution is finished, "perceptions of overall product quality that is sensitive to the worst performing feature", network investments, and adoption of standards (Brandts et al. 2015; Heinemann et al. 2009). The weak-link game has multiple Pareto-ranked Nash equilibria on every effort level that is simultaneously chosen by all players. The most efficient, *payoff-dominant equilibrium* is reached if all players contribute the highest effort level. However, the higher the individual effort level choice, the riskier the decision is, since it is uncertain what other players choose. If other players contribute less effort, a higher effort contribution gets devalued, while the costs of effort remain unchanged. In order to avoid costs of coordination failure, the player could choose the lowest effort level, which is a *secure option*. The secure option yields a lower payoff than what could be gained if all players choose simultaneously the highest effort level. Thus, the main feature of the weak-link game is the trade-off between the payoff-dominant and the secure equilibrium, and strategic uncertainty about the choices of others can undermine coordination on the efficient effort level (Van Huyck et al. 1990).

The turnaround game introduces organizational hierarchy to the weak-link game, after a failure to coordinate on the efficient equilibrium (Brandts and Cooper 2007). Therefore, the turnaround game has two parts. In the first part, groups of players have to coordinate repeatedly in a basic weak-link game. Usually groups establish a failure history and induce coordination on an inefficient equilibrium during this initial phase of the turnaround game (Brandts and Cooper 2006a). In the second part, the turnaround game introduces leadership, and leaders can have different instruments at their disposal, like varying earnings bonuses and communication. The task of the leaders is to try the "turnaround", leading the group from the inefficient low effort equilibrium to the efficient high effort equilibrium. The turnaround is harder to achieve if the group experienced a failure history (Devetag and Ortmann 2007).

[2] Throughout the paper, we use the term "effectiveness" to refer to the leader's successfulness. An effective leader is both efficient and credible. With "efficiency" we refer to the performance of the leader, which the leader has under control. For example, leaders who work harder, are abler, and send more relevant messages are more efficient. With respect to the group, groups, who coordinate on the same effort level, are said to coordinate more efficiently. With "credibility" we refer to the beliefs of the followers that the leader is followed by others as well. The follower, who believes that the leader is credible, will follow the leader. For example, elections, incoming follower messages, or being male can increase leader credibility, but such factors are not necessarily under the control of the leader. Credible leaders are followed, no matter what the message is, whereas effective leaders achieve the best and most desired outcome with their groups.

In our design, leaders are selected randomly and we reveal the gender of the leader. Leaders can use communication. Communication is voluntary and non-binding (cheap talk), and the messages are of free form. The treatment variations consist of the gender of the leader, and whether the communication is one-way or two-way. One-way communication means that we allow only leaders to send public messages visible to all of their followers. Two-way communication means that we allow first followers to send private messages to their leaders, while these follower messages are only visible for the leader, and subsequently we allow leaders to send public messages visible to all of their followers. Both communication conditions model a centralized organizational structure. Under the one-way communication condition the direction of communication is top-down only, for example, if a headquarters gives unified orders to the division managers. Under the two-way communication condition the leader keeps an "open door" and subordinates or division managers have the possibility to privately inform their leader before receiving unified orders.

One of the closest papers to our work is Kriss and Eil (2012). Using a weak-link game and assignment of a leader, who can use communication, the authors vary three communication conditions: follower messages are not possible (one-way communication), visible only to the leader (private two-way communication), or visible to all group members (public two-way communication). All messages have a low cost.[3] The authors find that private follower communication to the leader leads to greater leader credibility and more efficient coordination than prohibited follower communication, and prohibited follower communication leads to more efficient coordination than public follower communication. Public visibility of follower messages and public disagreement among followers can undermine leader communication trying to induce common beliefs and efficient coordination. An authoritarian leader could prohibit follower communication, which would lead to an increase in leader credibility. However, some followers might still not rely on others following the leader as well, since a leader message without any follower input is only a suggestion or request, and cannot credibly claim to represent follower intentions. Limiting the visibility of follower messages instead of prohibiting those, improves coordination. The authors call the possibility of private follower communication "open door" or "suggestion box" of the leader. The "open door" facilitates for followers to recognize their leader as a "coordination device", who can hear the followers and is perceived to have the support of the group, and this makes leaders more credible and more influential. Those leaders who recognize their pivotal role and that the "open door" improves leader credibility can use this fact to send messages suggesting the highest effort level aiming for maximum effectiveness. Though, incoming follower messages suggesting low effort levels might also undermine the confidence of the leader to take advantage of the boosted influence by the "open door". Kriss and Eil say that leaders whose messages are more a consequence of follower input lack "thick skin". The authors assume that successful real-world leaders develop such a "thick skin", which enables them to "appropriately disregard counterproductive input while still recognizing the broader value of communication within the organizational hierarchy".

Another close paper to our work is Cooper (2007), who studies communication strategies, and compares the behavior of experienced managers and undergraduate students. The proportion of men in his study was much higher among the experienced managers than the undergraduates. The

[3] In a related study, Kriss et al. (*forthcoming-b*) show that even small costs can reduce the use of communication. Due to Heath and Staudenmayer (2000), people tend to underestimate the difficulty of coordinating each others' actions and they undervalue communication. To exclude such an effect, we have no message costs in our design.

scarcity of women in Cooper's study reflects the male dominated manager population. Cooper finds that experienced managers are more likely to use the "good" communication strategy, and therefore they are more effective leaders. The "good" communication strategy is simple: "specifically request a high effort and point out the mutual benefits of high effort". Participants in our study can also send any kind of free form text messages. As a consequence we collect data on communication styles and this enables us to conduct a rich analysis of the message content. We lean our message content analysis on the systematic scheme invented by Brandts and Cooper (2007). We add gender-specific coding aspects based on the stereotypes of men tending to be more task-oriented and assertive, and women to be more relationship-oriented and expressive (Bem 1974; Bem 1993; Merchant 2012).

We also analyze the message content based on the classification of leadership styles in Eagly et al. (2003). Eagly et al. conducted a meta-analysis of 45 studies and classify three leadership styles: transformational, transactional, and laissez-faire leadership style. The transformational leadership style was first described by Burns (1978) and elaborated by Bass (1985, 1998). *Transformational* leaders are future-oriented, innovative leaders, who empower their followers to contribute more capably to their organization. They often act as role models, they emphasize the importance of the organization's mission, they exhibit optimism and excitement about goals and future states, and they attend to the individual needs of their followers, focusing on their development and mentoring. Many features of the transformational leadership style are in common with *charismatic* leadership (Conger and Kanungo 1998). Transformational leaders, in our context, would emphasize the mutual benefit argument and being part of the group along with sending a high effort level suggestion.

Leadership researchers contrasted the transformational style to the transactional style (Burns 1978; Avolio 1999; Bass 1998). While the transformational leadership style focuses on the individual development of followers, creating human capital and making use of it, the transactional style is a classic, more conventional leadership style. *Transactional* leaders appeal to the self-interest of their followers in that they establish exchange relationships with them: the transactional leader clarifies the responsibilities of the follower, and rewards followers if they meet objectives or correct them for failing to meet objectives. The transactional leadership style has three subcategories: contingent reward, active management by exception, and passive management by exception. Contingent reward provides rewards if follower performance meets objectives. In the context of our study, positive feedback, for example, could be seen as contingent reward. Active management attends to followers only if they fail to meet objectives. One example in our study are leaders who start to send social banter after the group reached a high effort level, instead of repeating seemingly unnecessary messages. Passive managers wait until problems get serious and intervene only then. As an example in our study, we could think of a leader who sends no message in the first message period, but intervenes later. Both transformational and transactional leaders recognize their pivotal role and take responsibility for the leadership. In contrast, the third leadership style defines leaders who lack recognition of their situation and authority, and let followers do as they want. The *laissez-faire* leader is frequently absent, lacks involvement during critical junctures, and generally fails to take responsibility for managing. In our study, this type would either send no message or only banter, without trying the turnaround.

Eagly et al. find that leader effectiveness relates positively to transformational leadership and contingent reward behavior as a subcategory of transactional leadership style.[4] Other subcategories of transactional leadership, namely active and passive management, and the laissez-faire leadership style relate negatively to the performance of the organization. According to Eagly et al., earlier studies on leadership styles before 1990 distinguished *democratic* and *autocratic* leadership styles (Lewin and Lippitt 1938; Vroom and Yetton 1973). Democratic leaders allow followers to participate in decision making, whereas autocratic leaders discourage their followers from participating in decision making. However, the democratic-autocratic classification of leadership style does not describe leader behavior, and the implications for leader effectiveness depend much on the organizational structure, for example, whether participation in decision making is feasible and allowed (Foels et al. 2000; Gastil 1994; Vroom and Yetton 1973). Therefore the above-described contemporary classification is more suitable to be linked to leader effectiveness.

Eagly et al. (2003) show small gender differences in leadership styles. The authors find that female leaders engage more often than male leaders in transformational leadership and in contingent reward behavior as part of the transactional leadership style, and both styles predict leader effectiveness. These findings are in line with earlier research findings and claims of a female leadership advantage (Sharpe 2000; Bass et al. 1996; Lowe et al. 1996). Later, Eagly and Johannesen-Schmidt (2007) reviews research and meta-analyses from the 1950s up to date. The authors find that even if gender differences in leadership styles are small, they are important. In a meta-analytic review of 25 years of research, Wang et al. (2011) find that the transformational leadership style is effective in most organizational context, and women, who engage in such leader behavior, gain at least some advantage. In a panel data investigation of U.S. corporations, Dezső and Ross (2012) find a positive performance effect of female representation in top management teams, yet, only in firms, whose strategies focused on innovation, facing complex challenges.

In our study, transformational leaders use the "good" communication strategy as defined in Cooper (2007). Cooper points out, that "good" communication is likely to be effective. We find that the suggestion of the highest effort level combined with assertive content and expressions of being part of the group are the most relevant message contents inducing effectiveness. Although, men communicate more often assertively than women and women with an "open door" send more often content emphasizing that the leader is coequally part of the group, we find no gender difference in the use of relevant message content. Moreover, men and women are equally credible and equally effective leaders, and thus we find no reason why women should not fill more of the top positions. Our findings further suggest that there will be a change on the global stage in the way leaders interact and communicate with followers if women would fill more top positions.

2 Experimental Design and Procedures

Each experimental session consists of 26 periods. At the beginning of a session, participants are randomly matched into groups of five and are informed that their group's composition will not change throughout the session. In each period, every participant $i \in \{1,2,3,4,5\}$ in group k

[4] Concerning charismatic leadership, Antonakis et al. (2015) show in a field experiment that charismatic speeches have a performance effect: workers increase their costly effort input and generate higher firm output.

simultaneously chooses an effort level $effort_i \in \{0,10,20,30,40\}$. Participant i's earnings are equal to:

$$\pi_i^k = 200 - 5 \times effort_i + 6 \times effort_k^{min},$$

where $effort_k^{\min}$ is the minimum effort chosen in the group. At the end of each period, participants are informed of their earnings and the group's minimum effort. Participants cannot observe others' effort choice, which makes it more difficult to escape coordination failure (Brandts and Cooper 2006b).

Each session is divided into two parts. Part 1 consists of periods 1 to 8 and Part 2 of periods 9 to 26. Participants know the session has two parts but are not given the specific instructions of Part 2 until they reach that part.

In Part 1, participants play the weak-link game without a leader. Given previous evidence with these parameters (Brandts and Cooper 2006a; Brandts et al. 2015), groups are likely to end up coordinating on the lowest effort level by the end of Part 1. By having groups fail to coordinate on high effort levels, the introduction of a leader is more meaningful.

In Part 2, we introduce leaders. At the beginning of period 9, one participant in each group is randomly selected to be the group's leader, which leaves the other group members as followers. The leader holds the position until the end of the session. Every three periods, which we refer to as message cycle, the leader has the option to send a written message visible to all followers. The message is sent before effort choices are made. Leaders can write anything they wish, including nothing, except for content that could be used to identify them. Messages are non-binding in that not following a message has no direct effect on earnings. Leaders make effort decisions and face the same incentives as followers.

2.1 Treatments

The study uses a 2x2 between-subjects factorial design. The two factors are whether followers can send messages and the gender of the leader. We randomly assigned groups to one of the four treatment conditions.

The first treatment variation is whether there is *One-way* or *Two-way* communication. In *One-way*, followers cannot send any messages, only leaders can do so. In *Two-way*, each follower can send a written message visible to the leader only. The message is sent before leaders can send their messages. Followers can write anything they wish, including nothing, except for content that could be used to identify them.

The second treatment variation is the gender of the leader. By randomly assigning the leadership position we randomly vary the leader's gender across the groups. To reveal information about gender, participants had to choose a profile picture they identified with. This occurred after they consented to take part in the study but before they read the instructions to avoid strategic selection of profile pictures. We created 12 generic profiles for each gender using the profile creator website pickaface.net (see the Appendix). All pictures have the same clothing, facial expression, face form, and eye color. We varied hair length, hair color, skin color, and did small modifications to the lips,

nose, eyes, and hairstyle to match generic racial features. We use profile pictures to preserve anonymity whilst revealing gender. We opted for pictures that also contain other cues such as race and hairstyles to distract participants from discerning the purpose of the study (Zizzo 2010), which can potentially lead to intentional changes in behavior (Camerer 2015). We displayed the profile picture of leaders along with their messages visible to their followers.

2.2 Predictions

The main interest in this paper is to show whether men and women differ in leader effectiveness across communication conditions and which channels explain leader effectiveness: a difference in leader behavior or a difference in follower behavior. Naturally, all our predictions refer to behavior in Part 2.

Since men have historically held most of the leadership roles in society, followers might hold biased stereotype beliefs that consider men to be better, more competent and more credible leaders (Ridgeway 2001; Day 2014; Goldin 2002; Akerlof and Kranton 2005). Reuben et al. (2014) show that, in an experimental market, stereotypes make both male and female participants twice more likely to hire a man than a woman when no other information is available than a candidate's appearance (which makes gender clear), despite the fact that on average both genders perform equally well. Grossman et al. (2016) show, using the turnaround game, that followers are more likely to follow men than women, holding leader messages constant. Even if followers do not hold biased beliefs themselves, it might be rational to adapt to the supposed bias if they expect others to follow men to a greater extent.

Prediction 1*: In both One-way and Two-way, male leaders are followed to a greater extent than female leaders.*

Leaders will tend to send messages in line with their own effort choices at least in the first period of the three-period message cycles. Otherwise leaders would not get feedback on their credibility, because they could not rule out that they are the only ones who do not follow their own message. In this sense, leaders face an exploration cost.[5] Further, leaders who do not follow their own message would lose credibility already in the first period. Since a possible loss in earnings in the first period is less severe than the loss of credibility for all the coming periods, leaders face a cost of screwing up others.[6] Given that requesting and choosing a high effort level is risky and requires high credibility, more risk averse or less confident leaders might request and choose a low effort level, whereas leaders who are more willing to take risks, or who are overconfident about the number of their followers, or both, are also the ones who will request and choose high effort levels. Men have been

[5] Leaders can minimize the exploration cost if they communicate a conditional strategy, for example, that they will follow their own message in the first period of the message cycle, but they will stop doing so if others do not follow as well.

[6] In a recent related paper, Cooper et al. (2016) study the "social credibility" of leaders in an experiment where leaders try to induce followers to invest in a joint venture. The authors find that "leaders manage social credibility by forgoing potentially profitable requests for investment in order to make it more likely that subsequent recommendations to invest are followed". Leaders in our study might also recognize the need to build up credibility, which is more than having other regarding preferences. For example, less confident leaders, who intend to choose a low effort level, can stay credible if they prevent their followers from losses by requesting a low effort level.

shown to be more willing to take risks and to be more overconfident (Reuben et al. 2012; Reuben et al. (in press); Dohmen et al. 2011).

Prediction 2*: In both One-way and Two-way, male leaders request more often the highest effort level than female leaders.*

If *Predictions 1* and *2* hold, men are more effective leaders than women in both *One-way* and *Two-way*. Kriss and Eil (2012) already showed that incoming follower messages in *Two-way* increase leader credibility, compared to *One-way*. Therefore leaders in *Two-way* might be more effective than leaders in *One-way*. Moreover, if, as assumed in *Prediction 1*, followers expect others to follow male leaders to a greater extent than female leaders, they might send lower effort level suggestions to female, rather than male leaders. The low follower suggestions can affect leaders negatively (Kriss and Eil 2012), and the credibility of female leaders who request high effort levels might be doubted more than the credibility of their male counterparts. Thus, in *Two-way*, the leader effectiveness of men, rather than women, might be boosted more.

Prediction 3*: In both One-way and Two-way, men are more effective leaders than women. In Two-way, the gender gap is larger.*

Merchant (2012) provides an overview on gender differences in communication styles and points out that men and women have different purposes when using communication. Women value the process of communication itself, whereas men view communication as a tool to reach a certain goal. Women use communication to enhance social connections, while men exert dominance and remain goal-oriented. In interactions, women tend to be more social, while men value independence and remain unemotional and less attached to conversations. Therefore, the language that women use is more expressive and polite, while men use a more assertive and dominating language, signaling status and overconfidence. Merchant (2012) show that gender differences in communication styles are often persistent in leadership styles as well.

Eagly et al. (2003) classify three leadership styles: transformational, transactional, and laissez-faire leadership style, described in more detail in the introduction. In our context, transformational leaders would request a high effort level and emphasize the mutual benefit argument and being part of the group. We call such “transformational” messages “relevant messages”, since Eagly et al. find that leader effectiveness relates positively to the transformational leadership style, and Cooper (2007) also finds the same message strategy leading to effectiveness. Transactional leaders, who engage in contingent reward behavior, would request a high effort level in an ordering style, give more often positive feedback, and encourage their followers with positive emotional content, rather than long explanations. Both transformational and transactional leaders recognize their pivotal role and are likely to communicate assertively. Eagly et al. (2003) show small gender differences in leadership styles. The authors find that female leaders engage more often than male leaders in transformational leadership and in contingent reward behavior as part of the transactional leadership style.

Prediction 4*: Male leaders send assertive, goal-oriented messages and clear orders more often than their female counterparts. Female leaders send more often relevant messages containing mutual benefit arguments, along with encouraging, positive emotional content and expressions emphasizing that the leader is an equal member of the group (using the personal pronoun “we” more often). We*

expect that these gender differences in communication and leadership styles are stable across communication conditions.

2.3 Procedures

The experiment was conducted at the Columbia Experimental Laboratory in the Social Sciences (CELSS) at Columbia University. Participants were recruited through ORSEE (Greiner 2015) and the experiment was programmed with z-Tree (Fischbacher 2007). A session lasted around 60 minutes. We used standard experimental procedures, including random assignment of subjects to treatments, anonymity, detailed instructions with control questions, dividers between the subjects' cubicles, and monetary incentives. Earnings were expressed in points and were converted to dollars at a rate of $1 per 345 points. Average earnings equaled $15.75. Detailed experimental procedures are available in the Appendix.

3 Results

In total, 165 participants took part in the study, of which 67 were male and 98 were female. In *One-way*, we had 80 participants in 16 groups with 8 male and 8 female leaders. In *Two-way*, we had 85 participants in 17 groups with 8 male and 9 female leaders.

Subsection 3.1 analyzes the treatment effects on leader effectiveness, subsection 3.2 analyzes the leader behavior, and subsection 3.3 the follower behavior and evaluation of leaders.

3.1 Treatment Effects

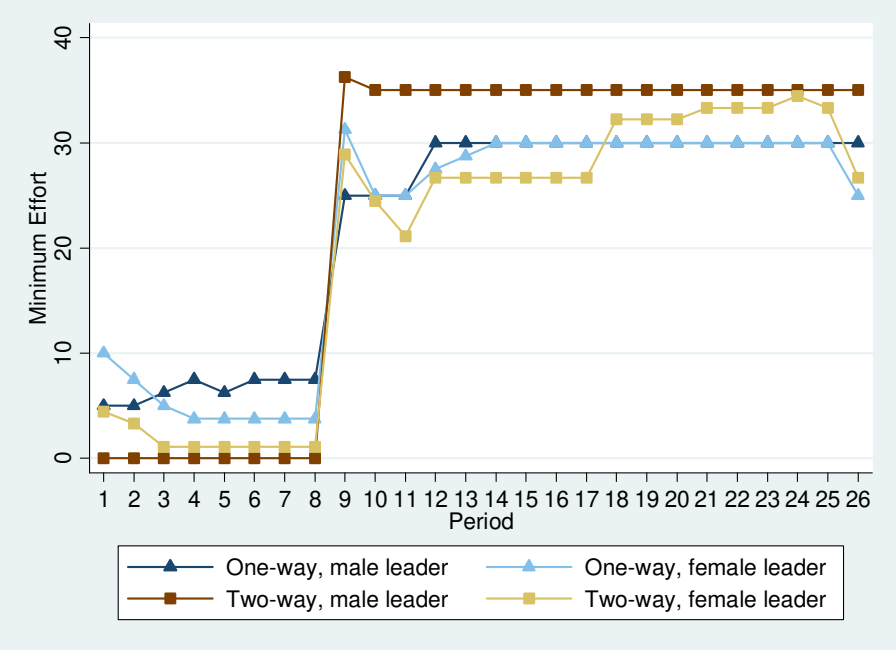
We start our analysis by looking at the treatment effects on leader effectiveness in the four different treatment conditions. Figure 1 depicts the group minimum effort by treatment conditions in all periods. In the initial periods 1 to 8, without a leader, 31 of the 33 groups converged to the most inefficient effort level, so the first part of the turnaround game replicated previous findings and reliably induced failure to coordinate on the efficient equilibrium (Brandts and Cooper 2006a). The second part of the turnaround game started in period 9, when all leaders successfully managed the turnaround on average at least above the minimum effort level of 20. Additional summary statistics are available in the Appendix.

Graphically, male leaders who receive follower messages are more effective than their female counterparts, or leaders in *One-way*. However, if we test whether these differences are statistically significant, we cannot confirm that groups in *Two-way* do better with male leaders than with female leaders ($p = 0.39$), or that groups with male leaders do better in *Two-way* than in *One-way* ($p = 0.44$). Similarly, groups in *One-way* do not differ depending on the leader's gender ($p = 0.99$). Moreover, we cannot confirm that groups in *Two-way* do better in general than in *One-way*, holding constant that leaders request the highest effort ($p = 0.65$).[7] In other words, we do not replicate previous findings that incoming follower messages in *Two-way* increase leader credibility, compared to *One-way* (Kriss

[7] Post-estimation test results after random effects GLS regressions for minimum effort. The regressions use treatment×leader's gender dummy variables and cluster standard errors on groups. All the regressions in the paper are available in the Appendix.

and Eil 2012). Finally, groups with female leaders do not differ depending on the communication treatment ($p = 0.98$).

Figure 1 Treatment effects by treatment conditions



***Result 1**: We find no support for Prediction 3. In both One-way and Two-way, men and women are equally effective leaders. In Two-way, although there is a difference in the leader effectiveness of men and women, which is noticeable in economic terms, the gender gap is far from significant.*

3.2 Leader Behavior

This subsection compares leader behavior across treatment conditions. A detailed description of the message content categories based on the scheme in Brandts et al. (2015) and a correlation analysis of the most often occurring categories are available in the Appendix. The subsequent analysis focuses on often occurring message categories and their combination, which are requesting the highest effort level (forty), assertive style, orders, expressions of being part of the group (coequality), positive emotional content, relevant (transformational) content, and transactional content. We exclude the mutual benefit argument from the subsequent analysis, because it occurs rarely and is highly correlated (0.49) with expressions of being part of the group. All variables are binary variables, taking the value of 1 if the message contains the certain category and 0 otherwise. For the combined content variables we use the sum of the codings. For example, if we combine forty with assertive style, we create a new binary variable that takes the value of 1 if the message contains both the suggestion of forty and is assertive and 0 otherwise. The suggestion to choose forty is positively correlated with all other categories, so that in relation to a suggestion of forty, none of the message categories has an independent effect. The highest correlation coefficient (0.72) is that of forty and the assertive style, which can be part both of the transformational and transactional leadership style. The combined variable for the relevant content contains forty, assertive style, and being part of the

group. The combined variable for the transactional content contains forty, assertive style, positive emotional content, and orders. Besides the fact that positive emotional content and orders can be viewed as part of the transactional leadership style, both variables are highly correlated with forty (positive emotional content with 0.46 and orders with 0.30) and assertive style (positive emotional content with 0.32 and orders with 0.45), while lowly correlated with each other (0.11).

Figure 2 depicts the frequency of requesting forty, assertive style, being part of the group, orders, relevant content, and transactional content by treatment conditions in leader messages in every message period. Men seem to request slightly more often forty and communicate more assertively than women. If we test whether these differences are statistically significant, we confirm that men communicate more assertively than women ($p = 0.05$ irrespective of the communication treatment, $p = 0.16$ in *One-way*, and $p = 0.15$ in *Two-way*) and the difference is around 0.23 percentage points. Moreover, leaders use the assertive style in *Two-way* more often than in *One-way* ($p = 0.07$) and the difference is around 0.21 percentage points.[8] However, we cannot statistically confirm any differences with respect to forty. Men and women are equally likely to request the highest effort level, in both *One-way* and *Two-way*. Leaders are slightly more likely to request forty in *Two-way* rather than in *One-way*, though the difference is noticeable in economic terms (around 0.16 percentage points), it is not statistically significant ($p = 0.15$).

Women express that they are part of the group more often in *Two-way* than in *One-way* ($p = 0.03$), while men use orders slightly more often than women, although this latter difference is not statistically significant ($p = 0.17$). Leaders are more likely to send positive emotional content in *Two-way* rather than in *One-way* ($p = 0.04$) and the difference is around 0.22 percentage points. It seems that incoming follower messages trigger the contingent reward behavior of the transactional leadership style, but we find no gender differences with respect to positive emotional content.

We find no statistically significant differences in the use of relevant content across treatment conditions. Men and women are equally likely to use the transformational style in both *One-way* and *Two-way*. As to the use of the transactional content, Figure 2 already suggests sharp differences. Leaders are more likely to use the transactional content in *Two-way* rather than in *One-way* ($p = 0.00$) and the difference is around 0.22 percentage points, which is most likely due to the positive emotional content component in the combined variable. Moreover, men are more likely to use the transactional content than women ($p = 0.02$) and the difference is around 0.15 percentage points, in both *One-way* and *Two-way*.

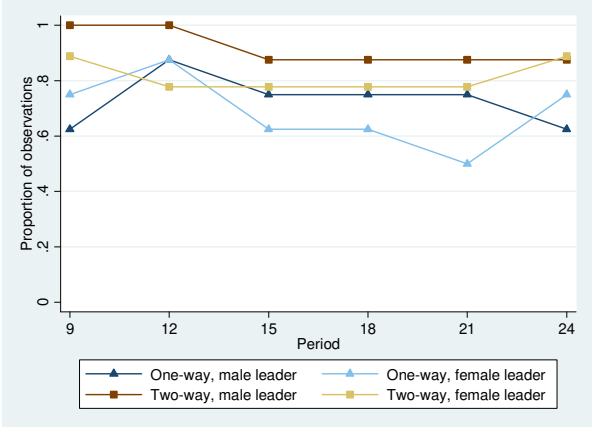
Figure 2 also reveals some trends in the use of message contents over time. Initially, probably until high effort levels are reached, leaders are likely to express that they are part of the group, which represents an important feature of the transformational style. Over time both the expressions of being part of the group and the relevant content decrease. The opposite is true for sending orders and the transactional content. Initially, the frequency of orders, which represents one feature of the transactional style, is very low. After establishment of a most likely successful group history, the frequency of orders and transactional content increases, as if leaders would increasingly handle their task by simple messages like "keep choosing 40". Of course, this is related to our design. After a successful turnaround, it is unnecessary for credible leaders to repeatedly suggest the same.

[8] Random effects GLS regressions for message content variables with treatment×leader's gender dummy variables and standard errors clustered on groups (see the Appendix).

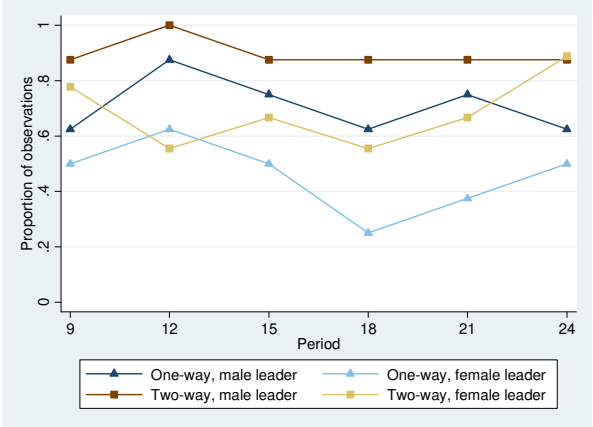
Similarly, in real life, teams that face less novelty during the production process might require the transactional, rather than the innovative transformational leadership style to sustain high performance.

Figure 2 Frequency of content in leader messages

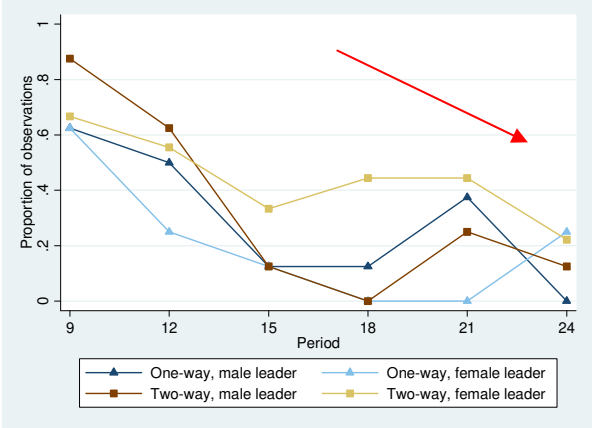
Suggestion to choose forty



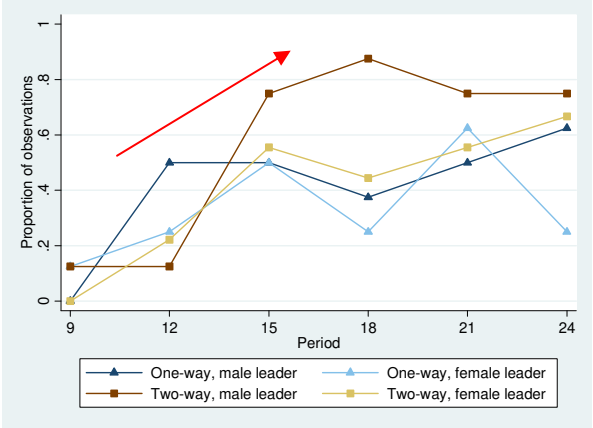
Assertive style



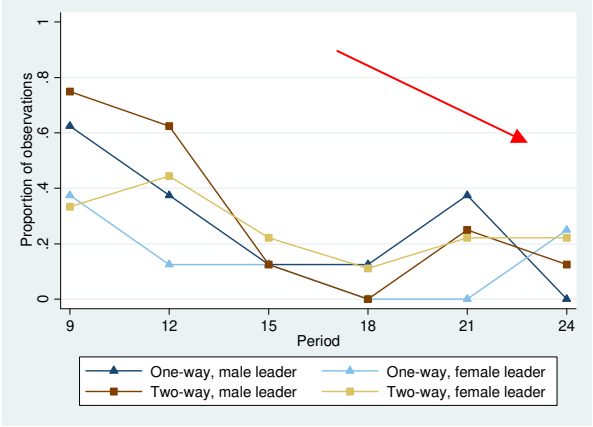
Being part of the group



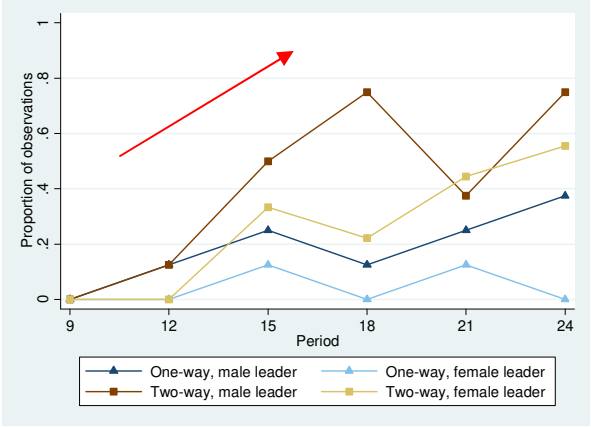
Orders



Relevant content



Transactional content



Next, we briefly report results of the analysis of follower message content, which might affect the content of leader messages in *Two-way*. The most often occurring content in follower messages is forty (0.59), no message at all (0.33), and positive emotional content (0.28).[9] Followers do not

[9] See the frequency of message categories in the Appendix.

differentiate between male and female leaders when they send forty (p = 0.51), no message at all (p = 0.96), or positive emotional content (p = 0.97).[10] The followers' suggestion of forty has a significant effect on the likelihood of leaders requesting forty (p = 0.06).[11] Female leaders depend more on the followers' suggestion of forty than male leaders (p = 0.07), as female leaders are by around 0.17 more likely to request forty if they received incoming forties, rather than another content or no message, from their followers (p = 0.11).[12] Although the effect of the followers' suggestion of forty is transmitted to the combined variables (relevant and transactional content), male and female leaders do not differ in their reactions to the incoming forties. It seems that leadership styles are stable traits, which cannot be influenced easily by followers. Whether followers send any message has no statistically significant effect on leaders' messages, neither on forty, nor on relevant or transactional content. If followers send positive emotional content, leaders are by around 0.08 more likely to request forty (p = 0.02). Followers' positive emotional content does not affect the likelihood of relevant content (p = 0.63), but it increases by around 0.20 the likelihood of transactional content (p = 0.00) in leader messages, no matter if the leader is male or female. There might be an emotional contagion between followers and leaders: positive follower emotions might increase positive leader emotions, which then transmit to the combined transactional content.

Result 2*: We find no support for Prediction 2, as male and female leaders are equally likely to request forty in both One-way and Two-way. We confirm the main points of Prediction 4. Men communicate more assertively than women. In Two-way, rather than One-way, women emphasize more often that the leader is an equal member of the group. The transformational leadership style is used equally often by men and women, but the transactional leadership style is used more often by men.*

3.3 Follower Behavior

This subsection analyzes leader credibility, which is the followers' reaction to the leaders' messages. Figure 3 compares the followers' reactions, measured by the minimum of followers' effort choices, to the leaders' message content in each period. The most interesting message contents are forty, assertive, relevant, and transactional content. Forty is important as it is the main driver of leader effectiveness, and as such it is part of the relevant and transactional content. The relevant content is the most effective style, especially in early periods, because it results in the highest minimum of followers' effort and reaches 40 already by period 12, as the figure shows. However, leaders can

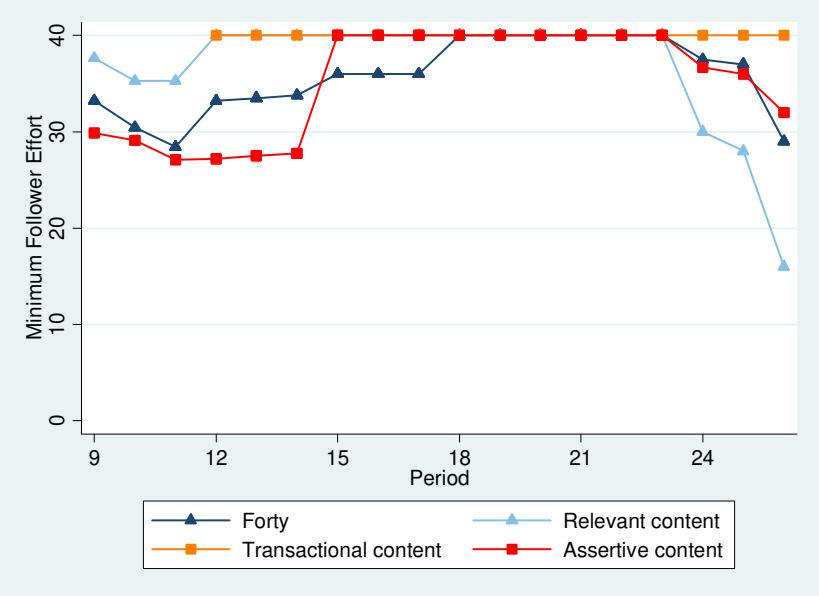
[10] Random effects GLS regressions for follower message content variables with a leader's gender dummy variable and standard errors clustered on groups (not reported in the Appendix).

[11] Post-estimation test results after random effects GLS regressions for leader message content (such as forty, relevant, and transactional content). The regressions use leader's gender× follower message content (forty, no message, and positive emotional content) dummy variables and cluster standard errors on groups (not reported in the Appendix).

[12] In period 9, followers send suggestions around 35 on average. As a reply, leaders send suggestions close to 40. In period 12, followers start to follow the communication of male leaders, and their suggestions converge to 40 over time, just in line with male leaders continuously suggesting 40. On the other hand, the communication of female leaders is not followed that easily. In period 12, followers suggest on average 30, and female leaders start to adapt their suggestions to this trend. In period 15, as if followers would recognize the importance of their input, they start to send higher suggestions, and the average converges to 35. From period 18 onwards, female leaders regain trust into their followers, and their suggestions converge to follower suggestions. Although female leaders initially recognized their pivotal role and ignored low follower suggestions, it seems that they had a harder time to gain the attention of their followers (see the figures on the distribution of the message values in the Appendix).

switch among leadership styles, and this might explain why we observe a decrease in the minimum of followers' effort in periods 24 to 26, as leaders of low performing groups might try a final turnaround using the transformational leadership style. The transactional content is an effective style as well, but one reason for this might be that the leaders use it only in case of successfully performing groups, which also explains why the style is used only from period 12 onwards. The assertive content alone, without requesting forty, is not enough to be as effective as leaders who request forty, as the figure shows in early periods. In later periods, it is only effective leaders who remain assertive, others do not communicate assertively anymore.

Figure 3 Comparison of followers' reactions by leaders' message content over time



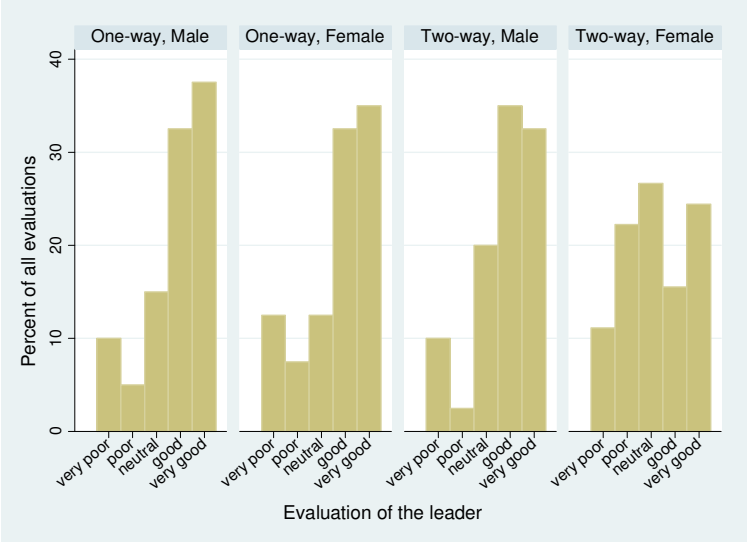
Next, we focus on leader messages that contain the request of forty, and we analyze treatment effects on leader credibility, using followers' individual effort choices. We find no statistically significant differences between male and female leaders, and *One-way* and *Two-way*, or a mixture of these treatment conditions.[13] Men and women who request forty are equally credible leaders. To test the robustness of this results, we repeat our analysis and restrict the data to observations of groups in periods 9 to 11 who did not coordinate on 40 in periods 5 to 8 (which was the case for 1 group), and to observations of groups in periods 12 to 26 if the group's average minimum effort in the previous message cycle was below 40, which means that leaders were previously not effective. We do this, because leader credibility is best observed after a previous history of coordination failure, when leaders can try the turnaround. Again, we find no statistically significant differences across treatment conditions.

Finally, we briefly report about leader evaluations, which we collected in the final questionnaire, explained in more detail in the Appendix. Figure 4 depicts the evaluations in percentages by treatment conditions. To ease comprehension, we interpret the original answer keys as "very poor" for 1, "poor" for 2, "average" for 3, "good" for 4, and "very good" for 5. On average, male leaders get an evaluation of 3.8, and female leaders get 3.4. With 3.2, female leaders in *Two-way* get the lowest average evaluation. If we test whether these differences are statistically significant, we confirm that

[13] Post-estimation test results after random effects GLS regressions for followers' effort. The regressions use treatment×leader's gender dummy variables and cluster standard errors on groups (see the Appendix).

female leaders in *Two-way*, who requested forty and whose groups succeeded to coordinate on forty, receive worse evaluations than their male counterparts (p = 0.04). Interestingly, female leaders in *Two-way*, who did not request forty, receive better evaluations than their male counterparts (p = 0.03).[14] In other words, followers in *Two-way* evaluate women harsher than men, if leaders were successful, but they blame women less than men, if leaders did not request the highest effort, which might be because followers perceive a greater influence on women rather than men if leaders have an "open door" for follower messages.

Figure 4 Evaluation of the leader in percentages by treatment conditions



***Result 3**: We find no support for Prediction 1. Men and women are equally credible leaders in both One-way and Two-way. Despite the differences in the communication paths of the leaders, followers do not differentiate between leaders based on their gender. Relevant message content is the most effective content.*

4 Conclusion and Discussion

Men communicate more assertively than women. If women have an "open door", they emphasize more often that the leader is an equal member of the group. Despite the differences in the communication paths of the leaders, men and women are equally likely to use the transformational leadership style, which relates positively to leader effectiveness. Men and women are equally effective leaders, who are equally credible and equally likely to request forty. We find a weak hint that women with an "open door" might have it harder to make their followers adapt to them, as followers might perceive to have a stronger influence on women rather than men. However, over time, women gain credibility and their groups become successful. As one of the female leaders expressed it "everyone is on board now", after she finally managed in the last periods that her group simultaneously chose the most efficient effort level.

[14] OLS regressions for the leader's evaluation with the leader's gender dummy variable and standard errors clustered on groups (not reported in the Appendix).

One shortcoming of our data is the relatively low number of observations and that our data might be over-fitted. Except for one male leader who sent once a suggestion of zero, all male leaders always requested forty if they sent any message. Thus, the limitation of our data is that we have no variation in the observations for male leaders. We cannot compare the whole spectrum of male to female behavior, or the impact of message value within gender. What we can compare is male leaders requesting forty to female leaders requesting forty, and follower behavior in these cases. Another shortcoming might be the presence of a negative experimenter demand effect (Zizzo 2010). Such an effect would go into the opposite direction than our prediction of male leaders being followed to a greater extent than female leaders, thus leaving room for an alternative interpretation of our results that does not rule out that men might be more effective leaders than women.

As pointed out in Eagly et al. (2003), the transformational leadership style is independent from the underlying organizational structure. In our study, the organizational structure is rather flat. The leader is closely tied to the group, and faces the same incentives as other group members. These circumstances might ease the use of the transformational leadership style for both men and women, and therefore might not bring gender differences onto the surface. A more competitive setting or a stronger hierarchy might magnify male-dominated leadership qualities. Nevertheless, due to the findings in our present study we see no reason why women should not be empowered to the same extent like men. Women have a different style to communicate and to lead, and since we showed that followers adapt over time to their leaders, we have reason to expect that an increase in the number of female leaders could also affect the structure and culture of organizations on the long run. Such a "feminine" development might lead to more balanced, more diverse, and more effective leader-follower interactions within organizations.

References

Akerlof GA, Kranton RE (2005) Identity and the economics of organizations. *Journal of Economic Perspectives* 19:9–32.

Antonakis J, d'Adda G, Weber R, Zehnder C (2015) "Just Words? Just Speeches?" On The Economic Value of Charismatic Leadership. *Working Paper*.

Avolio BJ (1999) Full leadership development: Building the vital forces in organizations. Thousands Oaks, CA: Sage.

Bass BM (1985) Leadership and performance beyond expectations. *Free Press*, New York

Bass BM (1998) Transformational leadership: Industry, military, and educational impact. *Erlbaum*, Mahwah, NJ

Bass BM, Avolio BJ, Atwater L (1996) The transformational and transactional leadership of men and women. *Applied Psychology: An International Review*, 45: 5-34.

Bem SL (1974) The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42:* 155-162.

Bem SL (1993) The lenses of gender: Transforming the debate on sexual inequality. *Yale University Press*.

Brandts J, Cooper DJ (2006a) A change would do you good: An experimental study of how to overcome coordination failure in organizations. *Amer. Econom. Rev*. 96(3): 669–693.

Brandts J, Cooper DJ (2006b) Observability and overcoming coordination failure in organizations: An experimental study. *Experiment. Econom*. 9(4): 407–423.

Brandts J, Cooper DJ (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.

Brandts J, Cooper DJ, Weber RA (2015) Legitimacy, Communication, and Leadership in the Turnaround Game. *Management Science* 61(11): 2627-2645.

Burns JM (1978) Leadership. *Harper & Row*

Camerer CF (2015) The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Critical Reply to Levitt and List. In *Handbook of Experimental Economic Methodology*, Oxford University Press, 249-95.

Conger JA, Kanungo RN (1998) Charismatic leadership in organizations. Thousand Oaks, CA: Sage.

Cooper DJ (2007) Are Experienced Managers Experts at Overcoming Coordination Failure? *The B.E. Journal of Economic Analysis & Policy* 6(2) Article 6.

Cooper DJ, Hamman J, Weber RA (2016) Fool Me Once: An Experiment on Credibility and Leadership. Available online at SSRN.

Day, DV (Ed.) (2014) The Oxford handbook of leadership and organizations. *Oxford University Press*

Devetag G, Ortmann A (2007) When and why? A critical survey on coordination failure in the laboratory. *Experimental Economics*, 10(2):171–178.

Dezső CL, Ross DG (2012) Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, 33: 1072-1089.

Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG (2011) Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, 9(3), 522–550.

Eagly AH, Johannesen-Schmidt MC, Van Engen ML (2003) Transformational, Transactional, and Laissez-Faire Leadership Styles: A Meta-Analysis Comparing Women and Men. *Psychological Bulletin* 129 (4): 569–591.

Eagly AH, Johannesen-Schmidt M (2007) Leadership style matters: The small, but important, style differences between male and female leaders. In D. Bilmoria & S. K. Piderit (Eds.), Handbook on women in business and management (pp. 279-303). Northampton, MA: Edward Elgar.

Eurostat (2014) Gender pay gap statistics

Fischbacher U (2007) z-Tree: Zurich Toolbox for Ready-Made Economics Experiments. *Experimental Economics* 10: 171-8.

Foels R, Driskell JE, Mullen B, Salas E (2000) The effects of democratic leadership on group member satisfaction: An integration. *Small Group Research*, 31: 676-701.

Gastil J (1994) A meta-analytic review of the productivity and satisfaction of democratic and autocratic leadership. *Small Group Research*, 25: 384-410.

Goldin C (2002) A pollution theory of discrimination: male and female differences in occupations and earnings. *NBER Working Paper* No. 8985.

Greiner B (2015) Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE, *Journal of the Economic Science Association* 1 (1): 114-125.

Grossman PJ, Eckel C, Komai M, Zhan W (2016) Gender, Coordination, and Reward in a Leadership Game. Working Paper.

Heath C, Staudenmayer N (2000) Coordination neglect: How lay theories of organizing complicate coordination in organizations. *Research in Organizational Behavior*, 22: 153-192.

Heinemann F, Nagel R, Ockenfels P (2009) Measuring Strategic Uncertainty in Coordination Games, *Review of Economic Studies*, 76 (1): 181-221.

ILO (2013) Key Indicators on the Labour Market: 7th edition. Geneva: ILO.

Kriss PH, Blume A, Weber R (*forthcoming-b*) Coordination with Decentralized Costly Communication. *Journal of Economic Behavior and Organization*

Kriss PH, Eil D (2012) Managers as coordination devices: Keep your door open, but your skin thick. In: *Essays on the Impact of Costly Communication on Coordination*, Dissertation by Peter Kriss. Carnegie Mellon University, Pittsburgh, pp. 106–140.

Lewin K, Lippitt R (1938) An experimental approach to the study of autocracy and democracy: A preliminary note. *Sociometry*, 1: 292-300.

Lowe KB, Kroeck KG, Sivasubramaniam N (1996) Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *Leadership Quarterly*, 7: 385-425.

Merchant K (2012) How Men and Women Differ: Gender Differences in Communication Styles, Influence Tactics and Leadership Styles. *CMC Senior Theses*. Paper 513.

OECD (2016) Labour Market Statistics: Labour force statistics by sex and age: indicators. *OECD Employment and Labour Market Statistics* (database).

OECD (2014) Employment Database

Reuben E, Rey-Biel P, Sapienza P, Zingales L (2012) The emergence of male leadership in competitive environments. *Journal of Economic Behavior & Organization* 83(1): 111-117.

Reuben E, Sapienza P, Zingales L (2014) How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences* 111 (12): 4403-4408.

Reuben E, Wiswall M, Zafar B (in press) Preferences and Biases in Educational Choices and Labor Market Expectations: Shirking the Black Box of Gender. *Economic Journal*

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

Sharpe R (2000) As leaders, women rule: New studies find that female managers outshine their male counterparts in almost every measure. *Business Week*.

United Nations (2015) A/RES/70/1 - Transforming our world: the 2030 Agenda for Sustainable Development. *United Nations - Sustainable Development knowledge platform*.

Van Huyck JB, Battalio RC, Beil RO (1990) Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. *American Economic Review* 80(1): 234-248.

Vroom VH, and Yetton PW (1973) Leadership and decision-making. Pittsburgh, PA: University of Pittsburgh Press.

Wang G, Oh IS, Courtright SH, Colbert AE (2011) Transformational leadership and performance across criteria and levels: A meta-analytic review of 25 years of research. *Group & Organization Management*, 36: 223-270.

World Economic Forum (2015) The Global Gender Gap Report.

Zizzo DJ (2010) Experimenter Demand Effects in Economic Experiments. *Experimental Economics* 13(1): 75-98.

Appendix

The first section of the Appendix contains more detail about the experiment's procedures, including a detailed timeline, a sample of the instructions, and screenshots of the computer program. The second section contains descriptive statistics as well as the regressions reported in the main body of the paper.

A.1 Detailed experimental procedures

After their arrival to the laboratory, participants were assigned randomly to seats. Before period 1, and before reading the instructions, everybody answered a short general questionnaire about gender, race, age, years of study, and major field of studies. Next, participants had to choose a profile picture. Figure A1 contains the 24 profile pictures.

Figure A1 Available profile pictures to female (left) and male (right) participants

We had separate instructions for Part 1 and Part 2, and participants read the instructions only prior to each part. To facilitate calculations for the participants, we handed out printed versions of the instructions for Part 1, which contained the Earnings Table showing how earnings were determined in each period. The same table applied in Part 2. Instructions were displayed on the computer screens and were read aloud by the experimenter. After reading the instructions for Part 1, participants completed a payoff quiz to check whether everybody understood the game's payoff structure. Instructions and screenshots can be found below.

The game was described using a workplace context to be in line with earlier papers, ease comprehension of the task, and enrich the wording and analysis of the free form messages (Cooper 2007; Brandts et al. 2015). As in Brandts et al. (2015), individual group members were referred to as "employees", and they were told that they were working for a "firm". The leader was called the "manager". Following Brandts et al. (2015), we did not use the term "effort" because of its strong connotation. Instead we asked participants to think of each period as a "workweek" lasting 40 hours and choose how many hours to devote to the firm's "bonus project".

Participants could enter their messages into a chat box, and they could either click on a button to send the message or click on a button labeled as "Send no suggestion". The leader's message, along

with the leader's profile picture, was displayed on all screens throughout the three periods of a message cycle. In *Two-way*, followers could see the profile picture of the leader already on their message entering screens and follower messages were displayed without followers' profile pictures and only on the leader's message screen. Participants knew from the instructions of Part 2 that the profile picture of the leader will be displayed.

In Part 2, after participants made their effort choice, we elicited their belief concerning the number of other group members who will follow the leader's message by asking "Out of the four other participants in your firm, how many will follow the Manager's suggestion?" Participants could enter guesses from 0 to 4, or they could indicate "Not applicable" for cases where the leader did not make any suggestion. Belief questions were only asked in the first period of each message cycle. Table A1 summarizes the sequence of events in the experiment.

Table A1 Timeline of the experiment

	One-way communication	**Two-way communication**
Before period 1	Demographics and choice of profile picture	
	Instructions for Part 1 and payoff quiz	
Periods 1-8	Effort choice	
	Feedback screen	
Before period 9	Instructions for Part 2	
	Random assignment of the leader role	
Periods 9-26		Followers send message to the leader (every third period)
	Leader sends message to the followers (every third period)	
	Effort choice	
	Belief question (every third period)	
	Feedback screen	
After period 26	Final questionnaire	

At the end of each period, participants saw their effort choice, the group minimum effort, their earnings in that period, and their accumulated earnings. Participants could not observe individual effort choices.

At the end of the experiment participants filled in a final questionnaire. We asked which role assignment ("Manager" or "Employee") participants would prefer if they could choose ("If you were to play Part 2 again and you could choose your role, which role would you choose?"). After reminding the individual average earnings of the participant in Part 2, we asked for the subjective evaluation about the performance of the leader on a five-point scale from "completely disagree" to "completely agree" ("How much do you agree/disagree with the following statement: My firm's performance in Part 2 is mostly due to the judgment of the Manager."). Finally, we asked a general risk attitude question ("How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?"), where participants could indicate their willingness to take risks on a scale from 0 for "not at all willing to take risks" to 10 for "very willing to take risks" (Dohmen et al. 2011).

At the end of the experiment, participants were shown their earnings separately for each part and in total. Participants were thanked and paid individually for their participation.

A.1.1 Sample instructions

Thank you for participating in this session. You are participating in a study on economic decision making and will be asked to make a number of decisions. Please read these instructions carefully as they describe how you can earn money. All the interaction between you and other participants will take place through the computers. Please do not talk or communicate in any other way with other participants. If you have a question, raise your hand and one of us will help you. The study is anonymous: that is, your identity will not be revealed to others and the identity of others will not be revealed to you. During the study your earnings will be expressed in points. Upon completion of the session, your accumulated earnings will be converted from points to dollars at a rate of $1 per 345 points. You will be paid these converted earnings in cash.

The study is divided into two parts. Your earnings today will equal the sum of earnings from each part. You will be randomly assigned to a firm of five participants. You will be grouped with the same five participants throughout Part 1 and Part 2. Part 1 consists of 8 periods and Part 2 consists of 18 periods. You will read the instructions for Part 1 below. You will receive the instructions for Part 2 once Part 1 has been completed.

Part 1

You are one out of five employees in a firm. Each period can be thought of as a workweek. Each employee spends 40 hours per week at their firm. In each period, every employee will be asked to choose how many hours to devote to the firm's bonus project. The available choices are 0 hours, 10 hours, 20 hours, 30 hours, and 40 hours. The earnings for an employee are determined in each period by how many hours that employee spends on the bonus project, and the minimum number of hours employees in his or her firm spend on the bonus project. Specifically, the employee's earnings are reduced by 5 points per hour that he or she spends on the bonus project. In addition, the employee also receives a bonus equal to the minimum number of hours any employee in his or her firm spends on the bonus project multiplied by 6 points. Each employee also gets a flat payoff of 200 points in each period. In other words, your earnings are given by the formula below:

200 – 5 × your hours in bonus project + 6 × minimum hours in bonus project by any employee

To facilitate your calculations, the following Earnings Table shows how your earnings depend on your choice and the minimum choice in your firm.

EARNINGS TABLE

		Minimum number of hours chosen in the firm				
		40	**30**	**20**	**10**	**0**
Your hours	**40**	240	180	120	60	0
	30		230	170	110	50
	20			220	160	100
	10				210	150
	0					200

Your earnings in each period are found by looking across from the number of hours you chose on the left-hand side and down from the minimum number of hours chosen in the firm by any employee. For example, suppose you spend 10 hours on the bonus project. Suppose the other four employees in the firm spend 20, 30, 40 and 40 hours. The minimum hours spent on the bonus project is 10 hours. Then your payoff equals: 200 - 5×10 + 6×10 = 210 points.

At the end of each period you will receive a summary of what happened in the period including the number of hours you spent on the bonus project, the minimum number of hours chosen in the firm, your payoff for the latest period, and your accumulated payoffs for the current part. The computer also provides a summary of this information for preceding periods. At no point in time will we identify the identity of any employees in the firm. In other words, the actions you take will remain confidential. To ensure your understanding of these instructions, click the "READY" button and answer the questions that will appear on your screen.

Part 2

Part 1 has ended. Read the instructions for Part 2 and click on READY once you are done. Part 2 is similar to Part 1. Now there will be 18 periods in Part 2. In each period, every employee will choose how many hours to devote to the firm's bonus project. Available choices are 0, 10, 20, 30 and 40 hours. The number of hours you choose and the minimum number of hours chosen in the firm will determine your earnings in that period. The Earnings Table is the same as in Part 1. Finally, your firm's composition has not changed. In other words, in Part 2 you will interact with the same firm of five people as in Part 1.

The difference between Part 1 and Part 2 is that one person in your firm will be randomly selected to be the Manager. In the next screen, you will be informed whether you have been assigned the role of the Manager or the role of an Employee for all of Part 2. Throughout Part 2, the Manager's profile picture will be displayed on the computer screen.

Every 3 periods, before everyone chooses a number of hours, there will be a message stage. In the message stage, first the employees will be able to send a written suggestion to the Manager. Only the Manager will be able to see these suggestions, other employees will not be able to see them. After the employees made their suggestions to the Manager, the Manager will be able to send a written suggestion to all employees. Both employees and Managers can alternatively decide not to send any suggestion. The suggestion cannot contain information that can be used to identify participants, such as a name, nickname, or any other identifying feature like clothing, or the desk number. Other than these restrictions, participants may write anything that they wish.

After the message stage, every employee will see the Manager's suggestion. Subsequently, each employee and the Manager enter the number of hours they wish to choose. Note that the suggestion does not commit you to any particular choice. That is, neither the Manager nor the other employees are required to choose the number of hours indicated in the suggestion.

A.1.2 Screenshots

You are the Manager Period: 2

MESSAGE STAGE

Suggestions from the employees:

Employee 1: let's choose 40
Employee 2: if we all choose 40, we can all earn the maximum
Employee 4: please suggest 40

Write your suggestion to all employees in the box below.

Remember to hit first the [Enter] key before you click on FINISHED SENDING SUGGESTIONS!
If you wish to send "No suggestion", click on SEND NO SUGGESTION and confirm with YES.

SEND NO SUGGESTION | FINISHED SENDING SUGGESTIONS

You are the Manager Period: 2

DECISION STAGE

The suggestion from the Manager to all Employees is:

pick 40

Please enter your hours
- 40
- 30
- 20
- 10
- 0

SUBMIT

A.2 Additional statistical analysis

Table A2 contains descriptive statistics of the leaders' suggestion (whether the leader sent any message, the message value, and the frequency of requesting forty), the followers' message value, the groups' minimum effort, the effort of individual participants, the individual earnings, and the individual belief about the share of followers depending on the communication treatment and the gender of the leader. The table reports the mean, standard deviation (in parentheses), and p-values of non-parametric Wilcoxon rank-sum test results. We test for treatment differences with the null hypothesis that the two samples come from the same population. The message value variable contains only positive numeric suggestions and has missing values if no message was sent. The belief variable contains missing values if participants indicated that the question was not applicable. The data contains observations from all 18 periods of Part 2, except in case of message and belief variables, which are collected only every third period. Since observations are history-dependent, we collapse by group.

Table A2 Descriptive statistics

	One-way communication		*Two-way communication*	
	Male leader	Female leader	Male leader	Female leader
Sent no message	0.25 (0.37)	0.19 (0.26)	0.08 (0.24)	0.02 (0.06)
	p = 0.82		p = 0.86	
	Males: p = 0.12; Females: p = 0.07			
Message value	35 (14.14)	34.58 (10.07)	40 (0)	34.59 (8.11)
	p = 0.64		p = 0.04	
	Males: p = 0.32; Females: p = 0.65			
Forty	0.73 (0.41)	0.69 (0.37)	0.92 (0.24)	0.81 (0.26)
	p = 0.70		p = 0.23	
	Males: p = 0.13; Females: p = 0.44			
Follower message value			37.11 (5.82)	32.79 (10.47)
			p = 0.49	
			Conover squared-rank variance test: p = 0.05	
Minimum effort	29.17 (18.15)	29.03 (17.20)	35.07 (13.95)	29.20 (15.44)
	p = 0.74		p = 0.06	
	Males: p = 0.24; Females: p = 0.88			
Individual effort	29.57 (17.68)	31.22 (14.64)	35.50 (12.73)	31.38 (13.32)
	p = 0.78		p = 0.06	
	Males: p = 0.24; Females: p = 0.88			
Earnings	227.15 (20.69)	218.06 (30.33)	232.92 (20.03)	218.27 (27.27)
	p = 0.40		p = 0.05	
	Males: p = 0.37; Females: p = 0.88			
Belief about the share of followers	0.89 (0.16)	0.89 (0.20)	0.96 (0.08)	0.92 (0.13)
	p = 0.83		p = 0.59	
	Males: p = 0.29; Females: p = 0.70			

Table A3 presents estimates from random effects GLS regressions testing the effect of the communication treatment and the leader's gender on group coordination. In all regressions, the dependent variable, in each period, equals the group minimum effort. In column (1) and (2), the independent variables are the communication treatment (*One-way* or *Two-way*) and the gender of the leader (male of female). In column (2), we restrict the model of column (1) to observations when leaders requested the highest effort level. In column (3), the independent variables are the interaction of the communication treatment (*One-way* or *Two-way*) and the gender of the leader (male or female). In all regressions, we cluster standard errors on groups and include data from all periods. Note that the omitted category is groups with a male leader in *One-way*.

Table A3 Treatment effects

Dependent variable: minimum effort	(1)	(2)	(3)
Two-way communication	2.95	-2.28	5.90
	(5.44)	(5.01)	(7.71)
Female leader	-3.09	-3.70	-0.14
	(5.41)	(5.00)	(8.42)
Two-way, female leader			-5.73
			(10.83)
Constant	30.64***	36.12***	29.17***
	(5.10)	(4.48)	(6.11)
Obs.	594	483	594
Clusters	33	31	33
Wald χ^2	1	1	1

Notes. Random effects GLS regressions. Standard errors (in parentheses) are corrected for clustering on groups. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

Table A4 contains the full list of all coded message content categories, indicating description and examples, and the percentage (frequency) of all observations including cases when the leader sent no message. All variables are coded as binary: variables take the value of 1 if the message contains a certain category and 0 otherwise. The message coding was carried out by the author and is available upon request.

The content scheme is based on the categories in Brandts et al. (2015). We distinguish ambiguous and indecisive content, however such messages are rare. Ambiguous content is ambiguous to interpret concerning its numeric value. Indecisive content is clear to interpret, but the leader communicates own doubts, so that the message contains several parallel options, for example "choose 20 or higher" or "choose 30 or 40". Such content is interpreted as indecisive, but not necessarily as ambiguous, if we assume that participants are risk averse, and choose the lower suggestion when receiving the uncertain message content. For example, if the suggestion says to choose "20 or higher", we coded the numeric message content as 20, not ambiguous, but indecisive; "work more hours" is interpreted as 40, ambiguous, but not indecisive (since the leader might have had a clear goal in mind, but communicated in a way that is ambiguous to interpret); and "devote the same number of hours" without any history in period 9 is interpreted as 0, ambiguous, but not indecisive. We include new categories such as the conditional strategy, which makes sense if leaders have to send messages for three sequential periods. We do not include the category for appeals to mutual trust, and instead of including categories for positive and negative feedback, we include categories for positive and negative emotional content. Such content might refer to previous

performance, for example, thanking others or praising team work, but it can also refer to the future choice, for example, encouraging words or desperate words spreading uncertainty.

Several categories are coded only for leaders. Based on the insights in Cooper (2007) and in Brandts et al. (2015), we code especially for the mutual benefit argument if an explanation emphasizes collective action and the resulting benefit for each group member. We code whether the leader emphasizes being part of the group, whether the content is assertive, whether it contains orders, and whether it is more requesting a favor. We also code whether the leader blames others for a coordination failure, and has a laissez-faire style, in the sense that the message content is not explicitly concerned with leading the followers. In *Two-way*, we code whether the leader refers to follower messages. One category is coded especially for followers. Specifically, whether followers send orders and try to rule their leaders in a persuasive style.

Table A4 Categories of messages for content analysis

Category	**Description and examples**	**Frequency in leader messages**	**Frequency in follower messages**
		n = 198	n = 408
No message sent		0.131	0.331
Forty	Suggestion to choose 40	0.788	0.586
Ambiguous to interpret	E.g. “raise the number of hours”	0.015	0.022
Indecisive, questioning	E.g. “choose at least 20”, “maybe we can do 40”	0.025	0.015
Conditional strategy	Choose 40 in first period, and switch to 0 if group fails in first period.	0.056	0.034
Positive emotional content	Commendation, encouragement	0.465	0.279
Negative emotional content	Blaming, scolding others	0.040	0.047
Explanation	Explanation for suggested effort (e.g. refer to history, repeat rules of the game)	0.308	
Mutual benefit argument	Explanation refers to mutual benefit	0.182	
Being part of the group	Using the pronoun “we”, “us”	0.323	
Assertive style	Confident, not aggressive, not passive	0.687	
Order	Clear order, no reasoning, no explanations	0.429	
Asking	Asking the group desperately (e.g. “I beg you”, strong “please”)	0.030	
Excuse	Blaming others or history	0.035	
Laissez-faire style	E.g. “you know what to do”	0.066	
Referring to followers	Refer to what followers suggested	0.245	
Social banter	Unrelated to the game (e.g. jokes)	0.106	0.071
Follower order	Followers tell the leader what to do, as an order (e.g. “say”, “tell”, but not “let’s choose”)		0.152

Note. The variable “referring-to-followers” is only coded in *Two-way* and contains 102 observations.

Table A5 contains the correlation analysis of the often occurring message content categories in leader messages. We include the mutual benefit argument instead of the explanation content, as this category is supposed to be more relevant for leader effectiveness (Cooper 2007).

Table A5 Correlation of message content variables

n = 198	*Forty*	*Positive emotions*	*Mutual benefit*	*Being part*	*Assertive style*	*Order*
Forty	1.00					
Positive emotions	0.46	1.00				
Mutual benefit	0.18	-0.12	1.00			
Being part	0.20	0.16	0.49	1.00		
Assertive style	0.72	0.32	0.12	0.14	1.00	
Order	0.30	0.11	-0.38	-0.49	0.45	1.00

Table A6 presents estimates from random effects GLS regressions testing the effect of the communication treatment and the leader's gender on leader message content. In column (1), the dependent variable, in every message period, equals one if the leader requests forty and zero otherwise. In column (2), the dependent variable, in every message period, equals one if the leader uses an assertive style and zero otherwise. In column (3), the dependent variable, in every message period, equals one if the leader sends a clear order without any explanation and zero otherwise. In column (4), the dependent variable, in every message period, equals one if the leader expresses to be part of the group and zero otherwise. In column (5), the dependent variable, in every message period, equals one if the leader sends positive emotional content and zero otherwise. In column (6), the dependent variable, in every message period, equals one if the leader uses relevant content and zero otherwise. In column (7), the dependent variable, in every message period, equals one if the leader uses transactional content and zero otherwise. In all regressions, we use as independent variables the interaction of the communication treatment (*One-way* or *Two-way*) and the gender of the leader (male of female). In all regressions, we cluster standard errors on groups and include data only of leaders. Note that the omitted category is groups with a male leader in *One-way*.

Table A6 Treatment effects on leader messages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*Forty*	*Assertive style*	*Order*	*Be part of the group*	*Positive emotional content*	*Relevant content*	*Transactional content*
Two-way communication	0.19	0.19	0.15	0.04	0.23	0.04	0.23**
	(0.16)	(0.16)	(0.12)	(0.11)	(0.16)	(0.11)	(0.10)
Female leader	-0.04	-0.25	-0.08	-0.08	-0.13	-0.13	-0.15**
	(0.19)	(0.18)	(0.11)	(0.11)	(0.15)	(0.11)	(0.07)
Two-way, female leader	-0.06	0.04	-0.07	0.19	-0.02	0.07	-0.01
	(0.22)	(0.23)	(0.18)	(0.15)	(0.22)	(0.15)	(0.12)
Constant	0.73***	0.71***	0.42***	0.29***	0.42***	0.27***	0.19***
	(0.14)	(0.13)	(0.09)	(0.09)	(0.12)	(0.09)	(0.07)
Obs.	198	198	198	198	198	198	198
Clusters	33	33	33	33	33	33	33
Wald χ^2	3	7	4	5	7	4	33

Notes. Random effects GLS regressions. Standard errors (in parentheses) are corrected for clustering on groups. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

Figure A2 contains figures on the distribution of message values by treatment conditions over every message period.

Figure A2 Distribution of message values

One-way, leader messages *Two-way, leader messages* *Two-way, follower messages*

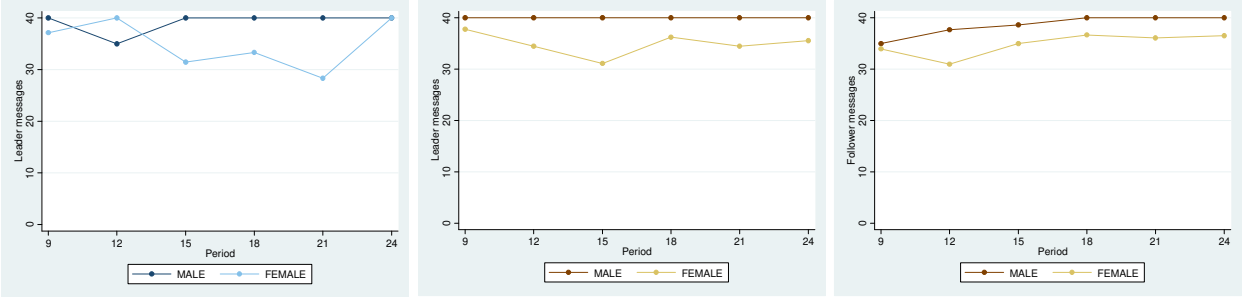


Table A7 presents estimates from random effects GLS regressions testing the effect of the communication treatment and the leader's gender on follower behavior. Follower behavior is measured by the followers' effort choices in each period, which is the dependent variable in both regressions. In both regressions, we use as independent variable the interaction of the communication treatment (*One-way* or *Two-way*) and the gender of the leader (male or female). In both regressions, we restrict the data to observations of followers in each period from 9 to 26 given that the leader requests the highest effort. In column (2), we restrict the data to observations of groups in periods 9 to 11 who did not coordinate on 40 in periods 5 to 8 (which was the case for 1 group), and to observations of groups in periods 12 to 26 if the group's average minimum effort in the previous message cycle was below 40, which means that leaders were previously not effective. We cluster standard errors on groups.

Table A7 Follower behavior

Dependent variable: minimum of followers' effort	*Any case*	*After failure cycle*
	(1)	(2)
Two-way communication	0.57	0.89
	(5.13)	(6.04)
Female leader	-0.08	0.27
	(5.41)	(6.37)
Two-way, female leader	-2.95	-3.94
	(7.43)	(8.47)
Constant	35.91***	35***
	(3.88)	(4.66)
Obs.	1872	468
Clusters	31	30
Wald χ^2	0	0

Notes. Random effects GLS regressions. Standard errors (in parentheses) are corrected for clustering on groups. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.