Purdue University Economics Department Working Paper No. **1330**

Gender, Beliefs, and Coordination with Externalities*

Timothy N. Cason[+], Purdue University

Lata Gangadharan[++], Monash University

Philip J. Grossman[#], Monash University; Rasmuson Chair, University of Alaska Anchorage

Abstract

Groups such as committees or boards make many important decisions within organizations. Many of these decisions affect external parties. This paper uses an experimental approach to study how the gender composition of three-person groups affects choices and beliefs in a Coordination game with selfish and prosocial equilibria. We find that the social preferences of group members are a key determinant of the group's coordination choice. Controlling for social preferences of the group, groups with more women are more likely to make choices that are kinder to external parties.

Both men and women believe that women will make kinder choices more frequently. Groups comprised of all men are expected to make 18 percentage points fewer kind choices than groups of all women. Men are also expected to be 9 percentage points less kind than women overall. These results have implications for public policies intended to increase gender diversity and women's representation on decision-making committees in the corporate sector, in politics, and in academia.

JEL Classification: C72, C92, D91

Keywords: Beliefs, Gender, Coordination, External Parties, Social Preferences, Laboratory Experiment

Funding provided by the Australian Research Council (DP130101695).

+ Department of Economics, Krannert School of Management, 100 S. Grant Street, West Lafayette, IN 47907-2076, U.S.A. Email: cason@purdue.edu

++ Department of Economics, Monash Business School, Wellington Road, Clayton, Victoria 3800, Australia. Email: lata.gangadharan@monash.edu

Department of Economics, Monash Business School, Wellington Road, Clayton, Victoria 3800, Australia. Email: philip.grossman@monash.edu

*The authors are grateful for helpful comments provided by James Bland, David Cooper, Martin Dufwenberg, Sarah Jacobson, Andreas Leibbrandt, Nikos Nikiforakis, Jean Paul Sobrino, seminar and conference audiences at Bath, European University Institute (Florence), Fordham, George Mason University, NYU Abu Dhabi, Penn, Shiv Nadar University, University of Alaska Anchorage, University of Western Australia, WZB (Berlin) and the Economic Science Association (Berlin, 2018), the editor, and an anonymous referee. For valuable research assistance we thank Stanton Hudja and Tennyson Yuan. Replication files are available on the Harvard Dataverse (https://doi.org/10.7910/DVN/YRUXVO).

1. Introduction

In recent years, there has been a move toward improving gender diversity in both the private and public sectors. In addition to providing opportunities to an underrepresented group, these measures could also potentially help improve decision making. Committees and groups make a large majority of important decisions in most organizations, so it is particularly important to understand how aspects of the group composition affect these decisions. The choices made by such committees are in part determined by the preferences and characteristics of their individual members, with gender an important and salient characteristic. This study therefore investigates how gender composition affects group decisions. Since group decisions often involve coordination, we explore this relationship using a Coordination game.

In many situations, the decisions made by committees and groups impose externalities on passive external parties, so our focus is on how the gender composition of a group influences choices that may reflect prosocial or selfish preferences. One prominent example is corporate board composition. Many decisions that are profitable for the board members or the corporations' shareholders may negatively affect others. For instance, corporate boards make decisions to invest in certain products or enter specific markets that may harm people in the community, and decisions by a board to close or relocate factories or merge with other companies can result in loss of jobs in a community and contribute to the slow decay and abandonment of small towns.[1]

The gender composition of groups such as boards is particularly relevant because it is an explicit policy choice. Norway in 2005 took the drastic approach of mandating publicly listed firms to have at least 40 percent women directors or be liquidated (Eckbo et al., 2014). Similar, but less strict policies have since been adopted in Belgium, France, Germany, Iceland, Italy, Malaysia, the

[1] Plant closings by General Motors in Flint, Michigan and by Hershey in Hershey, Pennsylvania are some illustrative cases of how board decisions can affect communities (see Armstrong, 2002).

Netherlands, and Spain, and most recently in California.[2] Such mandates are not limited to the corporate sector, however. In the political arena, eight European Union member states have legislated electoral gender quotas since 2000, requiring that a certain proportion of candidates be women. A further 14 have party quotas – voluntary commitments that a certain proportion of a party's candidates be women (Freidenvall and Dahlerup, 2013).[3] In Australia, both the major political parties have adopted proposals to increase the gender diversity in candidates for political office.[4] Many universities, of course, also require diverse perspectives on committees charged with making or implementing important policy decisions.[5]

In spite of this clear policy movement toward gender diversity on committees and other decision-making groups, little rigorous and causal evidence exists on the relationship between gender composition and group decisions (Azmat and Petrongolo, 2014). This paper investigates whether gender composition has an influence on decisions made in a group setting. Our research strategy relies on laboratory experiments that can help identify causal effects and the mechanisms underpinning this relationship. Other empirical approaches make it difficult to draw a clear link between gender composition and decision-making in groups. For example, data from board or committee meetings may not be publicly available and even if they are, key variables such as precise measures of meeting outcomes and beliefs about others' decisions are difficult to quantify.

---

[2] https://www.economist.com/blogs/economist-explains/2014/03/economist-explains-14 (accessed 22 March 2018); https://www.nytimes.com/2015/03/07/world/europe/german-law-requires-more-women-on-corporate-boards.html (accessed 6 August 2018); http://www.latimes.com/politics/la-pol-ca-governor-women-corporate-boards-20180930-story.html# (accessed 31 October 2018).

[3] Burkina Faso, Nepal, the Philippines, and Uganda have also included statutory candidate quotas in their constitutions, while Slovenia and Bosnia-Herzegovina have quotas written into their electoral laws (https://epthinktank.eu/2012/10/04/quotas-in-politics/ accessed 22 March 2018). India introduced affirmative action quotas for women in 1992, with one-third of all positions of the head of the village reserved for women.

[4] See https://probonoaustralia.com.au/news/2016/08/otoole-pushes-labors-gender-quota-line/ (accessed 23 March 2018) and http://www.abc.net.au/news/2016-09-06/liberal-partys-watershed-plan-to-attract-more-females/7819332 (accessed 22 March 2018).

[5] Woolston (2019) reports research that examines the hiring decisions of committees after a gender quota mandate in French universities (Deschamps 2018).

Moreover, gender composition is not randomly assigned in organizations, making it difficult to isolate and identify the impact of gender composition without confounding selection issues. Our experimental approach assigns participants randomly into groups to more clearly identify the underlying cause and effect of this relationship. Importantly, our groups are designed to have a majority and minority of men and women, which is often difficult to observe using observational data as fewer women are typically part of decision-making committees.[6]

The group decision we implement is a Coordination game. Individuals in many Coordination games have identical monetary payoffs over the set of possible outcomes, so their material interests are not in conflict and they are motivated solely to coordinate their strategies in order to obtain an outcome that is best for all of them. Many group decisions can be modeled as Coordination games and this is therefore a popular paradigm used in management and economics research (Cooper and Weber, 2020 and Devetag and Ortmann, 2007 survey the relevant experimental literature).[7] Managers seeking to coordinate actions of team members need to understand how the diversity of their groups' composition affects decisions. A Coordination game is therefore particularly useful for our investigation of gender composition, because choices in Coordination games depend partly on beliefs about the choices of others. If individuals believe that men and women make different choices, perhaps due to reliance on stereotypes (Bordalo et al., 2016), their best Coordination game choice may depend on their group's gender composition.

A novel feature of the Coordination game we study is that actions have payoff consequences for an agent who is external to the decision-making process. Our game therefore

[6] Bagues et al. (2017) examine the impact of greater representation of women in scientific committees. While composition was randomly selected, groups are not fully balanced. For instance, committees rarely have a female majority.

[7] For example, in many universities hiring and tenure decisions are based on consensus or the unanimity rule, which necessitates the need for coordination.

combines the incentive to coordinate with a tradeoff between the decision makers' own payoff and a desire to be prosocial. While significant evidence has accumulated that some individuals care about others' welfare in addition to their own material payoffs in many social dilemma and bargaining games, less evidence exists for Coordination games. Bland and Nikiforakis (2015) is one important exception. They examine if third-party externalities, positive and negative, can affect equilibrium selection in two-player Coordination games.

The presence of external parties who are affected by the group's decisions makes coordination more complex. In our study, three-member groups choose between two options, one of which provides them with a higher payoff but substantially reduces the payoff for the external party. In particular, if all three members choose the selfish option, they all receive a higher payoff but this hurts the external party. If instead they coordinate on a choice that gives them a modestly lower payoff, the external party's payoff increases by a substantial amount. If the choices of the three members do not match, then they and the external party all receive zero payoffs.[8]

Coordination failure is commonly observed and is one of the main reasons for the inefficient performance of groups (Brandts and Cooper 2006). To facilitate coordination, we allow group members to communicate with each other at the beginning of each round. Communication is anonymous, free-form, and nonbinding, akin to cheap talk. As demonstrated in the literature (e.g., Cooper et al., 1989, 1992; Blume and Ortmann, 2007; Cason et al., 2012), however, such communication can nonetheless assist groups in reaching an agreement about the option to choose.[9] An important advantage of communicating with others is that it can reduce the strategic

[8] For example, if a company is considering options to restructure, board members could choose an option such that they all receive a high payoff but this hurts the employees in the company as several lose their jobs. If instead they coordinate on retraining their employees, this may give them a modestly lower payoff as compared to the first option, but can increase the employees' payoffs by a substantial amount.
[9] Communication by a leader is another way that groups can coordinate on desired outcomes (Brandts et al., 2015).

uncertainty that individuals face about others' behavior. In our experiment, choices in the Coordination game affect others and trigger social preferences to different degrees across individuals, making the uncertainty more acute. Communication allows individuals to share their perspectives on the Coordination game, including norms about appropriate behavior, which could affect the relative amount of selfish and prosocial preferences expressed.

As we illustrate using a simple framework in Section 3, depending on the aggregate preferences of the individuals in the group, the group may coordinate on the prosocial or the selfish option. Further, even individuals who prefer the selfish option and do not care about the external party may decide to choose the prosocial option if they think that other members of their group will make the prosocial choice. Hence, in addition to their social preferences, individuals' beliefs about what others might choose (that are likely revealed during communication) help determine the final outcome in such Coordination games.

For example, if men and women have different social preferences, or if members hold the gender stereotype that women are more communal – more selfless and show concern for others (Eagly and Steffen, 1984), then groups with more women may have a higher likelihood of coordinating on the choice that increases the external party's payoff.[10] On the other hand, both men and women may also suppress their own social preferences and conform to the beliefs of the other gender's stereotype if they are keen to avoid coordination failure and a consequent payoff of zero. Bordalo et al. (2019) show that beliefs about gender differences in ability (on different types of knowledge questions) are biased by stereotypes and this harms group performance when submitting group answers. Our study also documents a stereotype bias in beliefs; but instead of

[10] Men are stereotyped as being more agentic, more self-assertive. According to social role theory (see Eagly and Wood, 1999 and Vogel et al., 2003), these stereotypes derive from the different roles women and men traditionally performed in their daily lives.

ability, we investigate the beliefs about the prosocial preferences across genders.

Our key objective is to examine how the gender composition of the group affects choices made by the group. We therefore exogenously vary the group in different rounds of the Coordination game so that participants are matched with different numbers of men and women. In some rounds, they are in mixed groups, with either a majority or minority of men, and in others they are in same-gender groups. This allows us to study if gender composition affects the group's choices over the selfish and prosocial options. Evidence suggests that women have different preferences as compared to men in the domain of risk and competition (Niederle and Vesterlund, 2007; Eckel and Grossman, 2008b; Croson and Gneezy, 2009). In terms of social preferences, Croson and Gneezy (2009) indicate that women exhibit more context specific prosociality and that their preferences are more malleable. Aguiar et al. (2009) and Brañas-Garza et al. (2018) show that women are expected to give more in a dictator game as compared to men; a recent meta-analysis finds that they give 13 percent more on average (Bilén et al., 2021). Building on this evidence, in most (but not all) rounds, we reveal the gender composition of the group along with information about another individual specific characteristic. In rounds in which this information is provided, as noted above, individuals' beliefs regarding the likelihood of their fellow group members making prosocial choices may depend on those members' gender. Our decision-making environment hence mirrors real-life situations in which opportunities exist for group members to communicate, consult, and advise each other about the decisions they take, while also observing some individual-specific characteristics of fellow group members. We also elicit individual social preferences using an Allocation task and beliefs about others' actions in the Coordination game.

We find that individual social preferences of members are a critical determinant of whether groups reach the prosocial equilibrium in the Coordination game. Increasing the number of

members who exhibit prosocial preferences in the individual Allocation task boosts the probability that the prosocial equilibrium is chosen by approximately 30 percentage points per person added. Controlling for the preferences of members, we find that increasing the number of women in the group augments the probability that the group chooses the prosocial equilibrium by 5-7 percentage points per woman member; consequently, groups with women in the majority choose the prosocial equilibrium by 11-14 percentage points more often than male majority groups. Uniform women groups choose the prosocial equilibrium up to 18 percentage points more often than uniform men groups. In the chat communications, women are more likely to agree to proposals of other group members, and they mention money less often than do men. This may be part of the "kernel of truth" (Bordalo et al., 2016) that is exaggerated in the beliefs, and we observe a stereotype bias in beliefs as women are expected to act prosocially 9 to 18 percentage points more frequently than men overall and in uniform gender groups.

Our findings have important implications. While gender composition of the group has an impact on prosocial choices by the group, the social preferences of individual members play a substantial and important role in shaping these choices. Prosocial preferences are however difficult to ascertain truthfully in the recruitment process, thus organizations interested in making decisions compatible with their social responsibilities may use gender as a predictor of prosociality. Consistent with this, the beliefs data clearly demonstrate that both men and women expect women to take more prosocial decisions in the presence of external parties. These perceptions can have an effect on the actual decisions taken in committees. Diversity in committees could therefore help lead to decisions that benefit others and perhaps even contribute toward reducing social and income inequality.[11]

[11] Greater diversity can also be perceived to lead to lower payoffs for the coordinating members of the group, consistent with the lower average performance observed in correlational studies of corporate board diversity (Adams

2. Related Literature

Our research contributes to two main strands of the literature. The first is the emerging literature in economics about the influence of gender composition of groups. The second is the well-established literature on Coordination games and communication.[12]

Gender composition of groups has been of increasing interest and has been examined using both experimental and observational data.[13] Apesteguia et al. (2012) find that, in a business game, women-only teams price less aggressively, invest less in research and development, and consequently earn lower profits than mixed or men-only teams. On the other hand, women-only teams invest more in social sustainability initiatives. Teams in their game however are not exogenously formed as the subjects register their own teams, thus they cannot control for the endogeneity of team formation. Hoogendoorn et al. (2013) instead randomly assign subjects to groups depending on their genders, and examine group performance (in terms of sales and profits) in a business venture. They find that teams with an equal gender mix perform better than men-dominated teams, although the comparison is less clear for all-women teams. Similarly, Dufwenberg and Muren (2006) find that in dictator games, all-women teams are more altruistic than all-men teams, but mixed-gender teams are the most altruistic. Baranski et al. (2021) find that fair bargaining agreements are highest in all-women groups and lowest in all-men groups.

---

and Ferreira, 2009). Board members who are critical of diversity policies could use this perception to justify the status quo. Post and Byron (2015), in their meta-analysis of 140 studies, however, conclude that women board representation is positively correlated with market performance in countries with greater gender parity, and that women representation is positively associated with financial monitoring. Credit Suisse Research Institute (CSRI, 2016) also indicates a positive correlation between diversity and business performance. The authors use data from 3,000 companies worldwide with a total of 27,000 senior managers. They find that in companies in which the majority in the top management are women, financial outcomes are superior (for example, they experience better sales growth, high cash flow returns on investments, and lower leverage).

[12] An extensive and insightful literature compares decisions made by individuals and by groups. For reviews of this literature see Charness and Sutter (2012) and Kugler et al. (2012).

[13] A large literature addresses behavioral differences by gender (i.e., risk attitudes, altruism, competitiveness, trust, bargaining). For a review, see Croson and Gneezy (2009).

Born et al. (2018) find that being in a male majority group exacerbates the tendency for women not to want to take on a leadership role. Ivanova-Stenzel and Kübler (2011) find that men perform better than women in an isolated memory task, but only in the presence of women. Grossman et al. (2015) find that woman leaders are more willing to take risks in a three-person investment game when playing in all-women groups. Keck and Tang (2017) show that confidence judgments by groups with at least one woman as member are significantly better calibrated than those by all-men groups and this is because groups with one or more women as members had a higher degree of opinion and information sharing. Hence these studies suggest that the success of gender diversity could depend on the outcomes that are being measured. But overall, in addition to bringing different preferences, knowledge base, and viewpoints to the table, the mere presence of women also alters the dynamics and the social sensitivity demonstrated by the group (Williams and Polman, 2015, Woolley et al., 2010).

In contrast to the existing experimental literature on gender composition of groups, our paper aims to examine decision-making in the context of Coordination games. This is of critical interest as many decisions are made by groups and groups are more effective if they can coordinate.

Two papers on the impact of board composition on firm performance and governance exploit data from the natural experiment in Norway, where publicly listed firms are mandated to have at least 40 percent woman directors. They provide mixed results. Ahern and Dittmar (2012) show that this affirmative action policy had a significantly negative impact on firm value and they attribute this to the newly added board members being younger and less experienced. Matsa and Miller (2013), by contrast, find that the policy did not affect corporate decisions in general, with the exception of employment policies; firms with more woman directors undertook fewer workforce reductions. Bagues et al. (2017) examine the role of evaluators' gender in scientific

committees using randomized natural experiments in Italy and Spain. Evaluators are randomly selected from a pool of eligible professors, thus enabling some (though not perfect) gender variation in group composition. Their main focus is on how this gender variation affects the evaluation of female versus male candidates. They find that having more women in the committee does not increase the quantity or quality of successful female candidates. Using data from the U.S, Kim and Starks (2016) show that women directors contribute additional expertise to corporate boards and this results in enhanced firm value by improving board advisory effectiveness. While their main contention is that women bring a diversity of skills, our mechanism importantly is based on the diversity of preferences and the beliefs group members have about these preferences, aspects that cannot be captured using observational data.

To our knowledge, only a few studies have examined the effect of gender on outcomes in Coordination games. Dufwenberg and Gneezy (2005) compare the performance of all-men and all-women 6-person groups in a 10-period, repeated play, minimum effort Coordination game. No preplay communication was allowed in their game and only the uniform gender composition of the groups was observable; subjects could see the other participants in the lab. Dufwenberg and Gneezy report no significant difference in chosen effort or group productivity. Di Girolamo and Drouvelis (2015) compare the performance of single-gender and mixed-gender 3- and 6-person groups in the same game as Dufwenberg and Gneezy. In the single-gender treatments, subjects are explicitly told the genders of their team members; in the mixed-gender treatment, subjects are unable to discern the gender mix of their team. While Di Girolamo and Drouvelis report no significant difference in chosen effort across the three treatments, it is worth noting that the all-women 3-person teams had the highest effort levels in every period except one. Holm (2000) shows that providing information about the opponent's gender in two-person Coordination games with

conflicting interests (such as battle of the sexes) makes subjects coordinate in ways that discriminate against women and decrease their earnings. In particular, both men and women subjects choose the outcome that benefits them more when they know that their opponent is a woman as compared to a man.

More recently, Babcock et al. (2017) examine gender differences in three-person groups, in which subjects coordinate on volunteering. While volunteering increases the payoffs of everyone in the group, it places the volunteer in a relative disadvantage because their payoffs are net of the costs of volunteering. They find that women volunteer more often than men in mixed gender groups, but when moving from a mixed-gender group to a uniform gender group, women see a decrease in their need to volunteer, while men see an increase, suggesting that beliefs about others' actions rather than individual preferences drive the gender differences in volunteering.

The Coordination game we study differs from these papers in several respects. The presence of a passive player and the externality resulting from coordination amongst members is a critical feature of our game. This introduces a tradeoff between maximizing own payoff and payoff to the passive player, creating a role for social preferences to shape decisions and leading to the emergence of selfish and prosocial equilibria. Further, in contrast to these studies, group members in our experiment have the opportunity to communicate with each other, facilitating the dissemination of preferences and beliefs. With respect to gender composition of the group, unlike Dufwenberg and Gneezy (2005) and Di Girolamo and Drouvelis (2015), our study varies gender systematically and we make it observable in both uniform and mixed gender groups, in addition to gender not being revealed in some rounds.

Finally, our paper also contributes to the literature studying the impact of communication in Coordination games. To our knowledge, this is the first attempt to document group members'

communication in a Coordination game with externalities. Thus, we extend the understanding of how groups can influence and resolve conflicting views. Communication could potentially lead to an increase in normative conflict with some encouraging others to take the selfish option and others publicly favoring the prosocial choice. When the gender composition of the group is known, gender specific preferences and beliefs could influence these communications and decisions.

3. Experiment Design and Procedures

3.1 Design

The experiment consists of four parts. In Part 1, players make decisions in a Coordination game. In Part 2, players participate in an Allocation task, while Part 3 assesses individual level risk preferences. Part 4 comprises demographic questions and also elicits beliefs about Part 1 decisions. Preferences and beliefs are separately elicited so as to help us understand behavior in the Coordination game.[14] Table 1 presents a timeline of the experiment. We discuss each of these parts in more detail below.

Each session employs 16 players. In the Part 1 Coordination game, the 16 players play in groups of four for 12 rounds. Three of the subjects in each group are randomly determined to be decision makers, occupying position C (referred to as type C in the instructions). The remaining subject in a group is in the Z position (referred to as type Z), and does not make a decision that affects payoffs. Subjects are rematched every round, but they retain their C or Z position for the entire experiment.

[14] We elicited beliefs in Part 4. Eliciting beliefs prior to Parts 1 and 2, or during Parts 1 and 2 would likely have influenced decisions made in those parts. As there was no feedback in Parts 1 and 2, those earlier decisions should have limited impact on beliefs.

Table 1: Timeline of Experiment

Part 0:	Initial questionnaire to collect gender and season of birth
Part 1:	12 rounds of the Coordination game • random rematching of groups • gender composition randomly varied across rounds • gender revealed for 9 of the 12 rounds (random order) • no feedback between rounds • all rounds paid
Part 2:	Individual Allocation decision • based on payoff used in the Coordination game • one randomly chosen group member's choice implemented for payment (no feedback)
Part 3:	Risk preference elicitation (no feedback)
Part 4:	Survey • (incentivized) belief elicitation concerning Coordination game choices: overall for each gender and by group gender composition • sociodemographic questions • payoffs for each stage revealed and paid

Note: Including preliminary data: 21 total sessions, 12 Coordination game rounds per session, communication for 60 seconds before each round, and gender revealed in nine of the 12 rounds. 336 subjects in total.

In each round, position C players individually choose between two actions, *M* and *J*. As shown in Table 2, a choice of *M* leads to a lower payoff and actually a negative payoff for Z, so for expositional reasons we refer hereafter to this choice as the *UNKIND-TO-Z* choice. Choice *J* is in contrast the *KIND-TO-Z* choice. Players, of course, only saw the neutral labels *M* and *J*. If all three position C players choose *UNKIND-TO-Z*, they each receive 7 experimental currency units (ECU). If they all instead opt for *KIND-TO-Z*, they each receive 5 ECUs. For position C players, *UNKIND-TO-Z* is the own-money maximizing choice. *KIND-TO-Z* is the prosocial choice; opting for *KIND-TO-Z* increases Z's payoff by 20 ECUs at the cost of lowering each position C player's payoff by 2 ECUs. If the choices of the three position C members do not match, then all four

players in the group receive 0 ECUs for that round.[15] Since the payoff to the player in position Z is affected by the choices of the position C players, the C players' decision generates a payoff externality on Z. The position Z player cannot influence this payoff.

Table 2: ECU Payoffs Earned for Coordination Game

All 3 position C choices are **M (UNKIND-TO-Z)**	All 3 position C choices are **J (KIND-TO-Z)**	All 3 position C choices **do not match**
position C all earn 7 each	position C all earn 5 each	position C all earn 0
position Z earns -16	position Z earns 4	position Z earns 0

Since our primary interest is to examine if the gender composition of the position C members of a group and information about the composition of the group affects decision-making, we vary the gender composition of the group across rounds, within sessions. The composition was varied across rounds randomly and differently in every session. In particular, in some rounds position C players are all men or all women and in others they are mixed: two men and one woman or vice versa.[16] Players are explicitly told that their decision screen may provide information describing the other two position C individuals in their group, i.e., players' gender and the seasons in which they were born while still maintaining anonymity. This information is obtained using a short survey at the beginning of the session. Players' characteristics are displayed using gender and season of birth icons, as illustrated in the experiment instructions in Appendix C. We include

[15] While the position C players are communicating and making their choices, the position Z player indicates what he thinks each of the C players will choose and why, and what he would himself have chosen if he had the role of a position C in the experiment. This helps ensure that player role and identity remain anonymous as all subjects are actively engaged in typing on their computers during this part of the experiment.

[16] We made the design decision to use three-person groups because we believed the decisions of men and women may differ depending on whether or not they are in the majority. For example, with a mixed-gender, three-person group, men, as the majority (minority), may be more (less) confident that the women will defer to the preference of men. Gender classification is based on self-identified gender and with this in mind, we refer to our subjects as men and women (gender as the social component), not male and female (which pertains to the biological component).

information on two player-specific characteristics instead of only focusing on gender to avoid priming gender as an artificially salient characteristic.[17] As we do not expect birth season to be correlated with subjects' choices, birth timing therefore allows for a placebo test to contrast with gender (see Section 4.2). To study the marginal effect of providing information about player characteristics we also include (in random order) some decision rounds in which this information is not revealed. Gender and birth season of fellow group members is revealed in nine out of twelve rounds and unrevealed for the remaining three rounds. The gender and birth season of the participant in position Z is never revealed to the group.

Evidence from previous experiments suggests that concerns toward inactive, external parties can be mixed, with some decision makers ignoring the presence of external parties and acting selfishly while others change their decisions to accommodate their concerns for the external parties (Charness and Jackson, 2009; Engel and Rockenbach, 2011; Humphrey and Renner, 2011; Andersson et al., 2014; Delaney and Jacobson, 2014; Blanco, et al., 2018). Bland and Nikiforakis (2015), for example, find that a large majority of subjects choose actions that increase their own payoffs even when external parties are substantially harmed. The mere presence of an external party can make the beliefs about the social preferences of other decision makers less confident (e.g., McDonald et al., 2013). To circumvent this and improve coordination, in all 12 rounds of the session, we allow the three position C players 60 seconds of anonymous, free-form, and non-binding preplay communication with each other prior to making their choices. Messages are only

[17] This was apparently successful, since mentions of gender in the chat communication (described later) were extremely rare, occurring in only 37 of the 8,455 total lines of chat. While the potential for experimenter demand effects always exists, we believe it does not affect our results because subjects received no cues about how gender (or birth season, for that matter) was related to objectives of the experimental investigation (Zizzo, 2010) until the survey at the end of the experimental session. Responses to the post-experiment survey indicate that 130 of the 336 subjects across all our datasets recalled the gender and birth season information when asked about it directly. A smaller number (80 of 336) indicated that gender affected their decisions and/or expectations about others' choices. But even if the display of personal characteristics primed some subjects to view gender as a salient aspect of the investigation, they could not determine what behavior was expected of them.

visible to the position C players. Although subjects remained anonymous throughout all sessions, gender and birth season labels were automatically shown on all chat statements in the rounds when these characteristics were revealed. Subjects followed some simple rules for this communication: to not identify themselves, be civil to each other, and avoid profanity. Apart from these restrictions, however, they could communicate about anything they wish.[18]

Even though theory suggests that nonbinding peer communication may not be effective as it is merely cheap talk and does not lead to credible ex-ante commitments (Farrell and Rabin, 1996), communication has been shown to significantly encourage coordination in many different situations.[19] The Coordination game in our setting has two clear Pareto-ranked Nash equilibria. *UNKIND-TO-Z* is payoff dominant for position Cs and if we assume that individuals only care about their own payoffs, we expect that position C decision makers will all choose *UNKIND-TO-Z*. However, if at least some proportion of individuals is motivated by equality, efficiency, and other prosocial concerns, this could lead position Cs toward choice *KIND-TO-Z*. This is particularly relevant because in our decision-making environment, position C players can communicate, thus improving their chances to coordinate on one of the two equilibria.

All players are paid for each of the 12 rounds in Part 1 but only receive feedback about their payoffs at the end of the session.[20] The experiment instructions (Appendix C) informed players that Part 1 consisted of 12 rounds, but players did not receive instructions to the later parts

[18] Mengel (2021) shows that this type of chat communication can increase gender bias when evaluating the performance of others. However, the effect of communication in a committee deliberation setting in which group coordination has payoff implications for external parties is as yet unexplored.

[19] For example, in the prisoners' dilemma: Loomis, 1959; Deutsch, 1960; Swensson, 1967; multiplayer prisoners' dilemma: Jerdee and Rosen, 1974; Dawes et al., 1977; public-good games using a voluntary-contribution mechanism: Isaac and Walker, 1988; Palfrey and Rosenthal, 1991; Pogrebna et al., 2011; Koukoumelis et al., 2012; Oprea et al., 2014; Jack and Recalde, 2015; trust games: Charness and Dufwenberg, 2006; Ben-Ner et al., 2011.

[20] We verified that this effectively limited spillover effects across rounds. The main regressions for the Coordination game choices shown in Table 3 indicate no significant time trends. We nevertheless control for potential cohort effects through session random effects.

until Part 1 was completed.

In Part 2, players participate in a simple Allocation task that measures their preferences over payoffs. The choice they face is the same as in Part 1, and again only position Cs make a decision. The key difference is that in Part 2, they do not need to consider the actions of other group members when making these decisions as they do not need to coordinate with others. For each four-person group, one individual Allocation decision, randomly selected, is implemented at the end of the experiment. This task is sometimes referred to as a random dictator game.

In Part 3, we elicit risk preferences using the Eckel and Grossman (2008a) risk task (see Appendix B). Players are asked to choose one out of five lotteries. Each lottery has two possible outcomes, both with an equal (50 percent) chance of occurring, that have increasing variance and expected value.

Part 4 is a questionnaire to obtain sociodemographic characteristics. The questionnaire also elicits incentivized beliefs about what percentage of men or women chose *UNKIND-TO-Z* or *KIND-TO-Z* for the Part 1 payoff case.[21]

3.2. Modeling framework

This subsection presents a simple modeling framework to illustrate how distributional social preferences may map into group decisions, as well as to motivate some specific assumptions required for a statistical power analysis. Consider a committee of N=3 individuals (position C

[21] Subjects are presented with three sets of questions. Subjects are told: "The percentages you are estimating are based on choices made previously by 96 Purdue students in this lab about 4 years ago, half men and half women;" and "One of the three sets of questions will be selected at random. For each answer to the randomly selected set of questions within 10% of the true percentage, you will receive 25 ECU$. For an answer between 10.01 and 20% of the true percentage, you will receive 10 ECU$. Any worse answer will receive 0 ECU$." Questions were of the type: "What percentage of Women do you think chose M?" "What percentage of Men do you think chose M in groups composed of all 3 Men?" and "What percentage of Women do you think chose M in groups composed of 2 Women and 1 Man?"

players) who have to choose between our two payoff cases: M = (7, 7, 7, -16) or J = (5, 5, 5, 4). Committee members can differ in their social preferences for payoff allocations M and J, which defines their type. The type indicates how much members prefer to be kind to the outside party, or instead look out for themselves.

The classic Fehr-Schmidt (1999) inequality aversion utility function is useful to model the preferences over payoff vectors $\mathbf{x}$ confronting the three group members:

$$V_i(\mathbf{x}) = x_i - \alpha_i \frac{1}{n-1}\sum_{j \neq i} \max[x_j - x_i, 0] - \beta_i \frac{1}{n-1}\sum_{j \neq i} \max[x_i - x_j, 0]$$

Individuals care about their own payoffs (x_i) and may also care about how these payoffs compare to those of others. For the two payoff vectors, the α_i term is not relevant because the decision-makers only face advantageous inequality. There is no inequality with respect to other (position C) committee members, so it is trivial to see that the utility for these two payoff allocations is:

$$V_i(M) = 7 - \frac{\beta_i}{3}23 \text{ and } V_i(J) = 5 - \frac{\beta_i}{3}.$$

Individuals prefer the Unkind allocation M over the Kind allocation J whenever the inequality aversion term is not too high; specifically, whenever $\beta_i < 3/11$. Also, when facing a choice between the Unkind allocation M and miscoordination (which gives a payoff of 0 to everyone), only an individual with a strong dislike for advantageous inequality would prefer miscoordination (specifically $\beta_i > 21/23$).

Fehr and Schmidt (1999) proposed the following three-value distribution for β_i, based on a calibration across a range of ultimatum game results (see their Table III): β_i = (0, 0.25, 0.60) in proportions (0.3, 0.3, 0.4). This implies that about 30% of the sample would have a $\beta_i = 0$; 30% would have a $\beta_i = 0.25$ and the remaining 40% a $\beta_i = 0.60$.

For our application, the self-interested type with $\beta_i = 0$ would prefer allocation M, and the

most inequality averse type with $\beta_i = 0.6$ would prefer allocation *J*. The intermediate type $\beta_i = 0.25$ is close to the critical value of $\beta_i = 3/11$ noted above. As this calibration is clearly a rough approximation, it seems sensible to consider this intermediate type as potentially supporting either allocation depending on the circumstances, others in their group, and/or their exact preferences.

When groups are formed in each round, they communicate about which payoff allocation to implement, thus revealing their preference to others in their group to coordinate to a common choice. For instance, if a majority (2 or 3 group members) have $\beta_i = 0$ they may choose *M*. If a majority (2 or 3 group members) have $\beta_i = 0.6$ they may be more likely to choose *J*. If exactly 2 group members are the intermediate type (i.e., in the neighborhood of $\beta_i = 0.25$) they may often choose the allocation preferred by the third member. In contrast, the choice is likely to be somewhat random if the group consists of one member of each type or if all 3 group members are the intermediate type $\beta_i = 0.25$. Hence depending on the preferences of the individuals in the group, the group may coordinate on the *KIND-TO-Z* or the *UNKIND-TO-Z* outcome.

Another potential channel through which group composition could affect group choices is through beliefs about others' actions. Even players who would prefer to choose the *UNKIND-TO-Z* option may select the *KIND-TO-Z* option if they think that other members of their group will make the prosocial choice, because failing to do so may lead to miscoordination and zero payoffs. While they do not observe the preference type of other members, the subject characteristics displayed may lead them to infer their type. This is how gender stereotypes and beliefs can affect choices in Coordination games as the group gender composition changes. In our decision-making environment, members can communicate with each other in every round, hence potentially revealing preferences and the choice they would like to make. This reduces strategic uncertainty and the need to fully rely on ex ante beliefs.

Which gender dominates the group decision is ambiguous. While the literature suggests that in some contexts women are more prosocial, it is unclear whether this will dominate in a group decision-making game. Therefore, we refrain from making a specific directional hypothesis about the impact of group gender composition. Instead we aim to explore behavioral regularities, which is also a strength of the experimental approach we employ. Communication allows for the possibility that group members share their views and update their beliefs about others' choices, thereby better coordinating on selfish and prosocial choices.

To summarize, we designed our experiment with the following research questions in mind:

1. Does the group gender composition influence the group's choices over selfish and prosocial options (towards the external party)?
2. How do individual preferences of group members influence this group choice?
3. Do beliefs about the proportion choosing the prosocial option vary according to individuals' gender and the gender composition of the group?

3.3 Parameter choice and power analysis

In addition to the payoff case (7, 7, 7, -16) vs. (5, 5, 5, 4) discussed in sections 3.1 and 3.2, in eight preliminary sessions (128 total subjects) we included two other payoff cases: (7, 7, 7, 4) vs. (5, 5, 5, 24) and (7, 7, 7, -4) vs. (5, 5, 5, 16). In both of these payoff cases, position C players face a choice between advantageous and disadvantageous inequality with respect to the position Z player. The disadvantageous inequality for the equilibrium with position C players earning 5 is sufficiently high that neither of the two *KIND-TO-Z* choices would be selected by anyone who has $\alpha_i > \beta_i$, which is a key assumption that Fehr and Schmidt make and empirically justify. This is also what we found in the preliminary sessions, as only 16% of position C players overall chose the *KIND-TO-Z* option for these two cases. (Appendix A provides more details.) We therefore focus only on

the payoff case: (7, 7, 7, -16) vs. (5, 5, 5, 4) for the main experiment.

The preliminary data also suggests that men and women have different distributions of β_i. Based on the allocation task choices in the preliminary dataset (in which a greater proportion of women chose the *KIND-TO-Z* option than men) and the proposed distribution by Fehr and Schmidt (1999), we calibrated a β_i distribution for men and women as follows:

β_i = (0, 0.25, 0.60) in proportions (0.40, 0.30, 0.30) for men.

β_i = (0, 0.25, 0.60) in proportions (0.30, 0.15, 0.55) for women.

These differing type distributions form the basis for the power analysis. We calculated the sample size needed based on the formal empirical analysis that we planned to conduct (e.g., choices aggregated at the group level, and based on regressions that account for within-session correlation). Our calibrated β_i distribution led to first and second moments of the choice distribution that approximately match those observed in the preliminary data. For a modified design with all 12 rounds focused on the single payoff case, and with random regrouping of individuals across rounds in 16-subject sessions, a power analysis based on 1,000 simulated samples indicates that data from 208 subjects in 13 sessions provide statistical power consistent with prevailing standards (80% power at 5% significance level). This is described in more detail in Appendix A.

3.4. Procedures

All sessions were conducted at the Vernon Smith Experimental Economics Laboratory at Purdue University, using z-Tree (Fischbacher, 2007). Subjects were undergraduate students, recruited across different disciplines at the university by email using ORSEE (Greiner, 2015). Equal numbers of men and women were recruited to each session, using the gender they indicated on their ORSEE registration. Subjects were only invited if they had not previously participated in any

similar coordination or allocation experiment. No subject participated in more than one session.

At the beginning of each experimental session, an experimenter read the instructions aloud while subjects followed along on their own copies. At the end of the instructions, subjects took a computerized quiz to test and reinforce their understanding of the instructions. Earnings in the experiment are denominated in ECUs, and these are converted to U.S. dollars at a pre-announced 10-to-1 conversion rate. Subjects' total earnings averaged US$21.67 each, with an interquartile range of $17.50 to $28.50. Sessions usually lasted less than one hour, including the time taken for instructions and payment distribution.

4. Results

We present results in four subsections. We begin first with the results for the nonstrategic Allocation task in section 4.1. We then discuss the Coordination game choices with the main outcome variable being the group choice in section 4.2 followed by Beliefs in section 4.3. Section 4.4 reports an analysis of the chat content during the Coordination game. Recall, a "*KIND-TO-Z*" choice refers to one in which the position C players earn 5 and the position Z player earns 4; and an "*UNKIND-TO-Z*" choice leads position C players to earn 7 and the position Z player to earn -16. In the discussion below, we often refer to the subjects who made *KIND-TO-Z* Allocation choices as "Kind types" and other as "Unkind types."

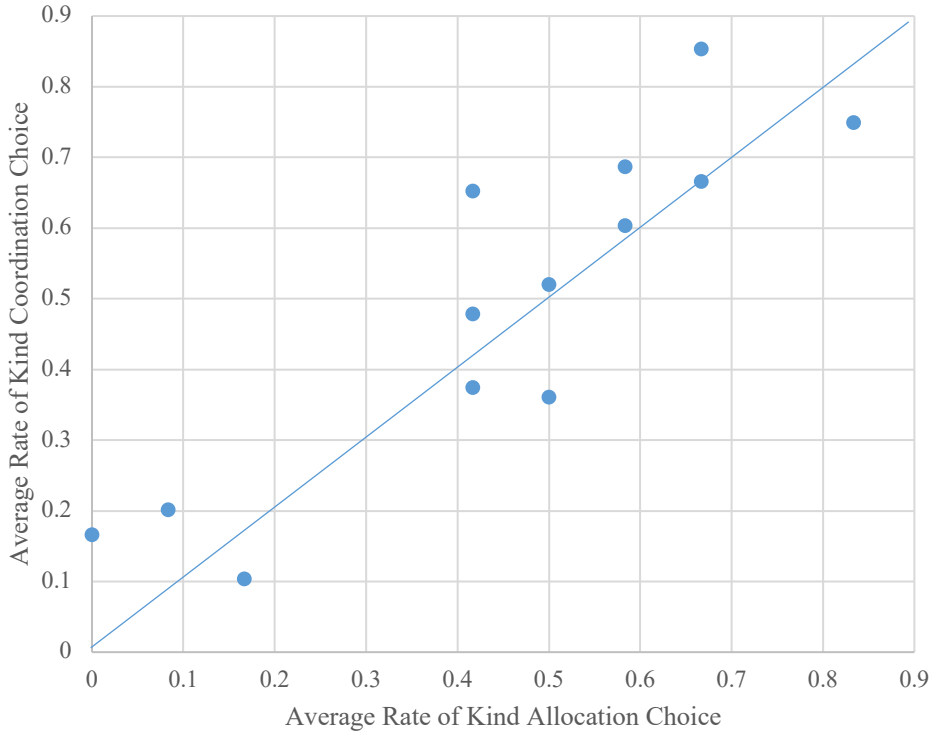
4.1 Allocation choices

Overall, 45% of the 156 Allocation choices (made by position C players) were *KIND-TO-Z*. Men were slightly more likely to make the *KIND-TO-Z* allocation (49%) than were women (41%), but

the difference is not statistically significant (p-value = 0.295).[22]

Substantial variation in the average rates that subjects made the *KIND-TO-Z* allocation choice exists across the 13 sessions. Figure 1 displays thirteen dots, one for each session, indicating a session's average Kindness rates for the Allocation and Coordination Game choices. Not surprisingly, this rate is highly correlated (correlation coefficient = 0.89). A Wilcoxon test (n=13 sessions) indicates no significant difference (p-value = 0.210) in Kindness rates across the Allocation and Coordination Game choices within sessions. This is our first indication that overall *KIND-TO-Z* choice rates in the Coordination Game are strongly influenced by the distribution of pro or antisocial preferences for the subjects randomly assigned to each session.

Figure 1: Correlation of Coordination Game and Allocation Game Kindness across Sessions



Note: Each dot indicates the average Kindness rates for the choices for one of the 13 sessions.

[22] Statistical test is based on a linear probability model with session random effects. The calculated p-value is identical for a random effects logit model.

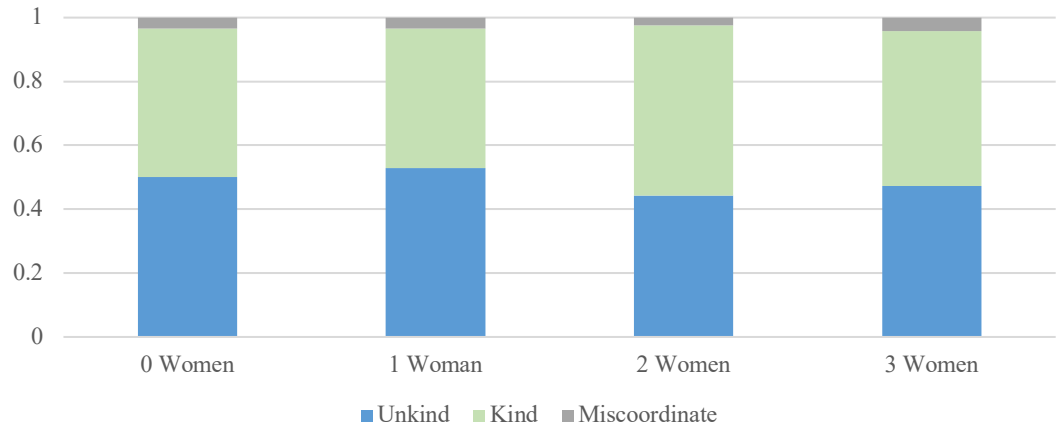
4.2 Coordination Game

Each group plays the Coordination game twelve times: nine times with gender and birth season revealed to other group members, and three times with this information withheld. Recall that any failure to coordinate by the three position C players results in a payoff of 0 for all four individuals in the group.

In all thirteen sessions, the position C players could first communicate in a computer-mediated chat room. This communication, not surprisingly, leads to a high rate of coordination; players coordinate on a positive payoff in all but 20 of the 624 group decisions (97 percent).[23] Roughly half (300) coordinate on the *KIND-TO-Z* choice.

Figure 2 reports the distribution of the group choices (*KIND-TO-Z*, *UNKIND-TO-Z* and miscoordination) by the number of women in the group. While a slightly higher frequency of *KIND-TO-Z* choices are made by groups in which women are in the majority, the difference is not stark.

Figure 2: Group Coordination Choices by Number of Women in Group



[23] There were no systematic differences in the gender composition of the 20 miscoordinating groups, as 3 groups were comprised of 3 men, 4 groups were all-women, 5 groups had 2 women, and 8 groups had one woman.

In contrast, what is strongly correlated with the group's *KIND-TO-Z* Coordination game choice is how many of the group members made *KIND-TO-Z* Allocation decisions. Figure 3 shows the clear monotonic relationship between the number of Kind types in the group and the group choice being Kind. In particular, if the group is comprised of members who all made *KIND-TO-Z* Allocation choices, in the Coordination game, they coordinate on the *KIND-TO-Z* choice 93% of the time. But if the group had no members who made *KIND-TO-Z* allocation choices, they coordinate on the *KIND-TO-Z* choice only 10% of the time.

Figure 3: Group Coordination Choices by Number of Kind Allocation Choice Types in Group

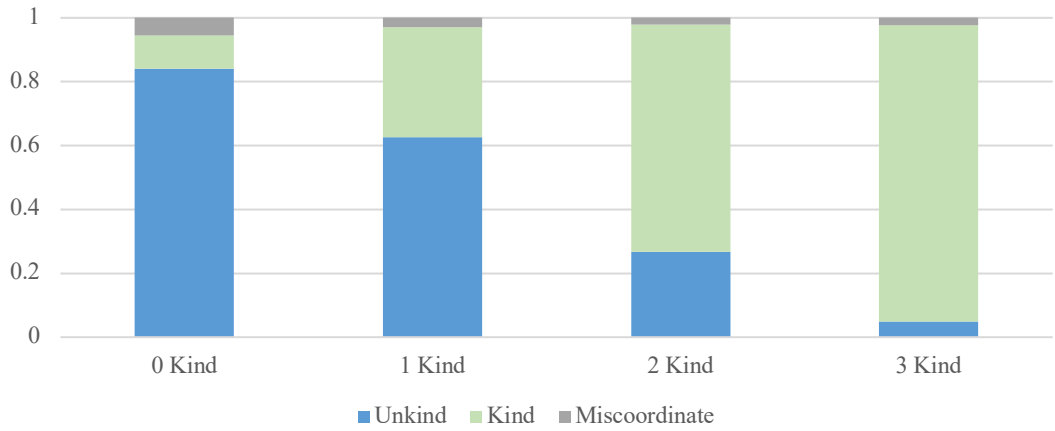
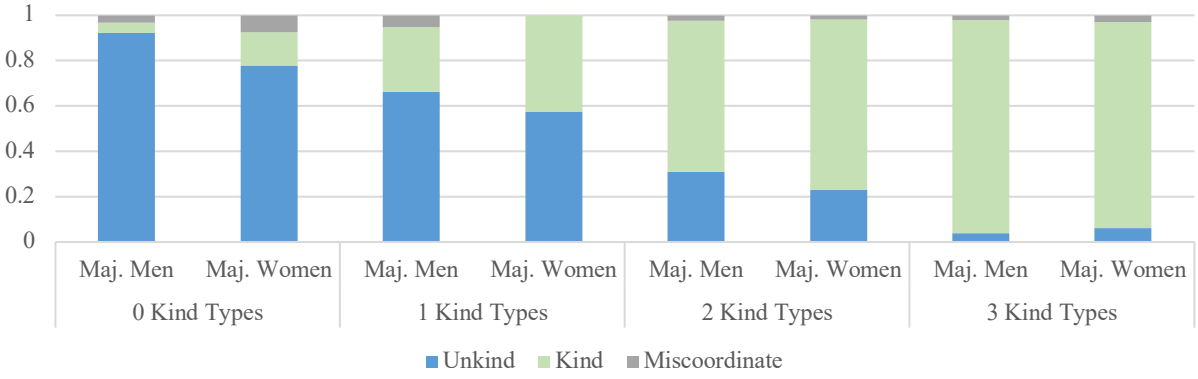


Figure 4 illustrates group choices controlling for the number of Kind types for majority men or women groups. When controlling for the Kind type distribution, the majority women groups select the Kind equilibrium more frequently than the majority men groups, except for the case of 3 Kind Types.

Table 3 reports Linear Probability Model regression results which bring these results together. The models are estimated with session random effects to account for within session correlation (Fréchette, 2012). The dependent variable is an indicator variable that denotes whether

Figure 4: Group Coordination Choices by Majority Men and Women Groups, Controlling for Number of Kind Allocation Choice Types in Group



the group coordinated on the *KIND-TO-Z* choice in the Coordination game.[24] Informed by the power analysis presented in Appendix A, we focus on three ways of capturing the impact of the gender composition of groups. The first explores group choice as a function of the number of women in the group (models 1-3) and the second as a function of whether women represent a majority in the group or a minority (a dummy variable, models 4-6). The third examines group choice as a function of uniform gender groups (a dummy variable indicating whether it is an all-women group, models 7-9). By definition this is a smaller sample because it excludes the rounds in which the gender composition is mixed.

In addition to the gender variables, in different specifications, we include the number of "Kind types" in the group (as determined by members' Allocation task choices), a dummy variable on whether information about member characteristics was provided in a round and a variable to capture any time trend (Round). To control for heterogeneity across groups, we also include risk preference and demographic variables as indicated in the table notes.

[24]In cases in which there is miscoordination (3% in our data), we use the majority choices to create the indicator variable (the miscoordinating groups are classified as KIND if 2 out of 3 members chose the *KIND-TO-Z* option, and as UNKIND if only 1 out of 3 chose the *KIND-TO-Z* option). We also examine robustness of results by excluding the group choices that did not lead to coordination and find that results remain consistent.

Table 3: Linear Probability Model Regression Results for Coordination Game Group Choices

	Number of women			Majority women			Uniform Gender		
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Number of women in group	0.033+ (0.020)	0.052** (0.018)	0.071** (0.019)						
Dummy=1 if majority women				0.075* (0.036)	0.106** (0.032)	0.139** (0.035)			
Dummy=1 if an all-women group							0.059 (0.067)	0.112+ (0.061)	0.184* (0.078)
Number of Kind allocation types		0.291** (0.021)	0.274** (0.019)		0.291** (0.021)	0.273** (0.018)		0.292** (0.035)	0.285** (0.035)
Dummy=1 if gender info provided			-0.010 (0.037)			-0.015 (0.037)			-0.021 (0.086)
Round			0.006 (0.005)			0.007 (0.005)			0.007 (0.009)
Constant	0.439** (0.072)	0.020 (0.051)	0.043 (0.114)	0.452** (0.058)	0.045 (0.045)	0.087 (0.107)	0.468** (0.073)	0.051 (0.068)	0.070 (0.215)
Demographic & Risk Preference Controls #	No	No	Yes	No	No	Yes	No	No	Yes
Observations	624						185		
Sessions	13								

Notes: Dependent variable: The group coordinated on the *KIND-TO-Z* choice in the Coordination game (1=Yes). Standard errors in parentheses (models estimated with session random effects); ** p<0.01, * p<0.05, + p<0.1 (all two-tailed tests); # included control variables: fraction of group who were the least risk averse (i.e., risked the most money in the risk preference task), fraction of group who were Business/Econ majors, fraction of group who were Engineering/Science majors, fraction of group who were not born in the USA, fraction of group who were nonwhite/Hispanic, fraction of group who were religious, fraction of group who self-reported a high GPA, and fraction of group who were freshman/sophomores.

Columns (1), (4), and (7) report the simple regression specifications for group outcomes that we employed in the power analysis, with only the group gender composition as an explanatory variable. They indicate a marginally significant impact of the number of women in the group, a significant impact if the group is majority women, but no significant impact for uniform gender groups (column 7). Controlling for the number of "Kind types" (as determined by group members'

Allocation choices), in columns (2), (5), and (8), the number of women in the group has a larger and more statistically significant impact on the group's *KIND-TO-Z* choice. Groups coordinate on the *KIND-TO-Z* equilibrium more often when the group has more women members. After adding controls for other group level variables (columns 3, 6, and 9), the group gender composition coefficients are 2-3 times larger as compared to the parsimonious specification with gender composition of the group as the only explanatory variable. Across the specifications reported in Table 3, the first two group gender composition variables led to a 5-14 percentage points increase in the probability that the group will make a Kind choice. The third specification, for uniform gender groups, has the advantage that gender specific beliefs about others' actions are homogeneous. Controlling for social preferences and other variables, the all-woman groups are 18 percentage points more likely than all-men groups to coordinate on the Kind choice.

While this is a substantial effect, it is notable that the impact of the Kind types is considerably larger, with an effect range of 27-29 percentage points per Kind group member. In Appendix B, we provide additional results relating to Table 3. Table B-1 displays results for each of the included demographic controls.[25] Table B-2 reports Logit models for all the specifications, and documents that results are qualitatively unchanged. Table B-3 reports results including data only from periods with gender information revealed, with findings qualitatively similar to those reported above. As we do not find gender differences in the individual preferences as measured by the Allocation task, the gender difference in group kindness observed in all specifications is likely due to gender differences in beliefs about others' actions. Overall social preferences of members have a stronger effect in our data than the effect of gender *per se*.

[25]Holding other variables constant, groups with a larger fraction of least risk averse members; a larger fraction born outside the US; and a larger fraction who are younger and inexperienced are more likely to choose the Kind option. Groups with a larger fraction who are business and economics majors and a larger fraction who are non-white, non-Hispanic are less likely to choose the Kind option.

Recall that we collected and displayed information about subjects' birth season as well as their gender, mainly to avoid making gender too salient as the only displayed characteristic. We did not expect birth timing to be correlated with group choices. Birth timing therefore provides a convenient placebo test to contrast with the significant gender difference. Table B-4 in the Appendix presents regressions analogous to Table 3, but with three variables relating to birth timing replacing the group gender composition variables. The three birth timing variables are: the number of group members who were born in the first half of the year; a dummy variable indicating whether a majority were born in the first half of the year; and a dummy if everyone in the group was born in the first half of the year. We find no birth timing impact in any of the specifications. Similarly, birth timing does not correlate with the individual Allocation choices (p-value=0.184).

An advantage of our experimental design is that we can examine the degree of internal consistency displayed by the position C players across the 12 rounds. As noted before, participants faced decisions with varying group compositions in a random order. We calculate the "average" frequency of the Kind decision (within subjects) across the 12 Coordination game rounds. If this average is 0 or 1 the subject is perfectly consistent; if it is 0.5, they are maximally inconsistent. We transform this into a consistency score, defined as the absolute deviation in the average choice frequency from 0.5. The average consistency score in our sample is 0.262 (with a standard deviation of 0.151). A session random effects regression, in which the dependent variable is the consistency score and the independent variables are a gender dummy and a dummy for the Kind type (measured by the subject's Allocation choice), shows that women are marginally less consistent than men (p-value =0.062). Kind or Unkind types are not more or less consistent. Thus, women appear to adjust their choices marginally more often than men, perhaps based on the gender composition of the group, consistent with Croson and Gneezy's (2009) observation that "gender

differences in other-regarding preferences … [arise because] women are more sensitive to cues in the experimental context than are men" (p. 463).

4.3 Beliefs about *KIND-TO-Z* choices by gender

Prior to revealing payoffs and any decisions of other subjects, we ask all subjects (position C as well as position Z) for their beliefs about what fraction of previous subjects they thought made *KIND-TO-Z* choices in the Coordination game (see footnote 21). We incentivized them to report accurately, and we asked about men and women separately depending on the group composition and overall.

Subjects overwhelmingly believe that women made the *KIND-TO-Z* choice more frequently overall, and in uniform gender groups. Overall, 134 of the 208 subjects (64 percent) believe that women were more often Kind than men, while only 27 subjects (13 percent) believe that men were Kind more often. (The remaining 23 percent believe that men and women were Kind with equal frequency.) Random effects regression models (not reported) strongly reject the hypothesis that men and women are believed to be equally likely to make the *KIND-TO-Z* choice (p-value < 0.01). Additional control variables indicate that the gender of the person reporting beliefs or the person's position (Z or C) do not influence these (incorrect) beliefs.

Figure 5 indicates how subjects' beliefs in the fraction of men and women choosing the *KIND-TO-Z* choice in the Coordination game depend on the group composition. On average subjects believe that about 30 percent of men will make the *KIND-TO-Z* choice in groups of 3 men, and they expect that this rate increases by about 10 percentage points for every additional woman who replaces a man in the group. They believe that about 48 percent of women will make the Kind choice in groups of 3 women, and that this rate decreases by 3 or 4 percentage points for

every man who replaces a woman in the group. The Linear Probability model regressions reported in Table 4 indicate that these shifts in beliefs due to the gender composition of the group are statistically significant.

Figure 5: Mean beliefs in % terms about coordination game choices for men and women, overall and by group gender composition

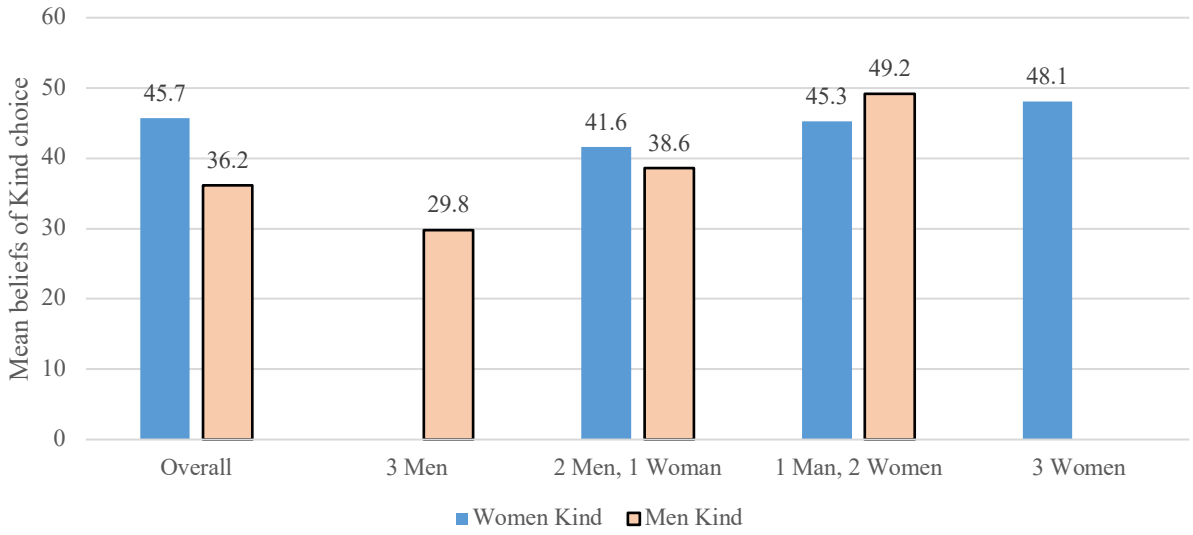


Table 4: Beliefs in Coordination Game Kind Choice by Gender, Depending on Group Gender Composition

Variables	(1) Belief Men Kind	(2) Belief Women Kind
Number of Women	9.69**	3.26**
In Group	(0.71)	(0.64)
Dummy = 1 if belief	0.71	2.29
is stated by a woman	(3.68)	(3.84)
Dummy = 1 if belief	-2.87	-7.56+
is stated by position Z	(4.24)	(4.43)
Constant	29.90**	39.18**
	(2.90)	(3.23)
Observations	624	
Number of subjects	208	

Notes: Dependent variable: The percentage of men or women supporting the *KIND-TO-Z* choice in the Coordination game. Standard errors in parentheses (models estimated with random effects on subjects); ** p<0.01, * p<0.05, + p<0.1 (all two-tailed tests).

Finally, we report in Table 5 Linear Probability model regressions that use the percentage difference in the reported *KIND-TO-Z* rate for men minus women. This measure indicates the difference in the expected kindness rate for the two genders. Since the dependent variable is the difference in kindness beliefs (men – women), the constant in this regression is the important variable to focus on. Results indicate that men are expected to be 8.7 (18.3) percentage points less kind than women overall (in uniform gender groups).

Table 5: Difference in (Men – Women) Beliefs in Kind Choice, by Group Gender Composition

Variables	Overall	Uniform Gender	2 Men, 1 Woman	1 Man, 2 Women
Dummy = 1 if belief is stated by a woman	-3.18 (2.46)	-3.63 (3.97)	-2.55 (2.28)	0.99 (2.65)
Dummy = 1 if belief is stated by position Z	3.07 (2.83)	7.52+ (4.57)	2.66 (2.63)	3.898 (3.06)
Constant#	-8.69** (2.46)	-18.29** (4.65)	-2.33 (1.74)	2.47 (2.02)
Observations	208			
Sessions	13			

Notes: Dependent variable: percentage difference in the reported *KIND-TO-Z* rate for men minus women. Standard errors in parentheses. Models estimated with random effects on sessions. ** p<0.01, * p<0.05, + p<0.1.

As a robustness test, we examine the data from the 52 position Z subjects separately. The beliefs for each of these subjects are statistically independent since position Z subjects neither make coordination or allocation decisions nor do they communicate with other subjects. They thus receive absolutely no feedback in the session. Similar to the results for the full sample, position Z subjects expect men to be significantly less kind Overall and for Uniform gender groups (Wilcoxon *p*-values<0.01).

4.4 Content of Communication

The position C players could exchange written chat messages for 60 seconds each round before submitting their Coordination game choices. In order to quantify their statements, we employed three Purdue student coders, who were unaware of the research questions addressed in this study, to independently read and classify all 5,877 lines of chat in the 624 chat rooms. During their training, they read the experiment instructions in order to understand the implications of the *UNKIND-TO-Z* and *KIND-TO-Z* choices the subjects made following the chats, but were unaware of the subjects' actual decisions. The coders judged whether each individual chat line fit into 15 different specific meaning categories and subcategories that were defined by the authors. Individual chat lines could be assigned to multiple categories. We used Cohen's Kappa (Krippendorff, 2004; Cohen, 1960) to assess category classification reliability, which nets out the level of coder agreement that can occur simply by chance.

Table 6 summarizes the mean frequency that coders identified different content categories across all chat statements. All categories meet at least the "moderate" agreement threshold (Kappa>0.4), and many are higher. Subjects' chats tend to focus on the choice between the *UNKIND-TO-Z* and *KIND-TO-Z* action, which is not surprising given the zero payoffs from miscoordination in the game. The most common statements mentioned M [*UNKIND-TO-Z*] or J [*KIND-TO-Z*], or agreed with previous proposals made to choose a particular action.

Notably, expressions of concern for the external party (Category 1) or statements about being nice or altruistic (Category 2) are relatively infrequent, as are statements mentioning money (3). Some (verbatim) examples of such statements are the following:

<u>Category 1 -- Concerns expressed for player Z's earnings/welfare/well-being:</u>

> we can't screw Z, they'll appreciate us not doing it
> kinda feel bad for z's tbh

Table 6: Average Frequency of Chat Statement Classifications

Category	Description	Mean Frequency	Cohen's Kappa
1	Concerns expressed for player Z's earnings/welfare/well-being	0.040	0.713
2	Mentions of being nice/altruistic	0.008	0.515
3	Mentions of money (generally and with specific goals)	0.055	0.769
3A	Expresses goal of maximizing own (position C) money earnings	0.014	0.436
3B	Indicates goal of making everyone better off / more money	0.013	0.768
4	Asks for proposal/advice	0.022	0.626
5M	Mentions choice M (agreements with Mare 9M below)	0.123	0.576
5J	Mentions choice J (agreements with J are 9J below)	0.109	0.583
6	Mentions gender	0.005	0.849
7	Subject mentions (birth) season	0.010	0.843
8	Mentions connection between gender and prosociality/ selfishness	0.001	0.416
9	Agrees to previous message in the group	0.382	0.574
9M	Agrees with M proposal	0.162	0.495
9J	Agrees with J proposal	0.155	0.537
11M	M either mentioned or agreed to (constructed from 5M and 9M)	0.285	0.827
11J	J either mentioned or agreed to (constructed from 5J and 9J)	0.264	0.841
10	Other (any statement not fitting into above categories)	0.285	0.804

Note: Kappa values between 0.40 and 0.60 are considered "moderate" agreement; values between 0.60 and 0.80 are considered "substantial" agreement; and values above 0.80 are considered "almost perfect" agreement (Landis and Koch, 1977).

at what point do we feel bad enough for the z's that we choose j [*KIND-TO-Z*]?

whoa we're feeling bad for Z here ?

Category 2 -- Mentions of being nice/altruistic:

you guys wanne be nice to Z this time?

we're nice folks around here

we going to be nice to Z lol?

Category 3 -- Mentions of money (generally and with specific goals):

Tell em to bring me my money

maximise our profit

Same but i need the $$

Keep choosing the option where everyone gets money

Explicit mentions of gender (Category 6) and the potential connection between gender and social preferences (Category 8) are rare, indicating that our manipulation of gender information did not substantially prime subjects to focus on the gender of their fellow group members. As indicated in Table 6 (Categories 6 and 7), mentions of gender are actually about half as frequent as mentions of birth season, although both are uncommon.

Some of the content communicated in the chat rooms differs significantly by gender. Men communicated more individual lines in each chat room (3.22 lines on average) compared to women (3.08 lines), but this difference is not statistically significant.[26] Table 7 reports the average frequency that chat statements are classified for the five categories in which women and men communicated significantly differently. Women mentioned gender less frequently, and agreed to previous chat messages more frequently than men.[27] Men are marginally significantly more likely to mention money, compared to women. The average number of statements about money increases monotonically from 0.41 for 0 men in the group, to 0.46 for 1 man, 0.58 for 2 men, and 0.62 for 3 men in the group (p-value=0.025 for a poisson regression with random session effects).

[26] Born et al. (2018) found that men speak more than women in their mixed gender groups, communicating face-to-face. Men submitted a similar number of chat lines in our experiment regardless of whether they were in mixed or uniform gender groups, but women tended to submit fewer lines (2.85 on average) in uniform gender groups than in mixed gender groups (3.18 on average). This difference is statistically significant based on a Poisson count regression (p-value=0.012). Men in our experiment submitted the first line of communication in the chat room at almost exactly the same rate as women.

[27] Women also agreed to previous statements at significantly higher rates than men in uniform as well as all kinds of mixed gender groups separately (p-value<0.021 for all cases except for groups with two women, in which p-value=0.063).

Table 7: Differences in Chat Statements by Gender

Category	Description	Men Freq.	Women Freq.	*p*-value
2	Mentions of being nice/altruistic	0.010	0.005	0.056
3	Mentions of money (generally and with specific goals)	0.065	0.046	0.060
3B	Indicates goal of making everyone better off	0.019	0.007	0.019
6	Mentions gender	0.007	0.002	0.020
9	Agrees to previous message in the group	0.351	0.414	0.000
Ave lines	Average number of lines of chat within each chat room	3.22	3.08	0.493

Note: Two-tailed p-values based on Poisson count regressions, with standard errors clustered on individual subjects.

5. Conclusion

Many of the most important decisions in organizations are made by groups. Coordination and communication are often critical for groups to make effective decisions. In this paper we focus on decisions that can affect group members as well as a passive external party. We examine if the gender composition of three-person groups affects choices and beliefs in a Coordination game with selfish and prosocial equilibria and we allow for opportunities to communicate. We also investigate individuals' prosocial preferences independent of their beliefs about others' choices, using a direct Allocation task.

We find that while there is no gender difference in the Allocation task, in the Coordination game, controlling for the social preferences of group members, groups with more women are significantly more likely to choose the prosocial equilibrium. This is observed in both uniform gender and mixed gender groups.[28] Moreover, both men and women strongly believe that women will make choices that are kinder to external parties, in line with gender stereotypes.

[28] This finding, in conjunction with the results from the literature (Fearon and Humphreys, 2018; Greig and Bohnet, 2009), provides credence to policies directing development funds in poor communities through women's groups.

Our findings suggest that the gender composition of the group and the beliefs that individuals have about others in the group can have important impacts on the outcomes not just for group members but also for others in society. Policies that encourage gender diversity in organizations could therefore have additional (and in some cases unintended or unplanned) impacts on social outcomes and inequality. These spillovers of gender diversity policies should be considered when advocating for or against such policies and when evaluating their effectiveness.

Another interesting finding that emerges from our study is that the number of Kind types in the group is a stronger predictor of Kind choices at the group level in the Coordination game, compared to the gender composition of the group. This indicates that firms seeking to promote corporate and environmental social responsibility initiatives, for example, may benefit from attracting decision makers with prosocial preferences. A challenge they may face is that prosocial preferences are not (fully) observable during the hiring process. However, due to the belief that women are more prosocial than men when coordinating on group outcomes, that men are more prosocial in the presence of women in the group, and that women are more malleable depending on the context, firms may see greater success if women participate in more decision-making roles. As beliefs about differences in prosociality across genders are exaggerated in our data, depending on the organization's objectives, it may be useful to consider other group characteristics besides gender. For instance, cognitive diversity, which captures differences in perspective or information processing styles in groups (Reynolds and Lewis, 2017), has been recently identified as key to solving new and complex problems.

References:

Adams, R.B. and D. Ferreira. 2009. "Women in the boardroom and their impact on governance and performance." *Journal of Financial Economics* 94, 291-309.

Aguiar, F., P. Branas-Garza, R Cobo-Reyes, N. Jimenez and L.M. Miller. 2009. "Are women expected to be more generous?" *Experimental Economics* 12, 93-98.

Ahern, K. and A. Dittmar. 2012. "The changing of the boards: The impact on firm valuation of mandated female board representation." *Quarterly Journal of Economics* 127(1), 137-197.

Andersson, O., H.J. Holm, J.R. Tyran, and E. Wengström. 2014. "Deciding for others reduces loss aversion." *Management Science* 62(1), 29-36.

Apesteguia, J., G. Azmat, and N. Iriberri. 2012. "The impact of gender composition on team performance and decision making: Evidence from the field." *Management Science* 58(1), 78–93.

Armstrong, S. 2002. "Corporate boards should represent a broader community of interests." *Knowledge@Wharton* (http://knowledge.wharton.upenn.edu/article/corporate-boards-should-represent-a-broader-community-of-interests/, accessed 22 March 2018).

Azmat, G. and B. Petrongolo. 2014. "Gender and the labor market: What have we learned from field and lab experiments?" *Labour Economics* 30, 32-40.

Babcock, L., M. Recalde, L. Vesterlund, and L. Weingart. 2017. "Gender differences in accepting and receiving requests for tasks with low promotability." *American Economic Review* 107(3), 714-747.

Bagues, M., M. Sylos-Labini, and N. Zinovyeva. 2017. "Does the gender composition of scientific committees matter?" *American Economic Review* 107(4), 1207-1238.

Baranski, A., D. Geraldes, A. Kovaliukaite, and J. Termewan. 2021. "An experiment on gender representation in majoritarian bargaining." Working Paper 20210060, NYU-Abu Dhabi.

Ben-Ner, A., L. Putterman, and T. Ren. 2011. "Lavish returns on cheap talk: Two-way communication in trust games." *The Journal of Socio-Economics* 40, 1-13.

Bilén, D., A. Dreber, and M. Johannesson. 2021. "Are women more generous than men? A meta-analysis." *Journal of the Economic Science Association* 7, 1-18.

Blanco, E., T. Haller, and J.M. Walker. 2018. "Provision of environmental public goods: Unconditional and conditional donations from outsiders." *Journal of Environmental Economics and Management* 92, 815-831.

Bland, J. and N. Nikiforakis. 2015. "Coordination with third-party externalities." *European Economic Review* 80, 1-15.

Blume, A. and A. Ortmann. 2007. "The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria." *Journal of Economic Theory* 132(1), 274-290.

Bordalo, P., K. Coffman, N. Gennaioli and A. Shleifer. 2016. "Stereotypes." *Quarterly Journal of Economics* 131, 1753-1794.

Bordalo, P., K. Coffman, N. Gennaioli and A. Shleifer. 2019. "Beliefs about gender." *American Economic Review* 109, 739-773.

Born, A., E. Ranehill, and A. Sandberg. 2018. "A man's world? – The impact of a male dominated environment on female leadership." Working paper, Stockholm University. Available at SSRN: https://ssrn.com/abstract=3207198.

Brañas-Garza, P., V. Capraro, and E. Rascon-Ramirez. 2018. "Gender differences in altruism on Mechanical Turk: Expectations and actual behavior." *Economics Letters* 170, 19-23.

Brandts, J. and D.J. Cooper. 2006. "Observability and overcoming coordination failure in organizations: An experimental study." *Experimental Economics 9*(4), 407-423.

Brandts, J., D.J. Cooper and R.A. Weber. 2015. "Legitimacy, communication and leadership in the turnaround game." *Management Science* 61, 2627-2645.

Cason, T., R. Sheremeta, and J. Zhang. 2012. "Communication and efficiency in competitive coordination games." *Games and Economic Behavior* 76 (1), 26-43.

Charness, G., and M. Dufwenberg. 2006. "Promises and partnership." *Econometrica* 74, 1579-1601.

Charness, G. and M.O. Jackson. 2009. "The role of responsibility in strategic risk-taking." *Journal of Economic Behavior and Organization 69*(3), 241-247.

Charness, G. and M. Sutter. 2012. "Groups make better self-interested decisions." *Journal of Economic Perspectives* 26, 157-176.

Cohen, J. 1960. "A coefficient of agreement for nominal scales". *Educational and Psychological Measurement* 20(1), 37–46.

Cooper, D. and R. Weber. 2020. "Recent advances in experimental coordination games." In: *Handbook of Experimental Game Theory*, C.M. Capra, R. Croson, M. Rigdon and T. Rosenblat (eds). Edward Elgar, 149-183.

Cooper, R., D.V. DeJong, R. Forsythe, and T.W. Ross. 1989. "Communication in the battle of the sexes game: Some experimental results." *RAND Journal of Economics* 20(4), 568-587.

Cooper, R., D.V. DeJong, R. Forsythe, and T.W. Ross. 1992. "Communication in coordination games." *The Quarterly Journal of Economics 107*(2), 739-771.

Croson, R. and U. Gneezy. 2009. "Gender differences in preferences." *Journal of Economic Literature* 47(2), 448–474.

CSRI, 2016. "The CS Gender 3000: The reward for change." Credit Suisse Research Institute, edited by Urs Rohner, Tidjane Thiam and Iris Bohnet.

Dawes, R.M., J. McTavish, and H. Shaklee. 1977. "Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation." *Journal of Personality and Social Psychology* 35, 1-11.

Delaney, J. and S. Jacobson. 2014. "Those outsiders: How downstream externalities affect public good provision." *Journal of Environmental Economics and Management 67*(3), 340-352.

Deschamps, P. 2018. "Gender quotas in hiring committees: A boon or a bane for women." *Sciences Po LIEPP Working Paper*, *82*.

Deutsch, M. 1960. "The effect of motivational orientation upon trust and suspicion." *Human Relations* 13, 123-139.

Devetag, G. and A. Ortmann. 2007. "When and why? A critical survey on coordination failure in the laboratory." *Experimental Economics* 10(3), 331-344.

Di Girolamo, A. and M. Drouvelis. 2015. "The role of gender composition and size of the group in a minimum effort game." *Economics Letters 137*, 168-170.

Dufwenberg, M. and U. Gneezy. 2005. "Gender and coordination." *Experimental Business Research*, 253-262.

Dufwenberg, M. and A. Muren. 2006. "Gender composition in teams." *Journal of Economic Behavior and Organization* 61(1), 50-54.

Eagly, A.H. and V.J. Steffen. 1984. "Gender stereotypes stem from the distribution of women and men into social roles." *Journal of Personality and Social Psychology* 46(4), 735-754.

Eagly, A.H. and W. Wood. 1999. "The origins of sex differences in human behavior: Evolved dispositions versus social roles." *American Psychologist* 54, 408–423.

Eckbo, B.E., K. Nygaard, and K.S. Thorburn. 2014. "Did Norway's board gender quota reduce firm value?" Working paper, Tuck School of Business at Dartmouth.

Eckel, C.C. and P.J. Grossman. 2008a. "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior & Organization* 68(1), 1-17.

Eckel, C.C. and P.J. Grossman. 2008b. "Men, women and risk aversion: Experimental evidence." *Handbook of Experimental Economics Results* 1, 1061-1073.

Engel, C. and B. Rockenbach. 2011. "We are not alone: the impact of externalities on public good provision." SSRN eLibrary.

Farrell, J. and M. Rabin. 1996. "Cheap talk." *The Journal of Economic Perspectives 10*(3), 103-118.

Fearon, J. D. and M. Humphreys. 2018. "Why do women co-operate more in women's groups?" In: *Towards Gender Equity in Development*, S. Anderson, L. Beaman, and J.-P. Platteau (eds.). Oxford Scholarship Online.

Fehr, E. and K.M. Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics* 114(3), 817-868.

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

Fréchette, G. 2012. "Session-effects in the laboratory." *Experimental Economics*, 15(3), 485-498.

Freidenvall L. and D. Dahlerup. 2013. Electoral gender quota systems and their implementation in Europe. European Parliament, Directorate General for Internal Policies, Policy Department C: Citizens' Rights and Constitutional Affairs.

Greig, F. and I. Bohnet. 2009. "Exploring gendered behavior in the field with experiments: Why public goods are provided by women in a Nairobi slum." *Journal of Economic Behavior and Organization* 70(1), 1–9

Greiner, B. 2015. "Subject pool recruitment procedures: organizing experiments with ORSEE." *Journal of the Economic Science Association* 1(1), 114-125.

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

Holm, H.J. 2000. "Gender-based focal points." *Games and Economic Behavior* 32, 292-314.

Hoogendoorn, S., H. Oosterbeek, and M. van Praag. 2013. "The impact of gender diversity on the performance of business teams: Evidence from a field experiment." *Management Science* 59(7), 1514-1528.

Humphrey, S.J. and E. Renner. 2011. 'The social cost of responsibility.' CeDEx discussion paper 2011-02.

Isaac, R.M. and J.M. Walker. 1988. "Communication and free-riding behavior: The voluntary contribution mechanism." *Economic Inquiry* 26, 585-608.

Ivanova-Stenzel, R. and D. Kübler. 2011. "Gender differences in team work and team competition." *Journal of Economic Psychology* 32(5), 797–808.

Jack, B. K. and M.P. Recalde. 2015. "Leadership and the voluntary provision of public goods: Field evidence from Bolivia." *Journal of Public Economics* 122, 80-93.

Jerdee, T.H. and B. Rosen. 1974. "Effects of opportunity to communicate and visibility of individual decisions on behavior in the common interest." *Journal of Applied Psychology* 59, 712-716.

Keck, S. and W. Tang. 2017. "Gender composition and group confidence judgment: The perils of all-male groups." *Management Science*, 1-22

Kim, D and L.T. Starks. 2016. "Gender diversity on corporate boards: Do women contribute unique skills?" *American Economic Review: Papers and Proceedings* 106(5), 267-271.

Koukoumelis, A., M.V. Levati, and J. Weisser. 2012. "Leading by words: A voluntary contribution experiment with one-way communication." *Journal of Economic Behavior & Organization* 8, 379-390.

Krippendorff, K. (2004). *Content Analysis: An Introduction to its Methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Kugler, T., E. Kausel, and M. Kocher. 2012. "Are groups more rational than individuals? A review of interactive decision making in groups." *Wiley Interdisciplinary Reviews: Cognitive Science* 3, 471-482.

Landis, J.R. and G.G. Koch. 1977. "The measurement of observer agreement for categorical data." *Biometrics* 33, 159-174.

Loomis, J.L. 1959. "Communication, the development of trust, and cooperative behavior." *Human Relations* 12, 305-315.

Matsa, D. and A. Miller. 2013. "A female style in corporate leadership? Evidence from quotas." *American Economic Journal: Applied Economics* 5(3), 136-169.

McDonald, I.M., N. Nikiforakis, N. Olekalns and H. Sibly. 2013. "Social comparisons and reference group formation: Some experimental evidence." *Games and Economic Behavior* 79, 75-89.

Mengel, F. 2021. "Gender bias in opinion aggregation." *International Economic Review* 62, 1055-1080.

Niederle, M. and L. Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *Quarterly Journal of Economics* 122(3), 1067-1101.

Oprea, R., G. Charness, and D. Friedman. 2014. "Continuous time and communication in a public-goods experiment." *Journal of Economic Behavior and Organization* 108, 212-223.

Palfrey, T.R. and H. Rosenthal. 1991. "Testing for effects of cheap talk in a public goods game with private information." *Games and Economic Behavior* 3, 183-220.

Pogrebna, G., D.H. Krantz, C. Schade, and C. Keser. 2011. "Words versus actions as a means to influence cooperation in social dilemma situations." *Theory and Decision* 71, 473-502.

Post, C. and K. Byron. 2015. "Women on boards and firm financial performance: A meta-analysis." *Academy of Management Journal* 58, 1546-1571.

Reynolds, A and D. Lewis. 2017. "Teams solve problems faster when they're more cognitively diverse" *Harvard Business Review*, 30 March.

Swensson, R.G. 1967. "Cooperation in the prisoner's dilemma game I: The effects of asymmetric payoff information and explicit communication." *Behavioral Science* 12, 314-322.

Vogel, D.L., S R. Wester, M. Heesacker, and S. Madon. 2003. "Confirming gender stereotypes: A social role perspective." *Sex Roles* 48, 519-528.

Williams, M. and E. Polman. 2015. "Is it me or her? How gender composition evokes interpersonally sensitive behavior on collaborative cross-boundary projects." *Organization Science* 26(2), 334-355.

Woolley, A.W., C. F. Chabris, A. Pentland, N. Hashmi and T.W. Malone. 2010. "Evidence for a collective intelligence factor in the performance of human groups." *Science* 330, 686-688.

Woolston, C. 2019. "How a hiring quota failed." *Nature, Career News*, 566, 287.

Zizzo, D.J. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics 13*(1), 75-98.